基于图像像素划分的主动隐写分析

刘 静* 汤光明
(解放军信息工程大学电子技术学院 郑州 450004)

摘 要：隐写分析的研究主要集中于隐写检测，而对主动隐写分析研究的较少。该文从隐写检测角度出发，将图像像素划分为不同类点，通过对信息嵌入、最低位平面置反带来的各类点频次变化的分析，提出一种针对空域图像 LSB (Least Significant Bit) 替换隐写术的主动隐写分析方法，解决了主动隐写分析中的密钥恢复问题。所提出的方法物理意义直观，实现简单。实验结果表明，该方法在一定嵌入率范围内均可成功恢复隐写密钥。

关键词：隐写；隐写分析；像素划分；最低位比特(LSB)替换；密钥恢复
中图分类号：TP309 **文献标识码**：A **文章编号**：1009-5896(2012)08-1928-06
DOI: 10.3724/SP.J.1146.2011.01422

Active Steganalysis Based on Pixels Classification in the Image

Liu Jing Tang Guang-ming
(*Institute of Electronic Technology*, *Information Engineering University*, *Zhengzhou* 450004, *China*)

Abstract: The research on steganalysis has mainly focused on hidden information detection, and there are few methods about active steganalysis. From the view of hidden information detection, the pixels are classified into different kinds. Based on the analysis of the effects on the frequencies of different kinds of pixels by message embedding and Least Significant Bit (LSB) plane flipping, an active steganalysis approach is proposed to recover the stego key of LSB replacement steganography in spatial domain of images. This method has remarkable physical significance and can be implemented conveniently. Experimental results show that this method can recover the stego key successfully in certain range of embedding ratio.

Key words: Steganography; Steganalysis; Pixel classification; Least Significant Bit (LSB) replacement; Stego key recovery

1 引言

20 世纪 90 年代以来，信息隐藏作为信息安全中的重要课题引起国际学术界的重视。隐写术作为信息隐藏的一个重要分支，由于它可以隐藏信息通信的事实，具有密码学所不具有的优点，在近十年来得到快速发展。隐写分析技术是隐写术的对抗技术，它一方面可以监控隐写术的非法滥用，另一方面还可促进安全性更高的隐写算法的诞生。

隐写分析可分为被动隐写分析和主动隐写分析两种[1]。前者主要指检测秘密信息的存在性[2,3]，后者在前者基础上，识别所用的隐写工具[4]、估计嵌入的秘密信息量[5]、恢复隐写密钥[6]，最终提取隐藏的秘密信息。目前隐写分析领域的工作主要集中于被动隐写分析研究，对主动隐写分析的研究相对较少。本文主要研究主动隐写分析中的隐写密钥恢复问题，该问题对数字取证等领域具有重要意义，同时该问题也是隐写分析领域中的难点问题。隐写密钥恢复不同于传统的密码分析。传统的密码分析处理的数据是密文序列，而密文序列中包含了密钥的信息，因此可以通过分析密文序列与加密算法恢复加密密钥。但隐写密钥恢复处理的是载密对象，载密对象与隐写密钥的联系一般并不像密文与加密密钥那样直接，如随机最低位比特(LSB)替换隐写算法是以隐写密钥作为随机数发生器的种子，生成一列随机位置，然后将消息嵌入到这些位置对应像素的 LSB 位生成载密图像，因此很难找到载密对象与隐写密钥之间可以利用的关系。从这个角度说，隐写密钥恢复比传统密码分析面对的问题更为困难。

过去隐写密钥恢复均通过在提取的数据中寻找可识别的模式来确定提取数据的隐写密钥是否为真密钥[7]，若信息经过加密，还得对每个隐写密钥提取出的序列再尝试所有的解密密钥，代价很高。

2011-12-31 收到，2012-05-02 改回
国家自然科学基金(61101112)，中国博士后基金(2011M500775)和河南省信息安全重点实验室基金(10ISLB001)资助课题
*通信作者：刘静 kimi_liujing@163.com

