doi:10.3969/j.issn.1001-4616.2024.04.014

基于预测误差绝对值标记的加密图像可逆信息隐藏方法

项　聪,朱　毅,陶永鹏,王兴田

(大连外国语大学,辽宁 大连 116044)

[摘要]　随着云计算的发展,加密图像可逆信息隐藏技术得到了越来越多研究者的关注,并广泛应用于司法、军事、医疗等领域. 本文提出了一种基于预测误差绝对值标记的加密图像可逆信息隐藏算法. 首先,使用中值边缘检测器获取原始图像像素预测值,并计算预测误差的绝对值;然后根据其二进制序列从最高有效位到最低有效位存在多个连续 0 值的分布状况,构建最优二叉树以生成自适应的哈夫曼编码;最后通过哈夫曼编码标记对应的预测误差绝对值像素,以腾出空间进行信息嵌入. 实验结果表明:该算法能够实现信息的完整提取和图像的无损恢复,与同类算法相比,能够有效地提高图像机密信息的嵌入率.

[关键词]　加密图像,可逆信息隐藏,预测误差绝对值,哈夫曼编码
[中图分类号]TP391　[文献标志码]A　[文章编号]1001-4616(2024)04-0126-09

Reversible Data Hiding Method in Encrypted Images Based on Prediction Error Absolute Value Labeling

Xu Cong,Zhu Yi,Tao Yongpeng,Wang Xingtian

(Dalian University of Foreign Language,Dalian 116044,China)

Abstract:With the development of cloud computing,reversible data hiding technology for encrypted images has attracted more and more attention from researchers and has been widely applied in the fields of justice,military,and medical treatment. This paper proposes an algorithm for reversible data hiding in encrypted images based on prediction error absolute value labeling. Firstly,the median edge detector is used to obtain the pixel prediction value of the original image and calculate the prediction error absolute value;Then,according to the distribution of multiple consecutive zeros in its binary sequence from the most significant bit to the least significant bit,the optimal binary tree is constructed to generate adaptive Huffman coding. Finally,the corresponding prediction error absolute value pixels are labeled by Huffman coding to make room for data embedding. Experimental results show that the algorithm can achieve complete data extraction and lossless image restoration,and can effectively improve the embedding rate of image confidential data compared with similar algorithms.

Key words:encrypted image,reversible data hiding,prediction error absolute value,Huffman coding

可逆信息隐藏(reversible data hiding,RDH)技术是将信息嵌入到载体图像中,并能无误地进行信息提取,无损地恢复原始图像[1-2]的技术. 常用的 RDH 方法有 3 种:基于无损压缩的可逆信息隐藏[3-4],基于直方图平移的可逆信息隐藏[5]和基于预测误差扩展的可逆信息隐藏[6-7]. 随着云计算和大数据的发展,人们更加关注隐私保护、数据安全和多媒体认证[8]. 加密图像的可逆信息隐藏(reversible data hiding in encrypted images,RDHEI)作为信息隐藏技术的一个分支,能够实现在加密图像中嵌入机密信息;在解密阶段,能够无误地提取机密信息,并且无损地恢复原始图像. 因此在云存储、医疗系统等注重隐私保护的领域中得到了广泛应用[9-12].

现有的 RDHEI 方法主要有:加密后腾出信息嵌入空间(vacating room after encryption,VRAE)和加密

收稿日期:2024-03-20.
基金项目:辽宁省教育厅科研项目(LJKMZ20221552).
通讯作者:项聪,副教授,研究方向:信息隐藏、图像安全. E-mail:xucongdlmu@163.com

前预留出可嵌空间(vacating room before encryption,VRBE)两大类.

Zhang[13]首先提出 VRAE 算法,将原始图像加密,分成若干图像块,并将图像块中半数像素的最后 3 个有效位翻转,嵌入 1 个信息位,但该算法的嵌入容量较低. Qin 等[14]通过置乱图像的位平面、图像的分块、块内像素的位置等方式加密图像,然后使用稀疏矩阵编码压缩加密后的图像,获取嵌入空间,该算法实现了信息提取和图像的无损恢复,并在一定程度上提高了信息的嵌入容量. 然而在基于 VRAE 的 RDHEI 算法中,由于图像经过加密后使得像素的相关性降低,难以实现更多机密信息的嵌入.

为了进一步提高 RDHEI 的嵌入容量,VRBE 类可逆信息隐藏算法被提出. 文献[15]中首次提出了 VRBE 算法,该算法利用预测差值的直方图位移方式在图像加密前预留空间. 文献[16-19]采用了对图像像素的最高有效位(most significant bit,MSB)预测的方式预留嵌入空间. 文献[20]进一步提出了一种基于像素预测和多个高阶有效位平面重排的算法,该算法通过对预测误差位平面分块后以重排的方式预留出可嵌空间,使得嵌入容量有了明显的提升. Yin 等[21]提出通过中值边缘检测器(median edge detector,MED)计算像素的预测值,并将预测值与原始值的二进制序列从 MSB 到最低有效位(least significant bit,LSB)进行比较,对相同的比特位数,用 Huffman 编码进行标记以预留嵌入空间. 在文献[21]的基础上,文献[22]进一步分析了预测值与原始值的二进制序列比较后所得的比特位数值的分布权重,改进了 Huffman 编码的生成算法,进一步提升了信息的嵌入容量. 然而这两种算法采用的像素原值和预测值的二进制序列的比较方式,仍然会造成较大的空间浪费.

