文章编号 1004-924X(2008)07-1295-08

可抵抗 SPA 分析的 HSBH 改进算法

张 健[1,2],于晓洋[2],黄海霞[2],范身申[2]
(1.东北林业大学 信息与计算机工程学院,黑龙江 哈尔滨 150000;
2.哈尔滨理工大学 测控技术与通信工程学院,黑龙江 哈尔滨 150040)

摘要:为了提高图像信息隐藏方法的鲁棒性以及增加隐藏容量,提出了一种可以抵抗 SPA 分析的 HSBH 算法。介绍了 SPA 方法的思想和原理,分析了高位空域隐藏算法 HSBH,得出其并不能抵抗 SPA 分析的结论。进而对 HSBH 算法进行了改进,提出了一种有效的可以抵抗 SPA 攻击的方法。实验结果表明:借助 SPA 方法的思想,可以在隐藏率>3%的情况下,准确检测其是否含有隐藏信息,正确检测率达到 95%以上。改进后的 HSBH 算法,无论嵌入量多少,SPA 分析都将得到一个非常小的估计值,并做出未隐藏信息的错误判断,从而达到安全隐藏信息的目的。

关 键 词:信息隐藏; LSB 算法; HSBH 算法; SPA
中图分类号:TP309.07 **文献标识码**:A

Improved HSBH algorithm against SPA

ZHANG Jian[1,2], YU Xiao-yang[2], HUANG Hai-xia[2], FAN Shen-shen[2]

(1. *Information and Computer Engineering College, Northeast Forestry University, Harbin* 150000, *China*;
2. *College of Measurement-control Technology & Communications Engineering, Harbin University of Science and Technology, Harbin* 150080, *China*)

Abstract: In order to improve the robust of image information hiding method and to increase the hiding capacity, a High Significant Bit Hiding(HSBH) algorithm against SPA (Sample Pair Analysis) analysis is proposed. The idea and the principle of SPA are introduced, and the HSBH algorithm is analyzed, results show that the HSBH drawback can not resist the SPA attack. So an improved HSBH algorithm is presented. The experimental results indicate that, considering the SPA idea, the correct rate of detection is nearly 95% when the embedding rate is more than 3%. And no matter how much message is embedded by the improved algorithm, SPA steganalysis will get a very small estimation value and form an incorrect judgment to hide information, so that the information can be hidden securely.

Key words: steganography; Low Significant Bit(LSB) algorithm; High Significant Bit Hiding(HSBH) algorithm; Sample Pair Analysis(SPA)

收稿日期:2007-11-27;**修订日期**:2008-01-08.
基金项目:国家自然科学基金资助项目(No. 30571455);东北林业大学校基金资助项目(No. 07005)

1 引　言

信息隐藏技术是近年来信息安全领域一个新的研究方向，该技术是将秘密信息隐藏在其它载体中，使人察觉不到，从而保证秘密信息的安全。信息隐藏按嵌入域分类可分为空域和变换域隐藏。在基于空域的图像信息隐藏技术中，最低有效位(LSB)算法信息容量大、易于实现，但是隐藏位置不可靠，鲁棒性很差[1]。文献[2]提出了一种利用颜色量化的方法来实现最高有效位(MSB)的嵌入算法，其鲁棒性有了较大提高。文献[3]又提出了一种新的利用 MSB 嵌入秘密信息的算法(HB 算法)，但是 HB 算法嵌入的信息容量依赖于载体，最大也只能达到图像大小的 30%左右。文献[4]提出了一种新的图像空域信息隐藏算法 HSBH(High Significant Bit Hiding)，该算法是对 HB 算法的扩展，把嵌入位由最高位扩展成最高四位，而且嵌入信息后，载体图像的像素值变化量极小，人眼无法察觉。嵌入信息量一般可达图像大小的 50%左右，即嵌入率最大可达到 50%，是一种比较好的空域隐藏算法。本文在分析 HSBH 算法原理的基础上，借助 SPA 方法的思想，准确地检测出其隐藏量，在隐藏率＞3%的情况下，检测的正确率达到 95%以上。同时对 HSBH 算法进行改进，提出了一种抗 SPA 分析的高位空域隐藏方法，经大量实验证明了该方法的可行性和有效性。

2 SPA 分析方法

SPA 算法是由 Sorina Dumitrescu 等人提出的，该算法通过分析载体信号样本对在 LSB 信息隐藏下的状态转移情况估算信息的嵌入率[5-7]。

用连续的样本 S_1，S_2，…，S_N 表示图像的像素值(下标表示样本在图像中的位置)，$P=\{(S_i,S_j)\mid 1\leqslant i,j\leqslant N\}$是样本对的集合，这里 $0\leqslant S_i,S_j\leqslant 2^b-1$，其中 b 是每个样本值的比特数，且 S_i,S_j 是两个相邻的像素值。还有一些具体的集合和参数分别为：