Fridrich 等人[6,8]根据密钥产生的嵌入路径上的数据特性进行穷尽搜索，证明了搜索密钥的复杂度仅与隐写密钥空间有关，与解密密钥空间无关。他们利用卡方检验对 JPEG 图像和空域图像的随机 LSB 隐写算法，给出了恢复隐写密钥的方法，但是该方法需要一定的检验参数。文献[9]基于密码学中的碰撞攻击方法，将密钥恢复问题转化为序列密码分析问题，完成了空域图像随机 LSB 替换隐写术的密钥恢复。由于密钥恢复的困难性，目前这一方面的研究成果仍非常少，近几年仍未见新的密钥恢复方法报导。

本文认为隐写密钥恢复问题本质上是一种特殊的隐写检测。因为对于一个载密对象，伪密钥产生的嵌入路径上的数据与载密对象具有同样的统计特性，而真密钥产生的嵌入路径上的数据与满嵌载密对象具有同样的统计特性。因此，若能区分满嵌载密对象与其他嵌入率的载密对象，也一定能区分真伪密钥。基于此，本文针对空域图像随机 LSB 替换隐写提出一种密钥恢复方法：首先根据像素与其邻域像素之差将像素划分为 H 点，L 点和 M 点；然后分析隐写操作与载密图像最低位置反操作对各类点频次的影响，得到检测满嵌载密图像的依据；最后将此依据映射到真伪密钥产生的嵌入路径上的数据中，从而恢复隐写密钥。

为叙述方便，对文中的一些名词术语进行如下约定：真(伪)路径指真(伪)隐写密钥产生的嵌入路径；嵌入率＝载体样本中携带的消息比特数/载体样本数，如图像嵌入率＝图像中携带的消息比特数/图像大小，路径上的嵌入率＝路径上携带的消息比特数/路径上元素个数。

2　图像像素划分

设一 $N_1 \times N_2$ 大小的图像 I 在位置 (i,j) 的像素灰度值为 $I(i,j)$，其中 $I(i,j) \in \{0,1,\cdots,255\}$，$i = 1,2,\cdots,N_1, j = 1,2,\cdots,N_2$。对图像非边界像素点，存在一个八邻域像素灰度值集合(顺时针方向)：$\{I(i-1,j-1), I(i-1,j), \cdots, I(i+1,j+1)\}$，则中心像素点 $I(i,j)$ 与其八个相邻像素点灰度值之差分别为 $I(i-1,j-1)-I(i,j), I(i-1,j)-I(i,j), \cdots, I(i+1,j+1)-I(i,j)$，记为向量形式 $\boldsymbol{D} = (D_1(i,j), D_2(i,j), \cdots, D_8(i,j))$。对于图像像素矩阵 4 个顶点上的像素，取其周围 3 个像素点作为其邻域像素计算 $\boldsymbol{D}$，而 4 条边上的像素，取其周围相邻 5 个像素点作为其邻域像素计算 $\boldsymbol{D}$。根据 $\boldsymbol{D}$ 中元素 $D_k, k \in \{1,2,\cdots,8\}$ 的大小以及中心像素点的奇偶性，可将图像像素点划分如下：

(1)若 $\max D_k < 0$，则称中心像素点为 H 点。在 H 点中，若中心像素点灰度值为偶数，则称中心像素点为 H^E 点；若中心像素点灰度值为奇数，则称中心像素点为 H^O 点。

(2)若 $\min D_k > 0$，则称中心像素点为 L 点。同理，L 点可根据中心像素点灰度值的奇偶性分为 L^E 点和 L^O 点。

(3)若 $\min D_k \le 0$ 且 $\max D_k \ge 0$，则称中心像素点为 M 点。同样，M 点包括 M^E 点和 M^O 点。

3　基于像素划分的密钥恢复

3.1　空域随机 LSB 替换隐写模型

在利用随机 LSB 替换隐写嵌入秘密消息时，首先根据待嵌入的消息 $m = \{m_1, m_2, \cdots, m_\ell\}, m_d \in \{0,1\}, 1 \le d \le \ell$，选择一幅自然图像 I，它包含 $N_1 \times N_2$ 个像素点，且 $N_1 \times N_2 \ge \ell$。然后，输入密钥 k_0，在 k_0 的控制下，伪随机数发生器(PRNG)产生嵌入路径 $\mathrm{Path}(k_0) = \{(i_1,j_1),(i_2,j_2),\cdots,(i_\ell,j_\ell)\}$。其中 $1 \le i_d \le N_1$，$1 \le j_d \le N_2$，$1 \le d \le \ell$。最后，选择修改 $I(i_d,j_d)$ 的最后一位以携带消息比特 m_d，具体修改方式见表 1。即如果 m_d 与 $I(i_d,j_d)$ 的最后一位相同，就不改变 $I(i_d,j_d)$；反之就改变 $I(i_d,j_d)$ 最后一位。待 ℓ 长消息完全嵌入到 I 中，得到载密图像 $\tilde{I}$。