为了进一步地提高图像的嵌入容量,本文提出了一种基于预测误差绝对值标记的 VRBE 算法. 该算法使用 MED 获取原始图像像素预测值,并计算像素的预测误差绝对值(prediction error absolute value,PEAV). 由于邻近像素的相关性,生成的预测误差绝对值图像其直方图分布具有明显的统计特征,通常会在较低差值附近出现最高峰值,并随着差值的增大而逐渐减小. 因此通过使用不等长的标签来标记不同的预测误差绝对值,就能够在图像加密前预留出大量可嵌空间,从而提高机密信息的嵌入量.

1　算法概述

算法的结构如图 1 所示. 整个过程可分为自适应预测误差绝对值标记(adaptive prediction error absolute value labeling,APEAVL)、图像加密、数据隐藏、数据提取和图像复原 4 个阶段.

(1)APEAVL 阶段:图像所有者计算原始图像 I 的预测误差绝对值,采用 APEAVL 对预测差值绝对值像素进行自适应标记.

(2)图像加密阶段:图像所有者使用标准流加密算法对预测误差绝对值图像 I_{P} 进行加密,生成加密图

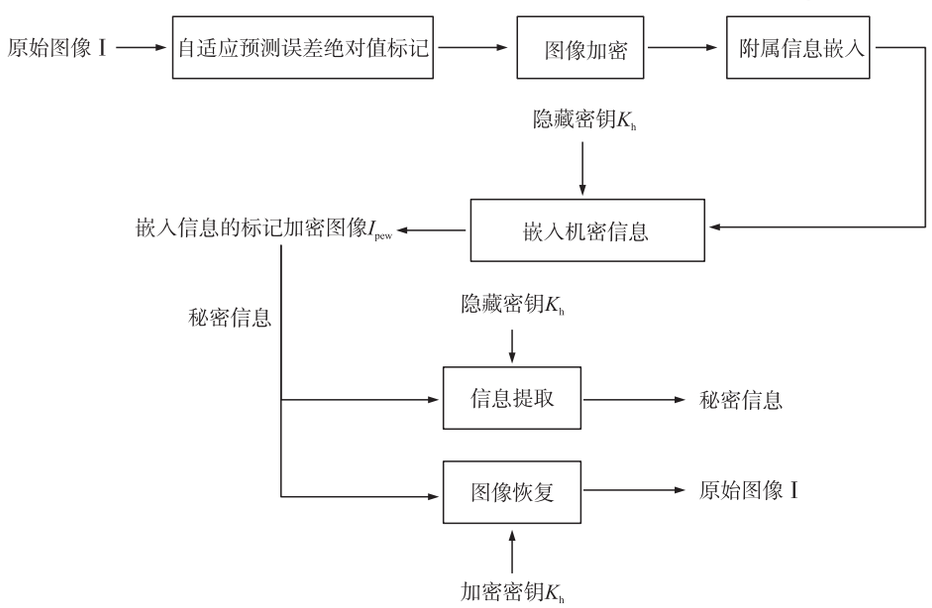


图 1　算法框架结构图

Fig. 1　Algorithm framework structure diagram

像 I_{pe},在加密图像中采用位替换的方式嵌入附加信息,生成嵌入附加信息的标记加密图像 I_{pem}.

(3)机密数据隐藏阶段:在加密图像 I_{pem} 中,信息隐藏者将经过信息隐藏密钥加密后的机密信息嵌入到经过 APEAVL 的 I_{pem} 中的可嵌像素的预留空间中,生成标记的加密图像 I_{pew}.

(4)数据提取和图像复原阶段:图像接收者基于拥有的不同密钥分离式完成机密数据的提取或原始图像 I 的无损恢复.

2　基于预测误差绝对值标记 RDHEI 算法

2.1　像素预测误差绝对值计算

设原始图像 I 大小为 $m\times n$,其中第一行和第一列像素作为参考像素保持不变,其余像素的预测值通过 MED 获得. 预测的像素和周边像素位置关系如图 2 所示,根据式(1)可计算 $X(i,j)$ 的预测值 $\hat{X}(i,j)$.

$$\hat{X}(i,j)=\begin{cases}\max(a,b), & c\leqslant\min(a,b)\\ \min(a,b), & c\geqslant\max(a,b)\\ a+b-c, & \text{otherwise}\end{cases}\tag{1}$$

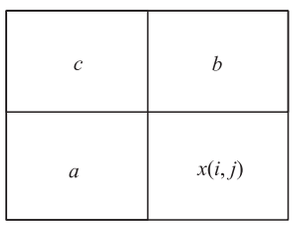


图 2　预测的像素和周边像素位置关系图

Fig. 2　Position relationship diagram between predicted pixel and surrounding pixels

根据式(2)计算像素 $X(i,j)$ 的预测误差绝对值 $\mathrm{ea}(i,j)$.

$$\mathrm{ea}(i,j)=|x(i,j)-\hat{x}(i,j)|\tag{2}$$

其中,$2\leqslant i\leqslant m$,$2\leqslant j\leqslant n$.