$D_n=\{(u,v)\in P\mid |u-v|=n\}$，$D_n$ 是 P 的子多重集，表示样本对两像素值的差的绝对值等于 n 的集合，其中 $0\leqslant n\leqslant 2^b-1$。

X_{2m+1}是 D_{2m+1}的子多重集，表示样本对两像素值的差的绝对值等于 $2m+1$，且两像素值中偶数较大的集合。

Y_{2m+1}是 D_{2m+1}的子多重集，表示样本对两像素值的差的绝对值等于 $2m+1$，且两像素值中奇数较大的集合。

X_{2m}是 D_{2m}的子多重集，表示样本对两像素值的差的绝对值等于 $2m$，且两像素值都是偶数的集合。

Y_{2m}是 D_{2m}的子多重集，表示样本对两像素的值差的绝对值等于 $2m$，且两像素值都是奇数的集合。

$|X_{2m+1}|$，$|Y_{2m+1}|$，$|X_{2m}|$，$|Y_{2m}|$分别表示自然载体图像多重集 X_{2m+1}，Y_{2m+1}，X_{2m}，Y_{2m}的势。

$|X_{2m+1}'|$，$|Y_{2m+1}'|$，$|X_{2m}'|$，$|Y_{2m}'|$分别表示嵌入信息后，载密图像多重集 X_{2m+1}'，Y_{2m+1}'，X_{2m}'，Y_{2m}'的势。

对自然图像而言，D_{2m+1}中的样本对的奇分量大的概率为$\frac{1}{2}$，即有下式的假设：

$$\mathrm{E}\{|X_{2m+1}|\}=\mathrm{E}\{|Y_{2m+1}|\}. \tag{1}$$

当有秘密信息嵌入时，$|X_{2m+1}|$与$|Y_{2m+1}|$的差值将增大，SPA 分析方法基于假设(1)，对样本对之间的转移关系进行统计，以嵌入率作为未知量建立方程，通过统计待检测图像中样本对之间的一些特征量，得到嵌入率，将隐藏的比率与设定的门限值进行比较来做出图像是否隐藏信息的判断。

3 HSBH 算法及 SPA 分析

3.1 HSBH 算法简介

HSBH 算法是一种空域信息高位隐藏算法，它是将信息隐藏到某一像素的高 4 位中，但不是在一个固定的高位嵌入，而是根据 Logistic 混沌映射公式 $Z_{n+1}=\lambda Z_n(1-Z_n)$计算出 Z_n 的值，将信息嵌入到载体图像的不同高位中。在适合高位隐藏的像素点上嵌入信息，当在第 k 位隐藏 1 时，若像素值的第 k 位为 1，则像素值不变，否则像素值的高 $8-k$ 比特不变，低 k 比特变为 2^{k-1}，即让第 k 位为 1；当在第 k 位隐藏 0 时，若像素值的第 k 位为 0，则像素值不变，否则像素值的高 $8-k$ 比特不变，低 k 比特变为 $2^{k-1}-1$，即让第 k 位为 0。

为使像素值在信息嵌入后改变不大,规定可隐藏信息的区间中的数只能是低 k 位比特的值在 2^{k-1} 左右的数,每个区间中元素的个数由参数确定。

设秘密信息是一幅灰度图像,其矩阵记为 $W=\{w(i,j)\ 1\leqslant i\leqslant M_1,1\leqslant j\leqslant M_2\}$,即秘密信息图像的大小为 $M_1\times M_2$。载体图像记为 $F=\{f(x,y)\ 1\leqslant x\leqslant N_1,1\leqslant y\leqslant N_2\}$,即载体图像的大小为 $N_1\times N_2$,嵌入信息后的载体图像称为载密图像,用 $F'=\{f'(x,y)\ 1\leqslant x\leqslant N_1,1\leqslant y\leqslant N_2\}$ 表示,其中(i,j)代表秘密信息图像的像素坐标;(x,y)代表原始载体图像和载密图像的像素坐标,$w(i,j)$、$f(x,y)$和 $f'(x,y)$分别代表相应位置的像素值。公式(2)说明在第 8 位隐藏信息时像素值的相应变化情况,其它高位以此类推。

设 $w(i,j)_n=0$ 则:

$$f'(x,y)=\begin{cases}127 & f(x,y)\in[128,128+r)\\ f(x,y) & f(x,y)\in[128-r,127]\end{cases}.$$

设 $w(i,j)_n=1$ 则:　　(2)

$$f'(x,y)=\begin{cases}f(x,y) & f(x,y)\in[128,128+r)\\ 128 & f(x,y)\in[128-r,127]\end{cases}.$$

