文章编号:1009-3087(2015)06-0139-05 DOI:10.15961/j.jsuese.2015.06.019

一种基于次优加权矩阵的灰度图像信息隐藏算法

彭振龙[1,2],桂小林[1*],安 健[1],冀亚丽[1],郭建宏[2]
(1.西安交通大学 电子与信息工程学院,陕西 西安 710049;2.泉州师范学院,福建 泉州 362000)

摘 要:为了增加信息隐藏后载体图像的不可感知性,理论上分析了在一个包含 n 个像素的灰度图像块内,最多改变2个像素,每个像素都只改变最不重要位的情况下,信息隐藏的容量上界,证明了不可能构建完全满足该条件的加权矩阵。提出了一种利用自然数列,构建次优加权矩阵的算法,该算法能将载体图像块内的像素修改值,放大相应倍数,从而用最小的修改,隐藏最多的数据,该算法还能推广至彩色图像。大量仿真实验表明,当分块不大于2 500万像素时,采用该算法构造的次优加权矩阵与最优矩阵进行信息隐藏的容量的差距最多为1个二进制位,该算法较好地平衡了信息隐藏的容量和不可感知性。

关键词:加权矩阵;信息隐藏;灰度图像;次优矩阵
中图分类号:TP391 **文献标志码**:A

An Information Hiding Algorithm of Gray Images Based on Suboptimum Weighted Matrix

PENG Zhenlong[1,2], *GUI Xiaolin*[1*], *AN Jian*[1], *JI Yali*[1], *GUO Jianhong*[2]
(1. School of Electronics and Info. Eng., Xi'an Jiaotong Univ., Xi'an 710049, China; 2. Quanzhou Normal Univ., Quanzhou 362000, China)

Abstract: In order to promote the imperceptibility of the host image after it was embedded the secret information, a gray image was divided into some blocks, each with n pixels. The upper bound capacity of each block was analyzed under the condition that not more than 2 pixels were modified and just the least significant bit was changed in every modified pixel. It was proved that such an optimum weighted matrix couldn't be constructed. Therefore a way to construct a suboptimum weighted matrix, which was a natural sequence, was proposed. The size of the modified pixel was enlarged certain times accordingly by the suboptimum matrix, thus the most hiding capacity could be brought by the least modification. At the same time, this method could be used in color images. Compared with the optimum weighted matrix mentioned above, one less bit could be hidden by the suboptimum weighted matrix in the worst condition that n was not more than 25 million. The better balance between the hiding capacity and the imperceptibility was reached.

Key words: weighted matrix; information hiding; gray image; suboptimum matrix

信息隐藏(information hiding/data hiding)[1],即将秘密信息嵌入到可以公开传输的载体(如图像、声音、视频等)中,利用人类感知的局限性和载体本身的冗余,使人无法察觉秘密信息的存在,从而达到隐秘传输或版权认证的目的。

数字图像具有较高的冗余,成为很好的载体。基于图像的信息隐藏技术有基于空间域的[2-7]、基于变换域的[8-9]、基于压缩域的[10]等。其中,基于空间域的最典型的算法是最不重要位(least significant bit,LSB)。LSB具有实现简单,嵌入容量大,不可感知性高,支持盲提取等特点而广受欢迎。Pan等[2]提出的算法在大小为 n 的分块内,最多修改2个像素值时,最大可隐藏 $\lfloor \mathrm{lb}(n+1) \rfloor$ 位数据。Nguyen等[3]利用多个位平面,改进了以往基于图像边缘的隐藏算法,具有较好的可视性和较大容量。Yang等[4]提出了基于 X 位LSB的自适应信息隐藏

收稿日期:2015-09-25
基金项目:国家科技重大专项资助(2012ZX03002001);国家自然科学基金资助项目(61472316;61172090;61502380);泉州市科技局重点项目资助(2014Z131);福建省青年教师项目资助(JA15401);福建省自然科学基金资助项目(2015J01286)
作者简介:彭振龙(1977—),男,博士生. 研究方向:移动群智感知;物联网;信息安全. E-mail:jxndpzl@163.com
* 通信联系人 E-mail:xlgui@mail.xjtu.edu.cn