原始图像 I 经过像素预测误差绝对值计算后生成预测误差绝对值图像 I_p(I_p 中的第一行第一列像素由 I 的参考像素构成,其余位置像素则由 I 对应位置像素的预测误差绝对值构成).

2.2　预测误差绝对值像素最大可嵌空间计算

图像 I 的预测误差绝对值在不同取值范围内,其二进制序列从 MSB 到 LSB 存在多个连续的 0 值. 如表 1 所示.

表 1　预测误差绝对值的二进制序列连续 0 值表

Table 1　PEAV binary sequence continuous 0-valute table

图像 I 预测误差绝对值 PEAV 的取值范围	PEAV 的二进制序列从 MSB 到 LSB 的连续 0 值个数 Q	图像 I 预测误差绝对值 PEAV 的取值范围	PEAV 的二进制序列从 MSB 到 LSB 的连续 0 值个数 Q
PEAV=0	8	15<PEAV<=31	3
PEAV=1	7	31<PEAV<=63	2
1<PEAV<=3	6	64<PEAV<=127	1
3<PEAV<=7	5	127<PEAV<=255	0
7<PEAV<=15	4		

可以看到随着预测误差绝对值的增大,连续 0 值个数逐渐减小,当预测误差绝对值在 31 以下时,其对应的连续 0 值个数保持在 3 个以上. 由于原始图像中相邻像素的空间相关性和 MED 的准确性,大量像素的预测误差绝对值会出现在 0 值附近. 使得图像 I_p 的直方图服从类似位置参数为 0 的单边拉普拉斯分布,这样的统计学特性可以被充分考虑和利用. 通过压缩编码去标记图像中不同预测误差绝对值中的连续 0 值,可以预留出较高的可嵌空间.

在图像 I_p 中,除了第一行第一列的参考像素外,其余像素的二进制序列从 MSB 到 LSB 存在连续的 Q 个 0,第 Q+1 位为反向值 1,因此这前 Q+1 位比特($0_1 0_2\cdots 0_q 1_{q+1}$)构成了像素的冗余空间,可用于信息嵌入,根据式(3)计算出像素的最大可嵌空间为:

$$E(i,j)=\min(Q(i,j)+1,8),\tag{3}$$

式中,($2\leqslant i\leqslant m$,$2\leqslant j\leqslant n$,$0\leqslant Q(i,j)\leqslant 8$),$Q(i,j)$ 表示像素 $I_P(i,j)$ 的二进制序列从 MSB 到 LSB 的连续 0 值个数.

2.3　图像 Huffman 编码

Huffman 编码通过使用可变长度编码对源输入符号进行编码,其中短编码用于表示出现概率高的符号,长编码用于表示出现概率低的符号. 这种编码方式可以减少用于所有符号编码的编码字符串,实现无

损压缩数据的目的. 本文通过构建最优二叉树的方式生成输入源符号的 Huffman 编码.

下面举例说明构建最优二叉树的步骤如下:

设有 5 个符号节点 ABCDE,将其作为叶子结点,假设其权重分别是 3、4、6、8、9,依据权重的大小将具有最小和次最小权重两个节点 A、B 的权重相加,生成权重等于 7 的节点,将该节点和余下的 3 个节点 E、C、D 再次依据权重大小进行最小和次最小权重节点的选择,生成新一轮参与选择的权重等于 13 的节点. 重复执行以上操作,直到生成具有最大权重 30 的根节点. 将各节点连接起来生成最优的二叉树,结果如图 3 所示.

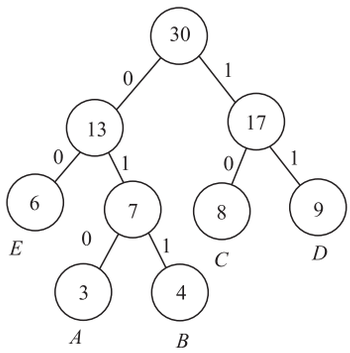


图 3　Huffman 树的生成图

Fig. 3　Huffman tree spanning graph

图 3 示例中圆圈内的值表示该节点的权重. 从根节点到要编码的叶子节点的 01 序列则为该叶子节点的 Huffman 编码.

根据上述方法,完成对图像预测误差绝对值的标记计算:

首先:对图像 I_{p} 中 Q 值(与表 1 的 Q 值含义相同)的分布进行统计,并以 Q 的不同取值(从 0 到 8)作为要编码的叶子节点,以该取值下的统计个数作为叶子节点的权重,按照上述方法构建最优二叉树,从而生成不同 Q 值所对应的 Huffman 编码. 表 2 给出图像 Lena 的 Huffman 编码示例.