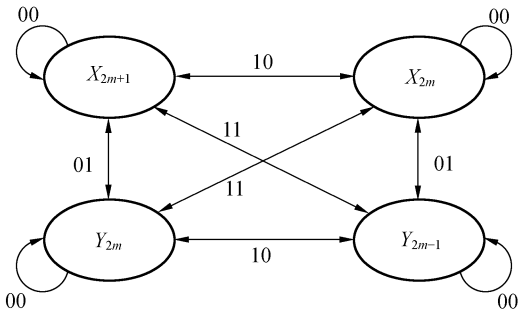
3.2　SPA 分析

HSBH 算法中参数 r 取值越小,隐藏信息后图像的不可见性就越好,但隐藏容量越小;r 取值越大,隐藏信息后图像的不可见性就越差,但隐藏容量也越大。可根据实际情况对 r 进行合理的取值,由于篇幅有限,文中不能对每种情况都进行公式推导,$r=1$ 的情况与 LSB 情况类似,基于有限状态机理论,而 $r\geqslant 2$ 的情况也类似,文中列出了 $r=2$的状态转移情况,推导了 $r=4$ 的嵌入率公式,其他情况类推。

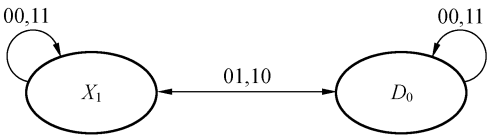
(1) $r=1$

可隐藏的区间有两个元素,以能进行第 8 位隐藏的区间[127,128]为例,当要隐藏的信息为 0 时,128 变成 127;当要隐藏的信息为 1 时,127 变成 128,即奇数变化时只有一种情况,奇数加 1 变成偶数;偶数变化时只有一种情况,偶数减 1 变成奇数。本文用数字 0 表示像素值不发生变化,数字 1 表示奇数加 1 变成偶数或者偶数减 1 变成奇数。

嵌入信息后,样本对中每个像素值有 0、1 两种变化,样本对有 4 种变化,修改模式 $\pi\in\{00,01,10,11\}$。这里不妨假设修改模式的第 1 位数字代表样本对中像素值小的点的变化情况,第 2 位数字代表样本对中像素值大的点的变化情况。当 X_{2m+1}中的样本对经过 10 变化时,像素值小的奇数经过 1 变化变成偶数,值增大 1,像素值大的偶数不变,所以样本对的值差变成 $2m$,且两个都为偶数,样本对属于 X_{2m},同理可得到 HSBH($r=1$)嵌入的有限状态图 1:



(a)$m\geqslant 1$



(b)$m=0$

图 1　HSBH($r=1$)嵌入的有限状态图

Fig. 1　Finite-state machine of HSBH ($r=1$) embedding

令嵌入率为 p,则 4 种修改模式的概率分别为 $\rho(00)=(1-\frac{p}{2})^2$,$\rho(01)=(1-\frac{p}{2})\frac{p}{2}$,$\rho(10)=\frac{p}{2}(1-\frac{p}{2})$,$\rho(11)=(\frac{p}{2})^2$。令 $E_m=X_{2m+1}+X_{2m}+Y_{2m}+Y_{2m-1}$,$E_m$ 是闭合的,即 $|E_m|=|E_m'|$。由状态图 1 和 $\mathrm{E}\{|X_{2m+1}|\}=\mathrm{E}\{|Y_{2m+1}|\}$可得:

$$\frac{(|E_m|-|E_{m+1}|)p^2}{4}-\frac{(|D'_{2m}|-|D_{2m+2}'|+2|X_{2m+1}'|-2|Y_{2m+1}'|)p}{2}+|X_{2m+1}'|-|Y_{2m+1}'|=0\quad m\geqslant 1\ ,\qquad(3)$$

$$\frac{(2|E_0|-|E_1|)p^2}{4}-\frac{(2|D_0'|-|D_2'|+2|X_1'|-2|Y_1'|)p}{2}+|X_1'|-|Y_1'|=0\quad m=0\ .\qquad(4)$$

解方程(3)或(4)中的任意一个即可得嵌入率 p 的估计值。

为了得到更准确的估计值,可将式(1)的假设

修改为公式(5)：

$$\mathrm{E}\{|\bigcup_{m=i}^{j} X_{2m+1}|\}=\mathrm{E}\{|\bigcup_{m=i}^{j} Y_{2m+1}|\}. \quad (5)$$

从而推导出更可靠的公式：

$$\frac{(|E_i|-|E_{j+1}|)p^2}{4}-\frac{(|D_{2i}{}'|-|D_{2j+2}{}'|+2\sum_{m=i}^{j}|X_{2m+1}{}'|-2|Y_{2m+1}{}'|)p}{2}+\sum_{m=i}^{j}(|X_{2m+1}{}'|-|Y_{2m+1}{}'|)=0, i\geqslant 1, \quad (6)$$

$$\frac{(2|E_0|-|E_{j+1}|)p^2}{4}-\frac{[|D_0{}'|-|D_{2j+1}{}'|+2\sum_{m=0}^{j}(|X_{2m+1}{}'|-|Y_{2m+1}{}'|)]p}{2}+\sum_{m=0}^{j}(|X_{2m+1}{}'|-|Y_{2m+1}{}'|)=0, i=0. \quad (7)$$