表 1 LSB 替换隐写方式

原始像素灰度值	$2x$		$2x+1$	
消息比特	0	1	0	1
隐写后像素灰度值	$2x$	$2x+1$	$2x$	$2x+1$

本文的主要任务就是在知道隐写算法的全部细节，并利用已有的方法估计出图像嵌入率的条件下，从载密图像 $\tilde{I}$ 中恢复出隐写密钥 k_0。

3.2　满嵌载密图像检测依据

图像经过隐写后，各类点转化情况如图 1 所示。

从图 1 可以看出，H 点与 L 点之间是独立的，LSB 隐写共导致的 16 类转化主要发生在 H 点内，L

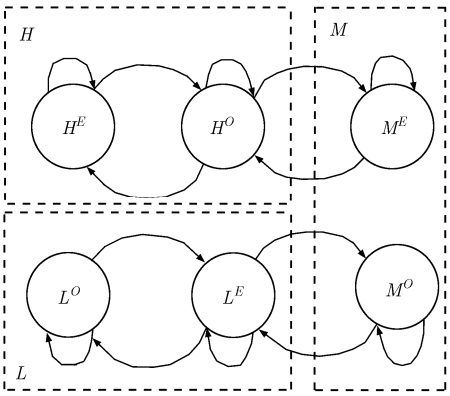


图 1　隐写后各类点的转化情况

点内，M 点与 H 点之间，M 点与 L 点之间。但只有其中的 4 类转化会导致图像中 H 点，L 点和 M 点的频次发生变化。这 4 类转化分别为

$$\left.\begin{aligned} &H^O(\max D_k=-1)\Rightarrow M^E(\max D_k=0)\\ &M^E(\max D_k=0)\Rightarrow H^O(\max D_k=-1)\\ &L^E(\min D_k=1)\Rightarrow M^O(\min D_k=0)\\ &M^O(\min D_k=0)\Rightarrow L^E(\min D_k=1)\end{aligned}\right\} \tag{1}$$

其中 $X(y)$表示满足条件 y 的 X 类点。为了方便描述，将这 4 类点按表 2 重新命名，(HM 表示 H 点中可能转化为 M 点的像素集合，其它可类推)。

表 2 集合的简单表示

符号	点集合
HM	$H^O(\max D_k=-1)$
LM	$L^E(\min D_k=1)$
MH	$M^E(\max D_k=0)$
ML	$M^O(\min D_k=0)$

由于 H 点，L 点和 M 点的频次总数为图像像素个数。因此，从 H, L, MH 和 ML 点出发，分析隐写对 H 点，L 点和 M 点频次的影响。

令 H 点，L 点，MH 点和 ML 点的频次总数记为 F, F_X 表示 X 点频次，用频率 $P_X=F_X/F$ 表示 X 点出现的概率，α 为所嵌入秘密消息中“1”的频率，β 为“0”的频率，$\alpha+\beta=1$，消息在嵌入前通常会加密，因此，$\alpha\approx\beta=1/2$，λ 为图像嵌入率。设载密图像中 X 点的频次为 $\widetilde{F}_X$，概率值为 $\widetilde{P}_X=\widetilde{F}_X/F$，则有如下定理：

定理 1 载密图像中，对 $\forall\lambda\in[0,1]$，当 $\alpha\approx\beta$ 时，H 点和 L 点的频次以 λ 为自变量增加，M 点的频次以 λ 为自变量减少。

证明 以 H 点为例：

$$\begin{aligned}\widetilde{P}_H=\widetilde{F}_H/F&=P_{H^E}+(1-\beta\lambda)P_{HM}+P_{H^O(\max D_k<-1)}\\&\quad+\alpha\lambda P_{MH}=P_H+\lambda(\alpha P_{MH}-\beta P_{HM})\end{aligned}\tag{2}$$