表 2　Lena 的预测误差绝对值对应的 Huffman 编码表

Table 2　Lena prediction error absolute value Huffman coding table

PEAV 的二进制序列从 MSB 到 LSB 的连续 0 值个数 Q	分布的个数统计值	Huffman 编码	PEAV 的二进制序列从 MSB 到 LSB 的连续 0 值个数 Q	分布的个数统计值	Huffman 编码
0	6 022	10000	5	31 493	101
1	9 926	11110	6	32 983	110
2	11 053	11111	7	20 623	1110
3	8 796	10001	8	125 104	0
4	15 121	1001			

其次:由于预测误差绝对值图像中除参考像素外,其余像素为图像 I 的预测误差绝对值,其 8 位比特中的前 Q+1 位为($0_1 0_2 \cdots 0_q 1_{q+1}$),所以通过使用 Q 值所对应的 Huffman 编码以位替换的方式替换前 Q+1 位比特中的部分比特位,可实现对预测误差绝对值像素的有效标记(接收方通过识别 Huffman 编码从而确定 Q 值,并进一步提取信息或复原图像),具体标记过程见文中 2.5.

2.4　图像加密

使用流加密方法加密图像 I_p.

首先,由加密密钥 K_e 生成一个伪随机矩阵 $\boldsymbol{R}$(大小为 $m\times n$),接下来,将图像 I_p 的像素 $I_p(i,j)$ 及其对应矩阵 $\boldsymbol{R}$ 中的元素 $r(i,j)$,根据式(4)、(5)转换为 8 位二进制序列,并标记为 $I_P^K(i,j)$ 和 $r^K(i,j)$.

$$I_p^k(i,j)=\left\lfloor \frac{I_p(i,j) \bmod 2^{9-k}}{2^{8-k}} \right\rfloor, k=1,2,3,\cdots,8 \tag{4}$$

$$r^k(i,j)=\left\lfloor \frac{r(i,j) \bmod 2^{9-k}}{2^{8-k}} \right\rfloor, k=1,2,3,\cdots,8 \tag{5}$$

根据式(6)将 $I_p^k(i,j)$ 和 $r^k(i,j)$ 进行异或运算.

$$I_{\mathrm{pe}}^k=I_p^k(i,j)\oplus r^k(i,j), k=1,2,3,\cdots,8 \tag{6}$$

根据式(7)将异或结果转换十进制,生成加密图像 I_{pe}

$$I_{\mathrm{pe}}(i,j)=\sum_{k=1}^{k=8} I_{\mathrm{pe}}^k(i,j)\times 2^{8-k}, k=1,2,\cdots,8 \tag{7}$$

2.5　附加信息嵌入

为了后续原始图像 I 的恢复,需要在加密的图像 I_{pe} 中嵌入必要的附加信息.

(1)Huffman 编码表 H:根据图像 I_{p} 不同 Q 值对应的 Huffman 编码(通过文中 2.3 方式生成),构建

Huffman 编码表 H. 由于 Huffman 编码的自适应性,不同图像的编码并不相同,构建规则如下:按照“4 位编码长度+Huffman 编码”的方式依次存储 Q 值从 0 到 8 的 Huffman 编码. 其中:4 位编码长度是用 4 位比特存储 Q 值的 Huffman 编码的长度值,如果该 Q 值不存在,则 4 位编码长度用“0000”表示. 例如根据表 2 的内容,结合上述规则构建图像 Lena 表 H 的各个 Q 值编码,结果如表 3 所示.

表 3　Lena 编码表 H 中的各个 Q 值编码

Table 3　Lena encoding table H for each Q-value encoding

Q 值	Huffman 编码	Huffman 编码长度	Q 值编码	Q 值	Huffman 编码	Huffman 编码长度	Q 值编码
0	10000	0101	010110000	5	101	0011	0011101
1	11110	0101	010111110	6	110	0011	0011110
2	11111	0101	010111111	7	1110	0100	01001110
3	10001	0101	010110001	8	0	0001	00010
4	1001	0100	01001001				

将各个 Q 值编码依次连接生成 Lean 的编码表 H.

(2)用于描述图像 I_p 中的像素(参考像素除外)是否可嵌的比特位 $p(i,j)$. 根据式(3)可确定像素 $I_p(i,j)$ 最大可嵌空间为 $E(i,j)$,则 $E(i,j)$ 至少需要存放 1 比特的可嵌说明位和对应的 Q 值 Huffman 编码,余下的部分方可用于存放额外的信息. 因此根据式(8)确定该像素是否为可嵌像素.

$$p(i,j)=\begin{cases}1, & \text{if } L(i,j)+1<E(i,j)\\ 0, & \text{if } E(i,j)\leqslant L(i,j)+1\end{cases} \tag{8}$$

式中,$L(i,j)$ 表示像素 $I_p(i,j)$,Q 值所对应的 Huffman 编码长度.