算法,根据像素间的差值动态地调整嵌入的数据量,提高了嵌入容量和不可感知性。Lu 等[5]提出了一种改进的插值和直方图平移算法,充分利用了参考像素的隐藏能力,提升了隐藏效率。郭萌等[6]在文献[2]的基础上,进一步提出了只修改1个像素值的算法即可达到文献[2]的隐藏容量,但文献[2,6]都基于二值图像,没有进一步扩展至灰度图像,且文献[6]中将像素分组的方法,只能适应于修改1个像素的情况,无法适应修改2个或以上像素的情况。

提出了一种构建加权矩阵的方法,该矩阵在最多修改2个像素的情况下,隐藏的容量大幅提升,且能方便地应用于灰度及彩色图像。

1 加权矩阵的嵌入原理与过程

1.1 嵌入原理

设待隐藏的信息 s 为长度为 w 的二进制位串,$\boldsymbol{F}$ 为一个含有 n 个像素的掩体分块。取 $\boldsymbol{F}$ 中各个像素的最不重要位(LSB)F_i,必然有 $F_i \in (0,1)$,$1 \le i \le n$,设置一个加权矩阵:

$$\boldsymbol{K} = [m_1, m_2, \cdots, m_i, \cdots, m_n], m_i \in N, i = 1,2,\cdots,n \tag{1}$$

$\boldsymbol{K}$ 的作用是赋予 $\boldsymbol{F}$ 中每个像素不同的权值,以便将 $\boldsymbol{K}$ 中的元素修改值放大相应倍数。

显然,在最多修改 r 个像素值的情况下,总共有 T 种可能组合,$T = C_n^0 + C_n^1 + C_n^2 + \cdots + C_n^r$,这 T 种状态最多可以表达的二进制位数为 $w = \lfloor \mathrm{lb}\ T \rfloor$,分析以下3种情况:

1)当 $r = n$ 时,$T = 2^n$,$w = n$,即待嵌入的密文以 n 位为单位嵌入;

2)当 $r = 1$,$\boldsymbol{K} = [1,2,3,\cdots,n]$ 时,结果与文献[6]一样;

3)当 $r = 1$,$n = 1$,$\boldsymbol{K} = [1]$ 时,结果与单个像素的 LSB 嵌入一样。

1.2 信息的嵌入与提取

提出的算法的信息嵌入主要步骤如下:

步骤1:$\boldsymbol{K}$ 与 $\boldsymbol{F}$ 点乘后之和再作模 2^w 运算,结果为 f,即 $f = \mathrm{mod}(\sum_1^n (F_i \cdot m_i), 2^w)$,将 f 转换成二进制位串。

步骤2:计算 $diff = s - f$。如果 $diff = 0$,则不需要修改任何像素;如果 $diff < 0$,则将 $diff$ 加上 2^w,即 $diff = diff + 2^w$。显然 $0 < diff < 2^w$。

步骤3:从加权矩阵 $\boldsymbol{K}$ 中寻找最多 r 个点,使这 r 个点之和等于差值 $diff$。并在 $\boldsymbol{F}$ 中找到与这 r 个点位置对应的像素,每个像素分别加1,如果加1后产生溢出,需将 $\boldsymbol{F}$ 中该像素值末位置0,再返回步骤1,直到完成信息的嵌入。

密文信息的提取过程只需要将 $\boldsymbol{F}$ 与 $\boldsymbol{K}$ 点乘后之和再模 2^w,结果就是密文。

2 加权矩阵的构建

加权矩阵 $\boldsymbol{K}$ 的构造,是提出算法的关键。

2.1 最优加权矩阵及其不可构建性

当最多修改2个像素,即 $r = 2$ 时,最理想的嵌入结果是:

$$w_2 = \lfloor \mathrm{lb}(C_n^0 + C_n^1 + C_n^2) \rfloor = \lfloor \mathrm{lb}\ \frac{n^2 + n + 2}{2} \rfloor \tag{2}$$