又因为 $\alpha\approx\beta$，由文献[10]知 $P_{MH}>P_{HM}$，所以，

$$\alpha P_{MH}-\beta P_{HM}>0 \tag{3}$$

令 $A=\alpha P_{MH}-\beta P_{HM}>0$，从而 $\widetilde{F}_H=(A\lambda+P_H)F$，即载密图像中 H 点的频次以 λ 为自变量增加。

同理可证 $\widetilde{F}_L=(B\lambda+P_L)F$，$B=\beta P_{ML}-\alpha P_{LM}>0$，即载密图像中 L 点的频次也以 λ 为自变量增加。

由于图像中 H 点，L 点和 M 点频次总数一定，为图像大小，记为 E，则

$$E=F_H+F_L+F_M=\widetilde{F}_H+\widetilde{F}_L+\widetilde{F}_M \tag{4}$$

因此：

$$\begin{aligned}\widetilde{F}_M&=E-(\widetilde{F}_H+\widetilde{F}_L)=-(A+B)F\lambda\\&\quad+(E-P_HF-P_LF)\end{aligned}\tag{5}$$

因为 $-(A+B)F<0$，所以载密图像中 M 点的频次以 λ 为自变量减少。 证毕

当图像嵌入率小于 1 时，真路径上的嵌入率为 1，伪路径上的嵌入率小于 1。从以上证明过程可以推出，载密图像中，真路径上 H 点，L 点的频次大于伪路径上相应点的频次，真路径上 M 点的频次小于伪路径上 M 点的频次。然而，根据 $\widetilde{P}_H=P_H+\lambda(\alpha P_{MH}-\beta P_{HM})$ 可知，载密图像中各嵌入路径上 H 点的频次依赖于载体图像中该路径上 H 点，MH 点和 HM 点的频次，由于载体图像中不同路径上同类点的频次并不一致，因此定理 1 不能映射为区分真伪密钥的依据，需要对载密图像作进一步处理。

对载密图像中单个图像像素定义一最低位置反操作：

$$\text{Flipping}: 0\leftrightarrow 1, 2\leftrightarrow 3, \cdots, 254\leftrightarrow 255 \tag{6}$$

设对载密图像中像素进行置反操作后，图像中 X 点的概率值为 $\widehat{P}_X$，频次为 $\widehat{F}_X$。则可得如下定理：

定理 2 载密图像中，对 $\forall\lambda\in[0,1]$，当 $\alpha\approx\beta$ 时，对像素进行置反操作后，H 点和 L 点的频次以 λ 为自变量减少，M 点的频次以 λ 为自变量增加。

证明 仍以 H 点为例进行证明，置反后有

$$\widehat{P}_H=\widetilde{P}_H-\widetilde{P}_{HM}+\widetilde{P}_{MH} \tag{7}$$

根据定理 1，得

$$\left.\begin{aligned}\widetilde{P}_H&=P_H+\lambda(\alpha P_{MH}-\beta P_{HM})\\ \widetilde{P}_{HM}&=\alpha\lambda P_{MH}+(1-\beta\lambda)P_{HM}\\ \widetilde{P}_{MH}&=\beta\lambda P_{HM}+(1-\alpha\lambda)P_{MH}\end{aligned}\right\} \tag{8}$$

因此，有

$$\begin{aligned}\widehat{P}_H&=P_H+\lambda(\alpha P_{MH}-\beta P_{HM})-\alpha\lambda P_{MH}\\&\quad-(1-\beta\lambda)P_{HM}+\beta\lambda P_{HM}+(1-\alpha\lambda)P_{MH}\\&=P_H+P_{MH}-P_{HM}+\beta\lambda P_{HM}-\alpha\lambda P_{MH}\end{aligned}\tag{9}$$