$p(i,j)$ 为 1 表明像素 $I_p(i,j)$ 是可嵌像素,否则为不可嵌像素. 统计所有不可嵌入像素的个数,并用 18 位比特 lp 记录其值.

(3)用于描述图像 I_p 中的各个可嵌像素的 Huffman 编码 $LA(i,j)$.

(4)用于描述图像 I 的像素(参考像素除外)预测差值符号的比特流 S:根据式(9)确定图像 I 各像素对应的符号位.

$$s(i,j)=\begin{cases}1 & \text{if } x(i,j)>\hat{x}(i,j)\\ 0 & \text{if } x(i,j)<\hat{x}(i,j)\end{cases} \tag{9}$$

$s(i,j)$ 为 1 表明符号位为正值,否则为负值. 将所有 $s(i,j)$ 顺序连接起来生成二进制流 S,用 18 位比特 ls 记录 S 的长度.

(5)对图像 I_p 使用文中 2.4 中的加密算法加密生成加密图像 I_{pe},将 H、lp、ls 依次连接起来以位替换的方式保存在加密图像 I_{pe} 第 1 行和第 1 列的参考像素中,替换下的像素值则作为附加信息 $L1$.

(6)将加密图像 I_{pe} 中的所有不可嵌像素的 MSB 替换为 0,替换下的比特位依次连接作为附加信息 $L2$;可嵌像素的 MSB 替换为 1,接着 $L(i,j)$ 比特位用该像素 Q 值对应的 Huffman 编码 $LA(i,j)$ 进行替换,完成对 I_{pe} 像素的标记,剩下的冗余空间则可用做预留空间进行信息嵌入.

(7)将附加信息 $L1$、$L2$、S 连接起来并按照光栅扫描顺序依次嵌入到加密图像 I_{pe} 可嵌像素标记后剩余的冗余空间中,生成嵌有附加信息的标记加密图像 I_{pem}. 嵌入公式如(10)所示.

$$I_{\text{pem}}(i,j)=\begin{cases}1*2^7+\sum\limits_{k=1}^{L(i,j)}LA(i,j)_k\times 2^{7-k}+\sum\limits_{k=1}^{E(i,j)-1-L(i,j)}b\times 2^{7-L(i,j)-k}+I_{\text{pe}}(i,j)\bmod 2^{8-E(i,j)}, & E(i,j)\leqslant 7\\ 1*2^7+\sum\limits_{k=1}^{L(i,j)}LA(i,j)_k\times 2^{7-k}+\sum\limits_{k=1}^{7-L(i,j)}b\times 2^{7-L(i,j)-k}, & E(i,j)==8\end{cases} \tag{10}$$

式中,$I_{\text{pe}}(i,j)$ 表示加密图像 I_{pe} 的可嵌像素;$E(i,j)$ 表示可嵌像素的最大冗余空间;$LA(i,j)_k$ 表示可嵌像素的 Huffman 编码的第 k 位比特;$L(i,j)$ 表示可嵌像素的 Huffman 编码长度;$E(i,j)-1-L(i,j)$ 表示可嵌像素标记后冗余空间,用于信息嵌入;b 表示要嵌入附加信息的比特值;$I_{\text{pem}}(i,j)$ 表示嵌入附加信息后生成的标记像素.

根据式(11)计算整个图像的嵌入容量.

$$EC=\sum_{i=1}^{z}(E_i-L_i-1)-lf \tag{11}$$

式中,z 是图像 I_{pem} 中可嵌像素的总数;E_i 表示第 i 个可嵌像素的最大冗余空间;L_i 表示第 i 个可嵌像素的 Huffman 编码长度;lf 表示图像 I_{pem} 的所有附加信息总长度.

2.6　机密信息嵌入

信息隐藏者须从嵌有附加信息的标记加密图像 I_{pem} 中提取附加信息,然后进行信息嵌入.

输入:嵌有附加信息的标记加密图像 I_{pem}

输出:嵌有机密信息的标记图像 I_{pew}

(1)从 I_{pem} 中的参考像素中获取比特流 H、lp、ls,根据表 H 确定 I_{pem} 中各像素的 Huffman 编码标记. 根据 H 的长度和 lp、ls 的值确定附加信息 $L1$、$L2$、S 的长度.

(2)以光栅扫描顺序扫描 I_{pem} 各像素(参考像素除外),如果像素 $I_{pem}(i,j)$ 的 MSB 为 0 则该像素是不可嵌像素,不做处理;如果 MSB 为 1,则该像素是可嵌像素,利用表 H 获取 Huffman 编码标签,根据编码标签确定该可嵌像素对应的 Q 值及最大冗余空间,$E(i,j)$ 根据式(10)提取该像素所嵌入的附加信息. 因为步骤(1)已经确定了附加信息长度,所以可从可嵌像素中提取图像所有者嵌入的附加信息,并定位到未使用的可嵌像素.

(3)根据式(10)以位替换的方式将经过隐藏密钥 K_h 加密后机密信息嵌入到未使用的可嵌像素标记后的剩余的可嵌空间中. 生成嵌有机密信息的标记图像 I_{pew}.