不失一般性,设 $\boldsymbol{K}$ 为一个行向量,取值见式(1)。隐藏容量要达到 w_2,要求设计的加权矩阵 $\boldsymbol{K}$ 满足以下条件:

①$\boldsymbol{K}$ 的大小与像素块 $\boldsymbol{F}$ 一致,元素个数为 n,且都是正整数。

② 必须保证 $\boldsymbol{K}$ 中任取1个元素或任取2个元素之和能组成连续的 $1 \sim \frac{n^2 + n + 2}{2}$ 中的任意数字。

③ 为了达到最优,必须保证 $\boldsymbol{K}$ 中任取1个元素或任取2个元素之和都不能相等,即必须同时满足 $m_i + m_j \ne m_p + m_q$,$m_i \ne m_p + m_q$,$m_i \ne m_j$,其中,$i, j, p, q \in \{1,2,\cdots,n\}$,且 $i \ne j$,$p \ne q$。

下面证明不可能构建这样的最优矩阵。

证明:

不失一般性,设图像分块 $\boldsymbol{F}$ 的大小 n 足够大,且当 $p < q$ 时,$m_p < m_q$。

为了达到上述条件②,$\boldsymbol{K}$ 中必须要有1、2,没有其他元素可以组合成1和2,因此,可得 $m_1 = 1$,$m_2 = 2$。

因为要保证连续,则下一个要组成的数字为3,又因为 $m_1 + m_2 = 3$,故 m_3 不能取3,否则,与上述条件③矛盾。

再下一个要组成的数字为4,无法由 m_1、m_2 组成,故只能取 $m_3 = 4$。

接下来,要组成的数字是5、6、7,因为 $m_1 + m_3 = 1 + 4 = 5$,$m_2 + m_3 = 2 + 4 = 6$,但7无法由前面的 m_1、m_2、m_3 中任取不多于2个元素相加组成,故只能取 $m_4 = 7$。

然后,要组成的数字是8,9,10,因为 $m_1 + m_4 = 1 + 7 = 8$,$m_2 + m_4 = 2 + 7 = 9$,但10无法由前面的

m_1、m_2、m_3、m_4 中任取不多于 2 个元素相加组成,故只能取 $m_5 = 10$。

这样,$\boldsymbol{K}$的前5个元素 $m_1,\cdots,m_5$ 分别为1、2、4、7、10,显然 $m_1 + m_5 = 1 + 10 = 11$,$m_3 + m_4 = 4 + 7 = 11$,数字 11 可由 2 个元素组合而成,违背了条件 ③。

证毕。

2.2　次优加权矩阵的构造算法

因为无法构造最优矩阵 $\boldsymbol{K}$,尝试构造次优矩阵。$\boldsymbol{K}$ 的元素个数等于分块 $\boldsymbol{F}$ 的像素个数,不失一般性,设 $\boldsymbol{K}$ 为一个行向量,$\boldsymbol{K}$ 的取值见式(1),且当 $p < q$ 时,$m_p < m_q$,参数 r 满足 $1 < r < n$。构建加权矩阵 $\boldsymbol{K}$ 如下:

前面 r 个元素取连续的正整数,第 $r+1$ 到 n 个元素变成首项为 $2r$、公差为 $r+1$ 的等差数列。即

$$\boldsymbol{K} = [m_1, m_2, \cdots, m_n] = [1,2,3,\cdots,r-1,r,2r,3r+1,4r+2,5r+3,\cdots,2r+(n-r-1)\times(r+1)]。$$