HSBH(r=1)嵌入信息时，像素值变化为奇数加 1 变成偶数，偶数减 1 变成奇数，与 LSB 算法的奇数减 1 变成偶数，偶数加 1 变成奇数相反，所以这两种算法嵌入信息后，对图像的统计特性影响不同，LSB 使$|Y_{2m+1}{}'|-|X_{2m+1}{}'|$随着嵌入率的增大而增大，而 HSBH($r$=1)使$|Y_{2m+1}{}'|-|X_{2m+1}{}'|$随着嵌入率的增大而减小，即$|X_{2m+1}{}'|-|Y_{2m+1}{}'|$随着嵌入率的增大而增大，由于这两种算法都是有限状态转移，可在只知道载密图像的情况下计算出嵌入率，嵌入率的求取公式也类似。

(2) r=2

可隐藏的区间有 4 个元素，以能进行第 8 位隐藏的区间[126,129]为例，当要隐藏的信息为 0 时，129 变成 127，128 变成 127。当要隐藏的信息为 1 时，126 变成 128，127 变成 128，奇数变化时有两种情况：(1)奇数加 1 变成偶数；(2)奇数减 2 变成奇数。偶数变化时有两种情况：(1)偶数减 1 变成奇数；(2)偶数加 2 变成偶数。本文用数字 2 表示奇数减 2 变成奇数或者偶数加 2 变成偶数的变化。嵌入信息后，样本对的每个像素值有 0，1，2 这 3 种变化，样本对有 3^2 种变化，修改模式 $\pi\in\{00,01,02,10,20,11,12,21,22\}$，表 1 为转移关系表。

表 1　HSBH 嵌入的状态转移表

Tab. 1　State transition of HSBH embedding

修改模式	修改模式出现的概率	X_{2m+1}中的样本对经过修改后归属的集合	Y_{2m+1}中的样本对经过修改后归属的集合	X_{2m}中的样本对经过修改后归属的集合	Y_{2m}中的样本对经过修改后归属的集合
00	p_1	X_{2m+1}	Y_{2m+1}	X_{2m}	Y_{2m}
01	p_2	Y_{2m}	X_{2m+2}	Y_{2m-1}	X_{2m+1}
02	p_3	X_{2m+3}	Y_{2m-1}	X_{2m+2}	Y_{2m-2}
10	p_4	X_{2m}	Y_{2m+2}	X_{2m+1}	Y_{2m-1}
20	p_5	X_{2m+3}	Y_{2m-1}	X_{2m-2}	Y_{2m+2}
11	p_6	Y_{2m-1}	X_{2m+3}	Y_{2m}	X_{2m}
12	p_7	X_{2m+2}	Y_{2m}	X_{2m+3}	Y_{2m-3}
21	p_8	Y_{2m+2}	X_{2m}	Y_{2m-3}	X_{2m+3}
22	p_9	X_{2m+5}	Y_{2m-3}	X_{2m}	Y_{2m}

其中 $p_1=(1-\frac{p}{2})^2$，$p_2=p_3=p_4=p_5=\frac{p}{4}(1-\frac{p}{2})$，$p_6=p_7=p_8=p_9=\frac{p^2}{16}$.

由表 1 的转移关系可求出：

$$|X_{2m+1}{}'|-|Y_{2m+1}{}'|=\left(1-\frac{p}{2}\right)^2(|X_{2m+1}|-|Y_{2m+1}|)+\frac{p}{4}\left(1-\frac{p}{2}\right)(2|X_{2m-1}|-2|Y_{2m+3}|+|D_{2m}|-|D_{2m+2}|+\frac{p^2}{16}(|X_{2m-3}|+|Y_{2m-1}|-|X_{2m+3}|-|Y_{2m+5}|+|D_{2m-2}|-|D_{2m+4}|)(m\geqslant 2).$$

由于 $\mathrm{E}\{|X_{2m+1}|\}=\mathrm{E}\{|Y_{2m+1}|\}$，公式等号右边第一项为 0，自然图像一般满足规律$|X_n|$、$|Y_n|$随着 n 的增大而减小，公式等号右边括号中的项>0，且$\frac{p}{4}(1-\frac{p}{2})$，$\frac{p^2}{16}$在 $p\in[0,1]$时单调递增，所以$|X_{2m+1}{}'|-|Y_{2m+1}{}'|$随着 p 增大而增大。