因为 $\alpha\approx\beta=1/2$，所以，

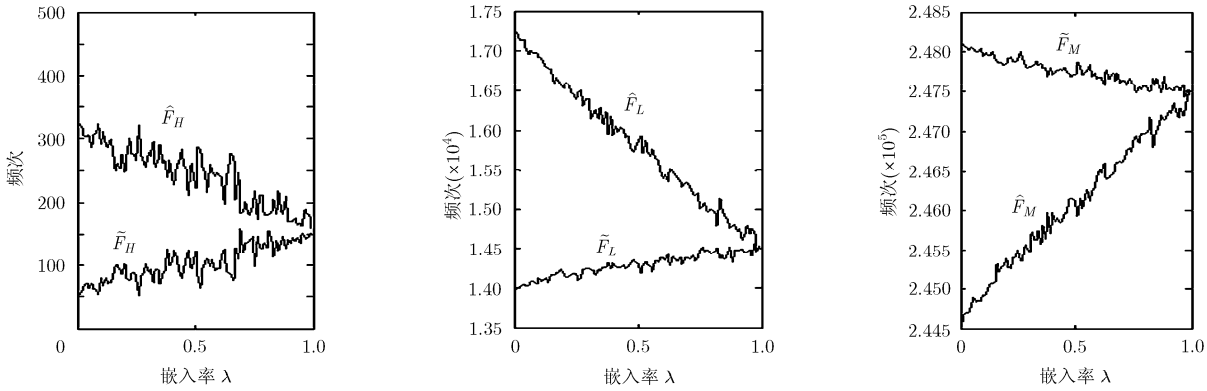
$$\widehat{P}_H=-\lambda(P_{MH}-P_{HM})/2+P_H+P_{MH}-P_{HM} \tag{10}$$

令 $A=(P_{MH}-P_{HM})/2>0, B=P_H+P_{MH}-P_{HM}$，则置反后 H 点的频次可表示为

$$\widehat{F}_H=\widehat{P}_HF=-AF\lambda+BF,\ AF>0 \tag{11}$$

即对载密图像像素进行置反操作后，H 点频次以 λ 为自变量减少，同理可得 L 点频次也以 λ 为自变量减少，M 点频次以 λ 为自变量增加。 证毕

图 2 为 512×512 大小的 lena 图像，经过隐写



(a) H点　(b) L点　(c) M点

图 2 H 点，L 点，M 点频次随 λ 的变化情况

和置反操作后，H 点，L 点，M 点频次随嵌入率 λ 的变化情况。

从图 2 中可以看出，进行置反操作之前，H 点，L 点，M 点频次随 λ 的变化情况符合定理 1 的描述；置反操作后，H 点，L 点，M 点频次随 λ 的变化情况符合定理 2 的描述；还可看出 $\widetilde{F}_H-\widehat{F}_H$，$\widetilde{F}_L-\widehat{F}_L$ 随 λ 的增加而增加，$\widetilde{F}_M-\widehat{F}_M$ 随 λ 的增加而减少。当嵌入率为 1 时，即对满嵌载密图像，有 $\widetilde{F}_H-\widehat{F}_H\approx 0$，$\widetilde{F}_L-\widehat{F}_L\approx 0$，$\widetilde{F}_M-\widehat{F}_M\approx 0$。据此，可得到不依赖于载体图像各类点频次的真伪密钥区分依据。

3.3 密钥恢复原理

设 $\widetilde{P}_X(k)$，$\widehat{P}_X(k)$ 分别为置反操作前后，载密图像中密钥 k 产生的嵌入路径上 X 点出现的概率。$F(k)$ 为密钥 k 产生的嵌入路径上的元素个数，也即消息长度。$S_X(k)$ 表示载密图像中密钥 k 产生的嵌入路径上 X 点频次与置反操作后 X 点频次的差值，即

$$S_X(k)=\left(\widetilde{P}_X(k)-\widehat{P}_X(k)\right)F(k) \tag{12}$$

则有如下定理：

定理 3　载密图像中，对 $\forall\lambda\in(0,1)$，当 $\alpha\approx\beta$ 时，$S_H(k_0)>S_H(k_j)$，$S_L(k_0)>S_L(k_j)$，$S_M(k_0)<S_M(k_j)$。其中 k_0 为正确密钥，k_j 为伪密钥。

证明　因为，对 H 点：

$$S_H(k)=\left(\widetilde{P}_H(k)-\widehat{P}_H(k)\right)F(k) \tag{13}$$

而

$$\widehat{P}_H(k)=\widetilde{P}_H(k)-\widetilde{P}_{HM}(k)+\widetilde{P}_{MH}(k) \tag{14}$$

所以有

$$\begin{aligned}\widetilde{P}_H(k)-\widehat{P}_H(k)&=\widetilde{P}_{HM}(k)-\widetilde{P}_{MH}(k)\\&=\alpha\lambda P_{MH}(k)+(1-\beta\lambda)P_{HM}(k)\\&\quad-\beta\lambda P_{HM}(k)-(1-\alpha\lambda)P_{MH}(k)\end{aligned} \tag{15}$$