在该数列保证满足上述条件 ① 和 ② 的情况下,组成最大的连续正整数为:

$$Z = m_r + m_n = r + (2r + (n - r - 1) \times (r + 1))。$$

为了能隐藏更多的数据,应使 Z 最大化,即

$$Z = -r^2 + (n+1)r + n - 1 = -\left(r - \frac{n+1}{2}\right)^2 + \frac{n^2 + 6n - 3}{4}。$$

当 $r = \dfrac{n+1}{2}$(n 为奇数)或 $r = \dfrac{n}{2}$ 或 $r = \dfrac{n+2}{2}$(n 为偶数)时,Z 取最大值 $\dfrac{n^2 + 6n - 3}{4}$。

因此,在所构造的次优加权矩阵下,能隐藏的最多位数为:

$$w_1 = \lfloor \mathrm{lb}\, \frac{n^2 + 6n - 3}{4} \rfloor \tag{3}$$

2.3　次优和最优加权矩阵的比较

根据式(2)可知,最优加权矩阵的隐藏容量为 w_2,设最优与次优加权矩阵的隐藏容量差值为:

$$c = w_2 - w_1 = \lfloor \mathrm{lb}\, \frac{n^2 + n + 2}{2} \rfloor - \lfloor \mathrm{lb}\, \frac{n^2 + 6n - 3}{4} \rfloor。$$

通过实验分析可知,在 $n \le 25\,000\,000$ 时,c 取最大值 1,即在分块大小不大于 2 500 万像素时,提出算法所构造的次优 $\boldsymbol{K}$ 与最优情况相比,信息隐藏的容量最大相差 1 个二进制位。

图 1 为 $n \le 100$ 时提出算法构造的次优矩阵与最优矩阵可隐藏信息位的比较。

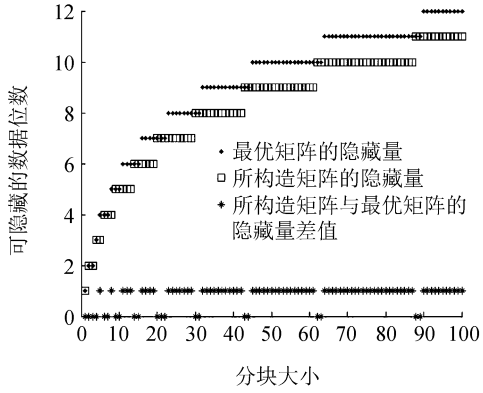

图 1　提出算法构造的矩阵与最优矩阵的隐藏容量比较

Fig. 1　Comparison of the hiding capacity between the proposed matrix and the optimum matrix

由图 1 可知,当 $c = 0$,即 $n = 1,2,3,4,6,7,9,10,\cdots,89$ 时,提出算法构造的矩阵与最优情况下的隐藏容量是一样的,其余情况下隐藏容量少 1 位。

3　实验与结果分析

根据前文和图 1 可知,当 $n = 9$ 时,提出的算法构造的加权矩阵与最优情况下的隐藏容量是一样的。现假设分块大小为 9,分块尺寸为 3 × 3。

3.1　加权矩阵 $\boldsymbol{K}$ 的构造

取 $r = \dfrac{n+1}{2} = \dfrac{9+1}{2} = 5$,故加权矩阵 $\boldsymbol{K} = [1,2,3,4,5,10,16,22,28]$,则 $w_1 = \lfloor \mathrm{lb}\, \dfrac{n^2 + 6n - 3}{4} \rfloor = \lfloor \mathrm{lb}\, \dfrac{9^2 + 6 \times 9 - 3}{4} \rfloor = \lfloor \mathrm{lb}\, 33 \rfloor = 5$,即一个图像块有 9 个像素,在最多改变 2 个像素的情况下,提出的算法可以隐藏 5 位二进制数。

3.2　信息隐藏过程

假设要隐藏的信息 $s = (11011)_2 = (27)_{10}$,且

$$\boldsymbol{F} = \begin{bmatrix} 165 & 166 & 165 \\ 173 & 172 & 176 \\ 175 & 174 & 177 \end{bmatrix}, \boldsymbol{K} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 10 \\ 16 & 22 & 28 \end{bmatrix}。$$