(3) r=4

可隐藏的区间有 8 个元素，嵌入信息后，像素值除了有 0、1、2 变化外，还增加了 3、4 变化，3 变

化表示奇数加 3 变成偶数或者偶数减 3 变成奇数;4 变化表示奇数减 4 变成奇数或者偶数加 4 变成偶数,所以修改模式有 5^2 种。与 $r=2$ 的情况一样,列出 $r=4$ 的状态转移表,并计算出各种修改模式出现的概率,通过推导,可求出下列方程:

$$\left(1-\frac{p}{2}\right)^2(|X_{2m+1}|-|Y_{2m+1}|)+\frac{p}{8}\left(1-\frac{p}{2}\right)(2|X_{2m-3}|+2|X_{2m-1}|-2|Y_{2m+5}|-2|Y_{2m+3}|+|D_{2m-2}|+|D_{2m}|-|D_{2m+2}|-|D_{2m+4}|)+\frac{p^2}{64}(|X_{2m-7}|+2|X_{2m-5}|+|X_{2m-3}|+|Y_{2m-5}|+2|Y_{2m-3}|+|Y_{2m-1}|-|X_{2m+3}|-2|X_{2m+5}|-|X_{2m+7}|-|Y_{2m+5}|-2|Y_{2m+7}|-|Y_{2m+9}|+|D_{2m-6}|+2|D_{2m-4}|+|D_{2m-2}|-|D_{2m+4}|-2|D_{2m+6}|-|D_{2m+8}|)+|Y_{2m+1}'|-|X_{2m+1}'|=0(m\geqslant 4)\ . \tag{9}$$

令 $A_m=2|X_{2m-3}|+2|X_{2m-1}|-2|Y_{2m+5}|-2|Y_{2m+3}|+|D_{2m-2}|+|D_{2m}|-|D_{2m+2}|-|D_{2m+4}|$,

$B_m=|X_{2m-7}|+2|X_{2m-5}|+|X_{2m-3}|+|Y_{2m-5}|+2|Y_{2m-3}|+|Y_{2m-1}|-|X_{2m+3}|-2|X_{2m+5}|-|X_{2m+7}|-|Y_{2m+5}|-2|Y_{2m+7}|-|Y_{2m+9}|+|D_{2m-6}|+2|D_{2m-4}|+|D_{2m-2}|-|D_{2m+4}|-2|D_{2m+6}|-|D_{2m+8}|$,

$C_m=|Y_{2m+1}'|-|X_{2m+1}'|$。

则公式(9)化简为:

$$\left(1-\frac{p}{2}\right)^2(|X_{2m+1}|-|Y_{2m+1}|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_m+\frac{p^2}{64}B_m+C_m=0\ . \tag{10}$$

同理得到下列公式:

$$\left(1-\frac{p}{2}\right)^2(|X_1|-|Y_1|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_0+\frac{p^2}{64}B_0+C_0=0,m=0, \tag{11}$$

其中,$A_0=2|Y_1|-2|Y_5|+|D_2|-|D_4|$,

$B_0=|X_1|+|X_3|+|Y_3|+|Y_5|-|X_5|-|X_7|-|Y_7|-|Y_9|+|D_2|+|D_4|-|D_6|-|D_8|$,

$C_0=|Y_1'|-|X_1'|$.

$$\left(1-\frac{p}{2}\right)^2(|X_3|-|Y_3|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_1+\frac{p^2}{64}B_1+C_1=0,m=1\ , \tag{12}$$

其中,$A_1=2|X_1|+2|Y_1|-2|Y_5|-2|Y_7|+2|D_0|+|D_2|-|D_4|-|D_6|$,

$B_1=2|X_1|+|X_3|+2|Y_1|+2|Y_3|+|Y_5|-|X_5|-2|X_7|-|X_9|-|Y_7|-2|Y_9|-|Y_{11}|+2|D_0|+2|D_2|-2|D_6|-2|D_8|-|D_{10}|$,

$C_1=|Y_3'|-|X_3'|$。

$$\left(1-\frac{p}{2}\right)^2(|X_5|-|Y_5|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_2+\frac{p^2}{64}B_2+C_2=0,m=2\ , \tag{13}$$

其中,$A_2=2|X_1|+2|X_3|-2|Y_7|-2|Y_9|+|D_2|+|D_4|-|D_6|-|D_8|$,

$B_2=2|X_1|+4|Y_1|+2|Y_3|-|X_7|-2|X_9|-|X_{11}|-|Y_9|-2|Y_{11}|-|Y_{13}|+4|D_0|+2|D_2|-|D_8|-2|D_{10}|-|D_{12}|$,

$C_2=|Y_5'|-|X_5'|$。

$$\left(1-\frac{p}{2}\right)^2(|X_7|-|Y_7|)+\frac{p}{8}\left(1-\frac{p}{2}\right)A_3+\frac{p^2}{64}B_3+C_3=0,m=3\ , \tag{14}$$