又因为 $\alpha\approx\beta=1/2$，因此，

$$\widetilde{P}_H(k)-\widehat{P}_H(k)=(\lambda-1)\left(P_{MH}(k)-P_{HM}(k)\right) \tag{16}$$

对于真密钥 k_0，其产生的嵌入路径上的嵌入率 $\lambda_0=1$，有

$$S_H(k_0)=(\lambda_0-1)\left(P_{MH}(k_0)-P_{HM}(k_0)\right)F(k_0)=0 \tag{17}$$

对于伪密钥 k_j，其产生的嵌入路径上的嵌入率 $\lambda_j<1$，因此，

$$S_H(k_j)=(\lambda_j-1)\left(P_{MH}(k_j)-P_{HM}(k_j)\right)F(k_j)<0 \tag{18}$$

同理有

$$S_L(k_0)=0,S_L(k_j)<0,S_M(k_0)=0,S_M(k_j)>0 \tag{19}$$

从而有

$$\begin{aligned}&S_H(k_0)>S_H(k_j),\ S_L(k_0)>S_L(k_j),\\&S_M(k_0)<S_M(k_j)\end{aligned} \tag{20}$$

证毕

在定理 3 的基础上，可设计针对空域图像随机 LSB 替换隐写的密钥恢复算法。由于 M 点的变化可反映 H 点和 L 点的全部变化情况，为了节省试验单个密钥的时间，只选择不等式 $S_M(k_0)<S_M(k_j)$ 作为区分真伪密钥的依据。因此，最小的 $S_M(k)$ 所对应的密钥即为真隐写密钥。

3.4 密钥恢复算法

假设已获得一幅经过空域随机 LSB 替换隐写的载密图像 $\tilde{I}$，它大小为 $N_1\times N_2$，密钥空间为 K。则密钥恢复的具体过程如下：

步骤 1　标记 $\tilde{I}$ 中的元素 $\tilde{I}(i,j),i=1,2,\cdots,N_1$，$j=1,2,\cdots,N_2$，若 $\tilde{I}(i,j)$ 为 M 点，则记 $W_1(i,j)=1$，否则记 $W_1(i,j)=0$，得到 0-1 矩阵 $\boldsymbol{W}_1$；

步骤 2　对 $\tilde{I}(i,j)$ 进行置反操作得 $\hat{I}(i,j)$，若 $\hat{I}(i,j)$ 为 M 点，则记 $W_2(i,j)=1$，否则记 $W_2(i,j)=0$，得 0-1 矩阵 $\boldsymbol{W}_2$；

步骤 3　计算 $\boldsymbol{W}=\boldsymbol{W}_1-\boldsymbol{W}_2$；

步骤 4 利用文献[6]中方法估计消息嵌入率 λ，并计算消息长度 $\ell = \lambda N_1 N_2$；

步骤 5 穷举密钥空间 K 中的密钥，对每个密钥 k，以 k 为种子利用隐写算法的随机数发生器生成路径 $\mathrm{Path}(k) = \{(i_1, j_1), (i_2, j_2), \cdots, (i_\ell, j_\ell)\}$，计算 $S_M(k) = \sum_{n=1}^{\ell} W(i_n, j_n)$，并记 $\boldsymbol{S} = \{S_M(k_i), k_i \in K\}$；

步骤 6 取 $S_{\min} = \min(\boldsymbol{S})$，记录 $\boldsymbol{T} = \{k | k \in K$ 且 $S_M(k) = S_{\min}\}$；

步骤 7 若 $|\boldsymbol{T}| = 1$，则 $\boldsymbol{T}$ 中的密钥为真密钥，结束；若 $|\boldsymbol{T}| > 1$，则算法未搜索到真密钥，结束。