2.7　信息抽取和图像恢复

2.7.1　只有隐藏密钥 K_h,获取机密信息

接收方收到加密的标记图像 I_{pew} 后,可以获取机密信息.

输入:嵌有机密信息的标记图像 I_{pew},隐藏密钥 K_h.

输出:机密信息.

(1)从 I_{pew} 图像的参考像素(第一行第一列)中根据 4 位编码长度+Huffman 编码的规则获取 Huffman 编码表 H;接着获取 18 比特的 lp 和 18 比特的 ls.

(2)使用文中 2.6 的方式定位到未嵌入附加信息的第一个可嵌像素中,根据式(10)并按照光栅扫描顺序依次获取后续所有可嵌像素嵌入的加密机密信息,使用隐藏密钥 K_h 解密后即获得原始的机密信息.

2.7.2　只有加密密钥,恢复原始图像

接收方收到加密的标记图像 I_{pew} 后,恢复原始图像.

输入:嵌有机密信息的标记图像 I_{pew},加密密钥 K_e.

输出:原始图像.

(1)接收方收到图像 I_{pew} 后,使用文中 2.6 的方式获取信息隐藏者嵌入的附加信息,并分解为 $L1$、$L2$、S 三部分.

(2)用 $L1$ 替换参考像素中的 H、lp、ls,以还原 I_{pe} 中的参考像素. 利用 K_e 生成伪随机矩阵 $\boldsymbol{R}$,根据式(6)还原 I_p 图像中参考像素原值.

(3)以光栅扫描顺序扫描 I_{pew} 中的像素(参考像素除外),根据 MSB 的值确定所有的可嵌和不可嵌像素,其中所有不可嵌像素的 MSB 依次用 $L2$ 中的比特位还原. 利用矩阵 $\boldsymbol{R}$,根据式(6)还原 I_p 图像中不可嵌入像素的原值. 可嵌像素则根据 Huffman 编码表 H,确定像素的 Huffman 编码标签,进一步确定对应的 Q 值. 如果 Q 值为 8 则像素用 8 个 0 替换;否则前 Q 个比特位用 0 替换,第 Q+1 位用 1 替换. 根据式(10),余下位置的比特值是经过加密的,并在信息嵌入过程中保持不变,因此余下位置的比特可利用矩阵 $\boldsymbol{R}$ 并根据式(6)还原,同前 Q+1 位比特组合后,则可还原 I_p 图像中可嵌入像素的原值.

(4)经过上述步骤后得到图像 I_p.

(5)图像 I_p 与原始图像 I 的参考像素相同,因此可以直接恢复原始图像的参考像素,接着按照光栅扫描顺序扫描图像 I_p 中的像素 $I_p(i,j)$(参考像素除外)并通过 MED 获取预测值,如果 $I_p(i,j)$ 的值为 0,则用预测值进行替换,恢复该位置上原始图像 I 的像素值;否则需要从附加信息流 S 中对应的顺序位获取符号

值并根据式(13)计算 $I(i,j)$,恢复该位置上原始图像 I 的像素值.

$$I(i,j)=\begin{cases}\hat{I}(i,j)+I_p(i,j), & s(i,j)==1\\ \hat{I}(i,j)-I_p(i,j), & s(i,j)==0\end{cases} \tag{13}$$

按照上述方式依次恢复原始图像 I 的所有像素,完成复原.

2.7.3　隐藏密钥和加密密钥进行数据提取和图像恢复

接收方接到加密的标记图像 I_{pew} 后,利用文中 2.7.1 获取机密信息,利用 2.7.2 恢复原始图像.

3　实验与结果分析

3.1　安全性分析

3.1.1　密钥空间分析

对于一幅大小为 $m\times n$ 的灰度图像,用长度为 $m\times m\times 8$ 的伪随机二进制序列对其进行加密,序列中的每个比特的值为 0 或 1. 密钥空间根据式(14)计算.

$$ks=2^{m\times n\times 8}. \tag{14}$$

在没有密钥的情况下,要从如此大的密钥空间获得完全正确的加密序列几乎是不可能的,因此加密算法有足够的安全性.

3.1.2　图像熵分析

图像熵能够反映图像灰度分布的离散程度,熵越大说明灰度分布越均匀,图像越安全. 理论上图像熵的最大值为 8,根据式(15)计算.

$$H(T)=-\sum_{i=0}^{255}p(x_i)\log_2(p(x_i)), \tag{15}$$

式中,T 表示有 256 个灰度等级的灰度图像,$p(X_i)$ 是灰度等级 $X_i(0\leqslant i\leqslant 255)$ 的概率. 表 4 列出了使用本文算法加密后的 6 幅灰度图像的熵.