其中,$A_3=2|X_3|+2|X_5|-2|Y_9|-2|Y_{11}|+|D_4|+|D_6|-|D_8|-|D_{10}|$,

$B_3=2|X_1|+|X_3|+2|Y_1|+2|Y_3|+|Y_5|-|X_9|-2|X_{11}|-|X_{13}|-|Y_{11}|-2|Y_{13}|-|Y_{15}|+2|D_0|+2|D_2|+|D_4|-|D_{10}|-2|D_{12}|-|D_{14}|$,

$C_3=|Y_7'|-|X_7'|$。

利用 $E\{|X_{2m+1}|\}=E\{|Y_{2m+1}|\}$ 可把公式(10)、(11)、(12)、(13)、(14)中的第一项消掉得方程:

$$\frac{p}{8}\left(1-\frac{p}{2}\right)A_m+\frac{p^2}{64}B_m+C_m=0(m\geqslant 0)\ . \tag{15}$$

为了提高估计精度,可将公式(1)的假设修改为公式(5),从而推导出更可靠的方程:

$$\frac{p}{8}\left(1-\frac{p}{2}\right)\sum_{m=i}^{j}A_m+\frac{p^2}{64}\sum_{m=i}^{j}B_m+\sum_{m=i}^{j}C_m=0(m\geqslant 0). \tag{16}$$

公式(16)中的参数 A_m,B_m 的值需要从原始载图中获得,在不知道原载体图像的情况下,为了计算出嵌入率的值,可用 $\frac{|X_{2m+1}'|+|Y_{2m+1}'|}{2}$ 代替 $|X_{2m+1}|$、$|Y_{2m+1}|$ 的值,$|D_{2m}'|$ 代替 $|D_{2m}|$,这是由于 HSBH 隐藏信息后 $|X_{2m+1}'|=|X_{2m+1}|+\Delta|X|$,$|Y_{2m+1}'|=|Y_{2m+1}|-\Delta|Y|$,所以 $\frac{|X_{2m+1}'|+|Y_{2m+1}'|}{2}=\frac{|X_{2m+1}|+|Y_{2m+1}|}{2}+$

$\frac{\Delta|X|-\Delta|Y|}{2}(m\geqslant 1)$，$\frac{|X_{2m+1}'|+|Y_{2m+1}'|}{2}\approx|X_{2m+1}|\approx|Y_{2m}|$，而$|D_{2m}'|\approx|D_{2m}|$，这样可近似求出 A_m，B_m 的值，从而利用公式(16)求出嵌入率 p，嵌入率的值为方程的两个解中绝对值小的那个解。

综上所述，HSBH 信息嵌入导致$|X_{2m+1}'|-|Y_{2m+1}'|$随着嵌入率的增大而增大，SPA 分析方法可检测 HSBH 算法是否隐藏信息。

3.3 实验结果与分析

为了验证 SPA 分析 HSBH 的可行性，本文用 $r=4$ 时的 HSBH 算法进行信息隐藏及 SPA 分析实验。选用了 8 幅 512×512 常用图像，如图 2 所示，对每幅图像嵌入 3.13%，6.25%，12.5%，25%，40.28%的信息，用公式(16)对载密图像进行嵌入率的估计，其中 $i=0$，$j=25$。表 2 给出了这些图像进行 HSBH($r=4$)隐藏后，嵌入率的值以及嵌入率的平均估计值。

图 2　实验图像组

Fig. 2　Test images

表 2　SPA 估计嵌入率

Tab. 2　Embedding rates estimated with SPA

图像 \ 嵌入率(%)	0	3.13	6.25	12.5	25	40.28
Lena	0.85	3.81	7.26	13.83	26.79	42.24
Mandrill	−1.78	2.70	6.86	13.46	29.89	51.32
Zelda	0.06	3.76	8.01	15.65	29.94	45.74
Boat	−0.72	2.22	5.46	12.42	23.06	40.68
Goldhill	−1.07	1.85	5.33	11.66	25.04	43.35
Cameraman	0.25	2.58	4.53	9.61	20.75	34.36
Peppers	−0.67	2.05	5.17	12.15	26.60	42.73
Jet plane	0.9	3.01	5.51	11.15	23.61	38.74
平均值	−0.27	2.75	6.02	12.49	25.71	42.39