具体的隐藏过程如下:

步骤 1:将 $\boldsymbol{F}$ 中所有元素先做模 $32(2^5 = 32)$ 运算,目的是为了减少后面的计算量,得

$$\boldsymbol{F}_2 = \begin{bmatrix} 5 & 6 & 5 \\ 13 & 12 & 16 \\ 15 & 14 & 17 \end{bmatrix}。$$

步骤 2:$\boldsymbol{K}$ 与 $\boldsymbol{F}_2$ 做点乘得 $\boldsymbol{F}_3$,即

$$\boldsymbol{F}_3 = \boldsymbol{K} \cdot \boldsymbol{F}_2 = \begin{bmatrix} 5 & 12 & 15 \\ 52 & 60 & 160 \\ 240 & 308 & 476 \end{bmatrix}。$$

对 $\boldsymbol{F}_3$ 的所有元素求和并做模32运算得:

$$f = \mathrm{mod}\left(\sum_1^9 (\boldsymbol{F}_3), 32\right) = \mathrm{mod}(1\,328, 32) = 16。$$

步骤3:计算 $diff = s - f = 27 - 16 = 11$。

步骤4:在 $\boldsymbol{K}$ 中找一个元素等于 $diff$,如果没有,就找2个元素之和等于 $diff$。本例中,可找到 $K_{11} + K_{23} = 1 + 10 = 11$。

步骤5:修改 $\boldsymbol{F}$ 中元素位置与 K_{11}、K_{23} 对应的元素,即 F_{11}、F_{23},并将这2个元素的像素值分别加1,得到藏秘后的图像块 $\boldsymbol{F}_4 = \begin{bmatrix} 166 & 166 & 165 \\ 173 & 172 & 177 \\ 175 & 174 & 177 \end{bmatrix}$。如果需要修改的像素值为255,则需要将原图 $\boldsymbol{F}$ 中的该像素值置为254,再重新返回步骤1,直到分块中要修改的像素值没有255为止。

3.3 信息提取过程

信息提取主要有以下步骤:

步骤1:得到藏秘图像块 $\boldsymbol{F}_4$,将 $\boldsymbol{F}_4$ 中所有元素先做模32运算,得 $\boldsymbol{F}_2' = \begin{bmatrix} 6 & 6 & 5 \\ 13 & 12 & 17 \\ 15 & 14 & 17 \end{bmatrix}$。

步骤2:计算 $\boldsymbol{K}$ 与 $\boldsymbol{F}_2'$ 点乘得 $\boldsymbol{F}_3'$,即

$$\boldsymbol{F}_3' = \boldsymbol{K} \cdot \boldsymbol{F}_2' = \begin{bmatrix} 6 & 12 & 15 \\ 52 & 60 & 170 \\ 240 & 308 & 476 \end{bmatrix},$$

对 $\boldsymbol{F}_3'$ 的所有元素求和并做模32运算,得:

$$f = \mathrm{mod}\left(\sum_1^9 \boldsymbol{F}_3', 32\right) = \mathrm{mod}(1\,399, 32) = 27,$$

即是密文 s。

详细的隐藏和提取过程,如图2所示。

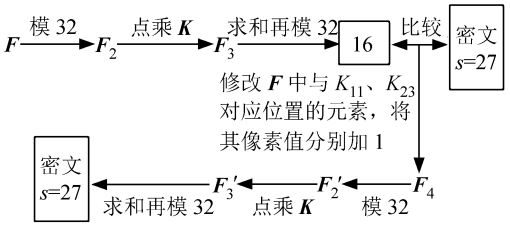

图2 分块大小为9时信息隐藏和提取过程

Fig. 2 Information embedding and extracting when the size of the block is 9

3.4 实验结果及分析

分别选取了经典的测试图像集中的lena(512×512)、fruits(512×480)等十几张图片,同时使用了几十幅自然图像,实验所采用的部分载体图像如图3所示。