4 实验结果与分析

为了验证算法的有效性，从 USC-SIPI[11]图像库中选择 30 幅载体图像，裁剪、转化为 512×512 大小的灰度图像。利用随机 LSB 替换隐写算法分别嵌入不同长度的秘密信息，嵌入的信息从 Advanced Encryption Standard(AES)加密的密文中截取。算法所用的随机数发生器是在 MATLAB7.0 中实现的，密钥长度为 16 bit。然后在配置为 Pentium® 4, 3.0 GHz, MEM-512M 的 PC 机上用本文的密钥恢复方法进行密钥搜索。图 3 为 lena 图像在嵌入率为 0.61 时的 $S_M(k)$ 值，图中箭头所指极小值对应的密钥为嵌入秘密信息时使用的隐写密钥。

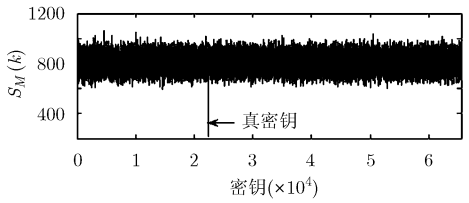


图 3 lena 图像 $S_M(k)$ 值

定理 3 证明了 $S_M(k_0) = 0$，从图 3 中可以看出 $S_M(k_0)$ 并不为零，这是因为推导过程中认为密钥 k 产生的嵌入路径上，经过嵌入率为 λ 的隐写后，$\alpha\lambda F(k)$ 个 MH 点成为 HM 点，$\beta\lambda F(k)$ 个 HM 点转化为 MH 点。同时，$\beta\lambda F(k)$ 个 LM 点成为 ML 点，$\alpha\lambda F(k)$ 个 ML 点成为 LM 点。然而，在实际的隐写中，HM 点与 MH 点的转化以及 LM 点与 ML 点的转化并不一定完全按此规则进行，从而使得 $S_M(k)$ 值与其理论推导值稍有偏离，但仍满足 $S_M(k_0) < S_M(k_j)$。

以上述 30 幅灰度图像为载体，分别利用本文方法、文献[8]中的卡方检验法和文献[9]中的碰撞攻击方法进行密钥恢复实验。图 4 显示的是 3 种算法的密钥恢复成功率，即某一嵌入率下能成功进行密钥恢复的图像数与载体数之比，表 3 列出了 lena 图像

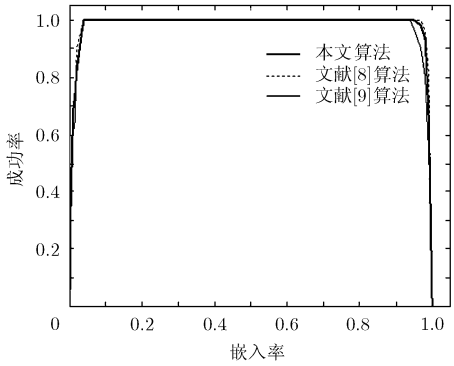


图 4 密钥恢复成功率

表 3 lena 图像的密钥恢复结果

嵌入率	本文算法	文献[8]算法	文献[9]算法
0.005	失败	失败	失败
0.010	成功	成功	失败
0.015	成功	成功	成功
0.400	成功	成功	成功
0.960	成功	成功	成功
0.990	成功	成功	失败
0.995	失败	失败	失败

各种嵌入率下的密钥恢复结果。

从图 4 和表 3 可以看出，由于本文方法，文献[8]中方法和文献[9]中方法都基于真伪路径上数据统计差异确定真密钥，性能有相似之处：都只能在一定嵌入率范围内成功恢复隐写密钥。由于 $S_M(k)$ 值与密钥 k 产生的嵌入路径上 MH 点与 HM 点频次差，ML 点与 LM 点频次差有关，当嵌入率较低时，真伪路径上这些点出现的频次都比较少，数据量不足从而导致密钥恢复成功率较低；当嵌入率较大时，由于真伪路径嵌入率较为接近，也会导致密钥恢复成功率较低。在较小嵌入率与较大嵌入率时，本文算法密钥恢复成功率与文献[8]中算法接近，稍优于文献[9]中算法。

另外，由于 3 种算法均是对密钥空间中的所有密钥进行试验，计算复杂度也是衡量算法性能的重要指标。三者计算复杂度均与嵌入路径上的元素数有关，随着嵌入率的增大而增加。但本文算法仅通过简单计算恢复密钥，复杂度较低。以 40%的嵌入率为例，卡方检验法密钥搜索速度为每秒 89 个密钥，碰撞攻击方法密钥搜索速度为每秒 169 个密钥，本文方法密钥搜索速度为每秒 165 个密钥。可以看出，本文方法的密钥搜索速度明显优于文献[8]中方法，与文献[9]中碰撞攻击方法接近。但本文是从隐