从表中可以看出，估算出的嵌入率都非常接近真实值，门限值取 1.8%时，在隐藏率＞3%的情况下，检测的正确率几乎达到 100%。这也说明基于 SPA 思想进行 HSBH 检测的合理性。这其中也存在微小的误差。由于公式(16)是基于自然图像的统计特性 $E\{|\bigcup_{m=i}^{j}X_{2m+1}|\}=E\{|\bigcup_{m=i}^{j}X_{2m+1}|\}$，对于单幅图像，$|\bigcup_{m=i}^{j}X_{2m+1}|$ 与 $|\bigcup_{m=i}^{j}X_{2m+1}|$ 不相等，这从嵌入率为 0 时的图像估计值就能看出，如果载体图像满足 $|\bigcup_{m=i}^{j}X_{2m+1}|=|\bigcup_{m=i}^{j}X_{2m+1}|$，则估算出的嵌入率应为 0，而由表中数据可看出，原载图在未嵌入信息的情况下，估算出的 p 不等于 0，但是值也接近于 0，所以公式基于 $|\bigcup_{m=i}^{j}X_{2m+1}|=|\bigcup_{m=i}^{j}X_{2m+1}|$ 的假设是很合理的，不过嵌入率估算有误差。当然公式是基于统计特性列出的，A_m，B_m 的值是近似值，这些也造成了估算的嵌入率的误差。

4　HSBH 改进算法

由于 HSBH 算法不能抵御 SPA 方法分析，

必须对其进行改进使其能抵抗 SPA 分析。文献[5]指出,检测得到的估计值随着偏离度 $\delta_{i,j}=\frac{|\bigcup_{m=i}^{j}Y_{2m+1}'|-|\bigcup_{m=i}^{j}X_{2m+1}'|}{|\bigcup_{m=i}^{j}Y_{2m+1}'|+|\bigcup_{m=i}^{j}Y_{2m+1}'|}$ 的增大而增大,所以 $|\bigcup_{m=i}^{j}Y_{2m+1}'|-|\bigcup_{m=i}^{j}X_{2m+1}'|$ 越趋于 0,得到的估计值也越小,若想第三方用 SPA 方法检测得到的估计值为一个低于门限的较小值,则必须使 $\delta_{i,j}$ 值较小。基于这个思想,由于 HSBH 隐藏信息后,$|\bigcup_{m=i}^{j}X_{2m+1}'|>|\bigcup_{m=i}^{j}Y_{2m+1}'|$,需要进行调整,使 $|\bigcup_{m=i}^{j}X_{2m+1}'|$ 值减小,$|\bigcup_{m=i}^{j}Y_{2m+1}'|$ 值增大,文献[5]用 DCLS 方法进行 LSB 补偿。通过在载密图像上动态地选择一块区域,对这个区域的全部像素值进行加 1 或者减 1,使像素值的奇偶性发生变化,使区域中的 X_{2m+1}' 与 Y_{2m+1}' 集合中的样本对互换,来达到 X_{2m+1}'' 与 Y_{2m+1}'' 的差值减小的目的。本文采用的方法则是在载密图像中动态选取区域,区域中的奇数减 1,偶数加 1。属于集合 X_{2m+1}' 的样本对的偶数值大于奇数值,修改后偶数值加 1 变成奇数,奇数值减 1 变成偶数,此时的奇数大于偶数,且两像素的差值由 $2m+1$ 变成 $2m+3$,即载密图像中 X_{2m+1}' 中样本对变成改进后载密图像中集合 Y_{2m+3}'' 中样本对,同理属于集合 Y_{2m+3}' 的样本对经改进修改后变成 X_{2m+1}'' 的样本对,修改使 X_{2m+1}' 与 Y_{2m+3}' 中的样本对互换,即在修改后的修改域中,$|X_{2m+1}''|=|Y_{2m+3}''|$,$|Y_{2m+3}''|=|X_{2m+1}'|$,所以修改后,图像的 $|X_{2m+1}''|$ 值减小,$|Y_{2m+1}''|$ 值增大,这样可消除 $|X_{2m+1}'|$ 与 $|Y_{2m+1}'|$ 的偏离度,区域的选择应该使 $|\bigcup_{m=i}^{j}X_{2m+1}''|$ 与 $|\bigcup_{m=i}^{j}Y_{2m+1}''|$ 的差值尽可能地小。与对选择区域的像素全部进行加 1 或者减 1 的 DCLS 算法相比,这种改进方法的优点在于需要进行修改的区域更小,这是因为 DCLS 是 X_{2m+1}' 与 Y_{2m+1}' 集合中的样本对互换,而本文中的改进方法是将 X_{2m+1}' 与 Y_{2m+3}' 中的样本对互换,$|X_{2m+1}'|$ 与 $|Y_{2m+1}'|$ 的差值比 $|X_{2m+1}'|$ 与 $|Y_{2m+1}'|$ 差值大,所以调节的力度更大,修改所需的区域更小。经过改进后,秘密信息的提取需要把修改区域的信息作为密钥,对修改区域进行偶数值加 1,奇数值减 1 计算,接下来的提取步骤与 HSBH 算法一样。