表 4　加密图像的熵

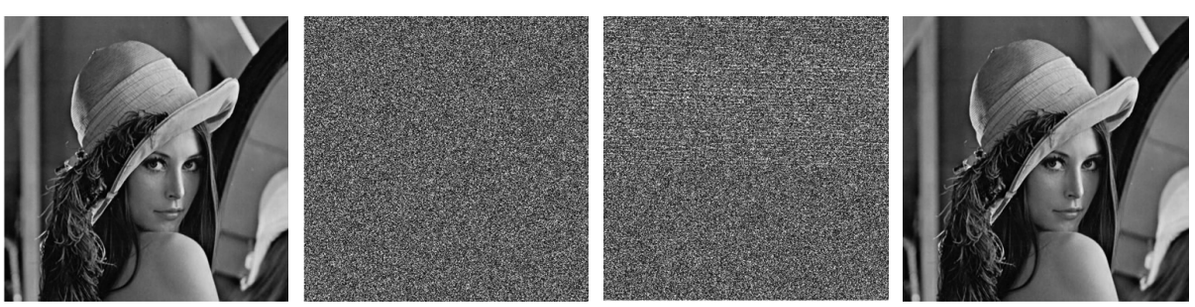
Table 4　Entropy of encrypted image

图像	加密图像的熵	嵌入机密信息标记加密图像熵	图像	加密图像的熵	嵌入机密信息标记加密图像熵
Airplane	7.999 2	7.996 9	Barbara	7.999 4	7.989 9
Lena	7.998 9	7.995 7	Boat	7.998 5	7.987 3
Peppers	7.999 2	7.992 7	Baboon	7.999 6	7.986 2

结果表明:加密图像及嵌入机密信息后的加密图像的熵值都接近 8,具有很高的安全性.

3.2　图像测试

选取 512×512 灰度图像 Lena 进行测试,结果如图 4 所示.



(a) 原始图像　(b) 加密图像　(c) 嵌入机密信息的标记图像　(d) 完全恢复的图像

图 4　图像的测试图

Fig. 4　Image test chart

图 4(a)所示为原始图像 I;图 4(b)所示为加密后得到的加密图像;图 4(c)所示为经过机密信息嵌入后标记的加密图像,其嵌入率 3.472 bpp;图 4(d)所示为经过加密密钥 K_e 复原图像,与原始图像相同,复原后的图像的 PSNR 接近+∞.

3.3　算法性能测试

为了更直观地评估嵌入容量,本文算法在 3 个图像数据库进行实验,结果如表 5 所示. 在 3 个图像数据库上,APEAVL 算法的 Average ER(Average Embedding Rate)分别为 3.982 bpp、3.784 bpp、3.191 bpp. 复原图像的 PSNR(Peak Signal to Noise Ratio)趋于 +∞,SSIM(Structural Similarity Index)等于 1,提取数据的 MSE(Mean Squared Error)为 0%.

表 5　不同数据库下的性能指标

Table 5　Performance indicators under different databases

Database	Average ER	PSNR	SSIM	MSE
BossBase	3.982	+∞	1	0%
BOWS-2	3.784	+∞	1	0%
UCID	3.191	+∞	1	0%

3.4　与同类算法的比较测试

为了验证本文算法的优越性能,与文献[20-22]进行对比实验测试. 结果如图 5、图 6 所示.