写检测的角度完成真伪密钥区分，文献[9]是利用密码分析方法。鉴于目前针对不同类型隐写方法的隐写检测算法日益增多，而一个好的检测算法一般都可转化为一个密钥恢复方法，因此，本文方法可为其他类型隐写术的密钥恢复方法提供更一般的思路。

5　结束语

本文从隐写检测的角度，提出一种基于图像像素划分的主动隐写分析方法，解决了空域图像随机 LSB 替换隐写的密钥恢复问题。通过对划分的各类点在隐写和置反操作后频次变化的分析，得出了区分真伪密钥的依据，从而设计了相应的密钥恢复算法。

由于本文的重点不是密码分析而是解决隐写检测与密钥恢复的转化问题，因此没有考虑伪随机数发生器的具体结构，所以搜索密钥时只是简单的穷举所有密钥。在后续工作中，我们将结合相应的密码分析方法降低密钥搜索复杂度，并研究其他常用隐写术的密钥恢复方法。

参 考 文 献

[1] Chandramouli R. A mathematical framework for active steganalysis[J]. *ACM Multimedia Systems Journal, Special Issue on Multimedia Watermarking*, 2003, 9(3): 301–311.

[2] Fridrich J, Kodovsky J, Holub V, *et al*.. Steganalysis of content-adaptive steganography in spatial domain[J]. *Lecture Notes in Computer Science*, 2011, 6958/2011: 102–117.

[3] Amirkhani H and Rahmati M. New framework for using image contents in blind steganalysis systems[J]. *Journal of Electronic Imaging*, 2011, 20(1): 013016-1–013016-14.

[4] Kodovsky J, Pevny T, and Fridrich J. Modern steganalysis can detect YASS[C]. SPIE Electronic Imaging Proceedings, Media Forensics and Security XII, San Jose, CA, 2010, 7541: 754102–754102-11.

[5] Fridrich J and Goljan M. On Estimation of secret message length in LSB steganography in spatial domain[C]. SPIE Electronic Imaging Proceedings, Security, Steganography, and Watermarking of Multimedia Contents VI, San Jose, CA, 2004, 5306: 23–34.

[6] Fridrich J, Goljan M, and Soukal D. Searching for the stego key[C]. SPIE Electronic Imaging Proceedings, Security, Steganography, and Watermarking of Multimedia Contents VI, San Jose, CA, 2004, 5306: 70–82.

[7] 王朔中, 张新鹏, 张卫明. 以数字图像为载体的隐写分析研究进展[J]. 计算机学报, 2009, 32(7): 1247–1263.
Wang S Z, Zhang X P, and Zhang W M. Recent advances in image-based steganalysis research[J]. *Chinese Journal of Computers*, 2009, 32(7): 1247–1263.

[8] Fridrich J, Goljan M, and Soukal D. Forensic steganalysis: determining the stego key in spatial domain steganography[C]. SPIE Electronic Imaging Proceedings, Security, Steganography, and Watermarking of Multimedia Contents VII, San Jose, CA, 2005, 5681: 631–642.

[9] 张卫明, 李世取, 刘九芬. 对空域图像LSB隐写术的提取攻击[J]. 计算机学报, 2007, 30(9): 1625–1631.
Zhang W M, Li S Q, and Liu J F. Extracting attack to LSB steganography in spatial domain[J]. *Chinese Journal of Computers*, 2007, 30(9): 1625–1631.

[10] 翟卫东, 吕述望, 刘振华. 基于广义高斯分布的彩色图像空域隐写检测算法[J]. 通信学报, 2004, 25(2): 33–42.
Zhai W D, Lv S W, and Liu Z H. Spatial stego-detecting algorithm in color images based on GGD[J]. *Journal of China Institute of Communications*, 2004, 25(2): 33–42.

[11] University of South California: The USC-SIPI image database [OL]. http://sipi.usc. edu/services/database/Database. html, 2011.1.

刘　静：女，1985 年生，博士，研究方向为信息隐藏、 隐写分析.

汤光明：女，1963 年生，教授，博士生导师，研究方向为信息安全、密码学、信息隐藏.