5　实验结果

按照 4 节中的改进算法,对前文所列的 8 幅图的载密图像进行改进,表 3 中给出了 Lena 改进前后,用 SPA 进行分析得到的嵌入率值。实验时,选择的区域是使用最简单的取图像前面几行的值进行修改,实际上修改的区域可以任意选取,只需把修改区域的信息作为密钥的一部分。实验发现只要适当地选取修改的区域,估算出的值就可取到非常小,远小于门限值 1.8%。根据视觉特性,人眼对图像平滑区的噪声较敏感,而对较复杂的纹理区的噪声不敏感,修改区域的合理选择,还可以使图像更好地满足不可见性要求。图 3 为 Lena 隐藏 40.28%信息的实验图像对比,其中(a)为密图,(b)为原载图,(c)为载密图,(d)为改进后载密图。

表 3　改进前后的嵌入率值比较

Tab. 3　Comparison of embedding rates before and after improvement

嵌入率(%)	改进前	改进后
0	0.85	0.85
3.13	3.81	0.16
6.25	7.26	0.12
12.50	13.83	0.058
25.00	26.79	0.23
40.28	42.24	0.28

(a)密图　(b)原载图

(a)Key image　(b)Original loading image

(c)载密图　(d)改进后载密图

(c)Key-loaded image　(d)Improved key-loaded image

图 3　实验图像对比

Fig. 3　Comparison of test images

6 结　论

本文针对目前的频域算法隐藏容量小以及空域算法鲁棒性不强的缺点，提出了一种可以抵抗SPA分析的HSBH算法。介绍了SPA方法和高位空域隐藏算法HSBH的基本原理，然后对HSBH进行了详细的分析，指出其不能抵抗SPA的分析，进而对HSBH算法进行改进，使之可以有效地抵抗SPA攻击。实验结果证明：借助SPA方法的思想，可以在隐藏率＞3%的情况下，准确检测其是否含有隐藏信息，正确检测程度达到95%以上。改进后的HSBH算法，鲁棒性有了很大的提高，达到了安全隐藏信息的目的。

参考文献：

[1] 张建伟.基于图像的LSB隐藏算法位平面分析及算法改进[J].武装指挥技术学院学报,2004,14(2):84-87.
ZHANG J W. Bit-plane analysis of LSB hiding algorithm based on image and improvement of algorithm [J]. *Journal of the Academy of Equipment Command & Technology*,2004,14(2):84-87. (in Chinese)

[2] 任智斌,隋永新,杨英慧,等.以图像为载体的最大意义位(MSB)信息隐藏技术的研究[J].光学 精密工程, 2002,10(2):182-187.
REN ZH B,SUI Y X,YANG Y H, *et al*.. Study of the MSB information_hiding technique in a carrier image[J]. *Opt. Precision Eng.*, 2002,10(2): 182-187. (in Chinese)

[3] 罗大光,范明钰,郝玉洁,等.一种基于图像最高位(MSB)的水印嵌入算法[J].计算机应用,2004,24(12):88-89.
LUO D G,FAN M Y,HAO Y J, *et al*.. An watermarking algorithm based on MSB of image[J]. *Computer Applications*, 2004,24(12):88-89. (in Chinese)

[4] 于晓洋,徐贵森,张健.一种基于混沌的信息高位隐藏新方法[J].电子器件,2007,30(5):1677-1680.
YU X Y, XU G S, ZHANG J. A new method of information high-bit hiding based on chaos[J]. *Chinese Journal of Electron Devices*, 2007,30(5):1677-1680. (in Chinese)

[5] 罗向阳,陆佩忠,刘粉林.一类可抵御SPA分析的动态补偿LSB信息隐藏方法[J].计算机学报,2007,30(3):463-473.
LUO X Y,LU P ZH,LIU F L. A dynamic compensation LSB steganography method defeating SPA[J]. *Chinese Journal of Computer*, 2007,30(3):463-473. (in Chinese)

[6] 陆佩忠,罗向阳,汤庆阳,等.基于三次方程的LSB隐藏信息的盲检测[J].电子与信息学报,2005,27(3):392-396.
LU P ZH,LUO X Y,TANG Q Y,*et al*.. Blind detection of LSB steganography based on a cubic equation[J]. *Journal of Electronics & Information Technology*, 2005,27(3):392-396. (in Chinese)

[7] SORINA D,WU X L,WANG ZH. Detection of LSB steganography via sample pair analysis[J]. *IEEE Transactions on Signal Processing*, 2003, 51(7):1995-2007.

作者简介：张　健(1980－)，男，黑龙江哈尔滨人，博士研究生，主要研究方向为信息安全和混沌理论等。E-mail:zhangjianok00@163.com

于晓洋(1962－)，男，黑龙江哈尔滨人，教授，博士生导师，主要研究方向为信息安全，图像处理，混沌理论等。E-mail:yuxiaoyangok@163.com

黄海霞(1984－)，女，福建省福安人，硕士研究生，主要研究方向为信息安全，图像处理，混沌理论等。E-mail:huanghaixia198404@163.com

范身申(1984－)，女，硕士研究生，主要研究方向为信息安全、图像处理、混沌理论。