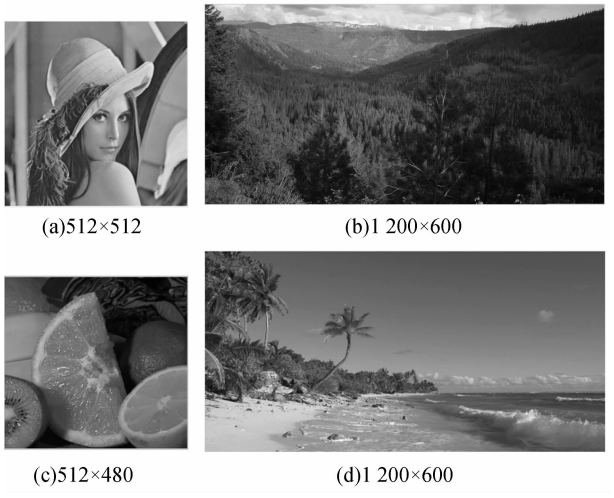

图3 部分载体图像

Fig. 3 Some host images

隐藏的信息主要是二值图像(如签名等),实验所用部分密文图像如图4所示。

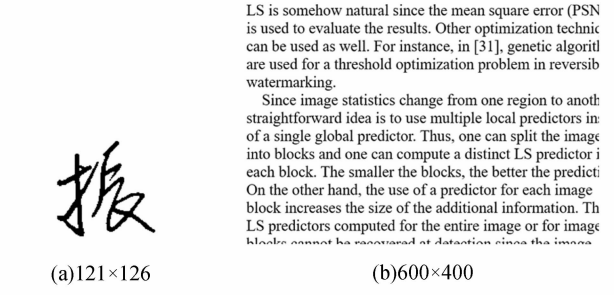

图4 部分密文信息

Fig. 4 Some cipher text information

选取部分实验结果如表1所示。由表1可知,提出的算法具有较高的隐藏容量,相比文献[6],提出算法的最高隐藏量平均高出66.7%,同时,提出算法的 $PSNR$(峰值信噪比)平均只降低1.202%。

根据式(3),当分块 $\boldsymbol{F}$ 的大小为 n 时,可隐藏的信息位数为 w_1,其最多可表达 2^{w_1} 种状态。故分块不用修改的概率为 $1/2^{w_1}$;修改1个像素的概率为 $C_n^1/2^{w_1}$;修改2个像素的概率为 $1 - (1/2^{w_1} + C_n^1/2^{w_1})$。当 n 增加时,不用修改的概率呈指数级速度减少。即说明密文信息的序列越长,载体图像与其匹配的可能性越小。

表 1　部分实验结果

Tab. 1　Some experimental results

算法	载体图像	密图	最高隐藏量/位	隐藏率/(位·像素$^{-1}$)	分块尺寸	分块隐藏量/位	所用像素块	修改的像素个数	*MSE*	*PSNR*
文献[6]方法	3(a)	4(a)	86 700	0.058 2	3×3	3	5 082	4 410	0.016 8	65.871 8
提出的方法			144 500	0.058 2	3×3	5	3 049	5 042	0.019 2	65.290 2
文献[6]方法	3(b)	4(a)	81 600	0.058 2	3×3	3	5 082	4 415	0.018 0	65.586 6
提出的方法			136 000	0.058 2	3×3	5	3 049	5 055	0.020 6	64.998 7
文献[6]方法	3(c)	4(b)	240 000	0.333 3	3×3	3	80 000	70 115	0.097 4	58.246 0
提出的方法			400 000	0.333 3	3×3	5	48 000	79 542	0.110 5	57.698 2
文献[6]方法	3(c)	4(c)	240 000	0.244 4	3×3	3	58 666	51 399	0.071 4	59.594 6
提出的方法			400 000	0.244 4	3×3	5	35 200	58 279	0.080 9	59.049 0
文献[6]方法	3(d)	4(a)	240 000	0.021 2	3×3	3	5 082	4 450	0.006 2	70.220 5
提出的方法			400 000	0.021 2	3×3	5	3 049	5 021	0.007 0	69.696 2

4　结　论

与文献[6]相比,提出算法构造的次优加权矩阵,在 *PSNR* 降低很少的情况下,信息的隐藏容量得到很大的提升,该算法能很方便地应用于彩色图像。

要达到式(2)中的容量,只需满足从加权矩阵 $\boldsymbol{K}$ 中任取不多于2个元素即可组成2^w种组合,2.1节中 ①②③ 是充分非必要条件,为进一步探索更优的加权矩阵构造算法提供了可能。

但是,论文也有一些不足之处,如:怎样构建更好的加权矩阵 $\boldsymbol{K}$?如何证明 $\boldsymbol{K}$ 是最优的?只在理论和实践上研究了 $r=2$ 的情况。当 $2<r\leq n$ 时,加权矩阵的构造更为复杂,而且随着 n 的增大,其计算量也进一步增大,n 与 r 如何取值才能达到隐藏信息与计算量的最优?隐藏容量与不可感知性是否有内在的某种平衡关系?这些内容还值得进一步研究。

参考文献:

[1] 葛秀慧,田浩,郭立甫,等. 信息隐藏原理及应用[M]. 北京:清华大学出版社,2008.

[2] Pan H K, Chen Y Y, Tseng Y C. A secure data hiding scheme for two-color images[C]//Proceedings of the Fifth IEEE Symposium on Computers and Communications (ISCC 2000), 2000. Antibes-Juan les Pins: IEEE, 2000: 750 – 755.

[3] Nguyen T D, Arch-int S, Arch-int N. An adaptive multi bit-plane image steganography using block data-hiding [J/OL]. Multimedia Tools and Applications, 2015: 1 – 27 [2015 – 07 – 02]. http://link.springer.com/article/10.1007%2Fs11042 – 015 – 2752 – 9. DOI: 10.1007/s11042 – 015 – 2752 – 9.

[4] Yang C H, Weng C Y, Wang S J, et al. Adaptive data hiding in edge areas of images with spatial LSB domain systems [J]. IEEE Transactions on Information Forensics and Security, 2008, 3(3): 488 – 497.

[5] Lu T C, Chang C C, Huang Y H. High capacity reversible hiding scheme based on interpolation, difference expansion, and histogram shifting[J]. Multimedia Tools and Applications, 2014, 72(1): 417 – 435.

[6] Guo Meng, Zhang Hongbin, Wei Lei. Data hiding in binary images[J]. Acta Electronica Sinica, 2009, 37(11): 2409 – 2415. [郭萌,张鸿宾,魏磊. 二值图像中的数据隐藏算法[J]. 电子学报,2009,37(11):2409 – 2415.]

[7] Parah S A, Sheikh J A, Bhat G M. Data hiding in intermediate significant bit planes, a high capacity blind steganographic technique [C]//Proceedings of the International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET). Tiruchirappalli, Tamilnadu, India: IEEE, 2012: 192 – 197.

[8] Qin C, Chang C C, Chen Y C. Efficient reversible data hiding for VQ-compressed images based on index mapping mechanism[J]. Signal Processing, 2013, 93(9): 2687 – 2695.

[9] Yamawaki K, Nakano F, Noda H, et al. Improvement of JPEG compression efficiency using information hiding and image restoration [J]. IEICE Transactions on Information and Systems, 2013, E96-D(5): 1233 – 1237.

[10] Wang Kan. High capacity lossless data hiding technique for compressed domains[D]. Guangzhou: South China University of Technology, 2013. [王衍. 压缩域图像大容量无损信息隐藏技术研究[D]. 广州:华南理工大学,2013.]

(编辑　赵　婧)