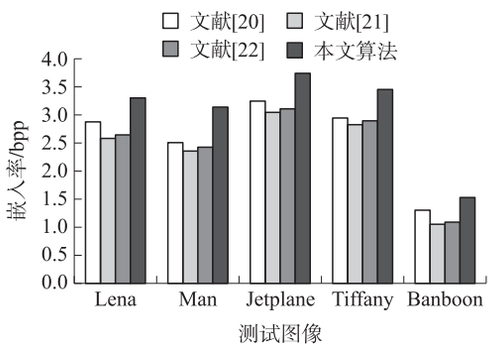


图 5　测试图像上嵌入率的比较

Fig. 5　Comparison of ER on test images

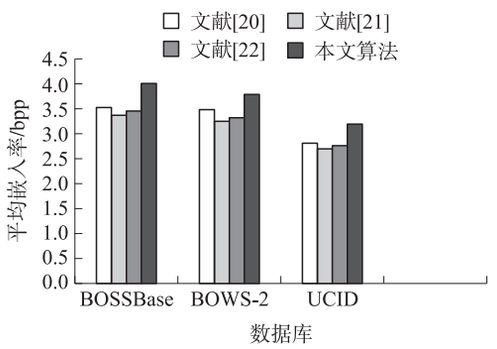


图 6　数据集上平均嵌入率的比较

Fig. 6　Comparison of average ER on three databases

结果分析:文献[20]计算原始图像的预测误差,并将其用于位平面的重排和压缩以提升图像的可嵌空间. 但该算法没有考虑到利用预测差值的分布特征,而本文算法则利用预测差值的分布特点进行编码压缩,能够进一步提高预测差值冗余空间的压缩率,获得了更高的嵌入率. Yin 等[21]在计算出像素预测值后,将预测值与原始值从 MSB 到 LSB 进行比较,个数相同的情况有 9 种,分别用预定义的 9 种 Huffman 编码表示,以预留出可嵌空间. 但是仍然存在相当部分的像素,虽然它们的预测差值较小,但预测值和原值在比较的高阶有效位上出现了不一致,致使相同比特位的计数值较低,无法生成更多的可嵌空间. 浪费了具有较小预测差值像素的冗余空间. 文献[22]在文献[21]的基础上改进 Huffman 算法,能够根据特定的图像生成自适应的 Huffman 编码以节省像素标记的长度进而提高信息的可嵌容量. 然而并没有解决文献[21]存在的空间浪费问题. 而本文算法使用 MED 获取预测差值,进而生成预测差值绝对值图像. 在充分考虑预测差值绝对值的分布统计特征后,以此为基础构建自适应的 Huffman 编码算法进行像素空间压缩. 一方面减少了复原图像所需附加信息;另一方面通过对预测差值绝对值的 Huffman 编码标注方式也在图像中预留出了更多的可嵌空间. 因此在图 5 的测试图像和图 6 的数据集的实验测试中,与上述这些算法相比取得了更好的性能指标. 说明 APEAVL 算法优于现有算法,具有更高的嵌入率.

4　结论

本文提出了一种基于预测误差绝对值标记的高容量的 RDHEI 方法. 通过构建最优二叉树,生成图像自适应的 Huffman 编码,标记预测误差绝对值中的连续 0 值,预留可嵌空间. 实验结果表明,与现有算法相比,本文算法能够获得更高的信息嵌入容量. 在今后的工作中,将主要围绕以下 3 个方面进行改进提高:首先,改进加密算法进一步提高图像的安全性;其次通过使用更精确的预测器缩小预测差值来进一步提高算法嵌入容量;最后进行算法的优化以降低计算的复杂度.

[参考文献]

[1]　CHEN H S,NI J Q,HONG W. High-fidelity reversible data hiding using directionally enclosed prediction[J]. IEEE signal

processing letters,2017,24(5):574-578.
[2] HONG W,CHEN T S,CHEN J. Reversible data hiding using delaunay triangulation and selective embedment[J]. Information sciences,2015,308(1):140-154.
[3] FRIDRICH J,GOLJAN M,DU R. Lossless data embedding—new paradigm in digital watermarking[J]. Journal on applied signal processing,2002(2):185-196.
[4] CELIK M U,SHARMA G,TEKALP A M,et al. Lossless generalized-LSB data embedding[J]. IEEE transactions on image processing,2005,14(2):253-266.
[5] NI Z C,SHI Y Q,ANSARI N,et al. Reversible data hiding[J]. IEEE transactions on circuits and systems for video technology,2006,16(3):354-362.
[6] THODI D M,RODRIGUEZ J J. Expansion embedding techniques for reversible watermarking[J]. IEEE transactions on image processing,2007,16(3):721-730.
[7] TIAN J. Reversible data embedding using a difference expansion[J]. IEEE transactions on circuits and systems for video technology,2003,13(8):890-896.
[8] LI R,LI X Y. Pixel value ordering reversible data hiding algorithm based on image block selection[J]. Journal of image and graphics,2017,22(12):1664-1676.
[9] LIAO X,LI K,ZHU X,et al. Robust detection of image operator chain with two-stream convolutional neural network[J]. IEEE Journal of selected topics in signal processing,2020,14(5) 955-968.
[10] SHI Y Q,LI X,ZHANG X,et al. Reversible data hiding:advances in the past two decades[J]. IEEE access,2016(4):3210-3237.
[11] LIAO X,YU Y,LI B,et al. A new payload partition strategy in color image steganography[J]. IEEE transaction on circuits and system for video technology,2020,30(3):685-696.
[12] LIAO X,YIN J,CHEN M,et al. Adaptive payload distribution in multiple images steganography based on image texture features[J]. IEEE transaction on dependable and secure computing,2020. doi:10.1109/TDSC.2020.3004708.
[13] ZHANG X P. Reversible data hiding in encrypted image[J]. IEEE signal processing letters,2011,18(4):255-258.
[14] QIN C,QIAN X K,HONG W,et al. An efficient coding scheme for reversible data hiding in encrypted image with redundancy transfern[J]. Information sciences,2019,487(6):176-192.
[15] MA K,ZHANG W,ZHAO X,et al. Reversible data hiding in encrypted images by reserving room before encryption[J]. IEEE transactions on information forensics and secur,2013,8(3):553-562.
[16] PUTEAUX P,PUECH W. An efficient MSB prediction-based method for high-capacity reversible data hiding in encrypted images[J]. IEEE transactions on information forensics and security,2018,13(7):1670-1681.
[17] PUTEAUX P,PUECH W. Epe-based huge-capacity reversible data hiding in encrypted images[C]//2018 IEEE International Workshop on Information Forensics and Security(WIFS),IEEE,2018:1-7.
[18] PUTEAUX P,PUECH W. A recursive reversible data hiding in encrypted images method with a very high payload[J]. IEEE transaction on multimedia,2021,23:636-650.
[19] CHEN F,YUAN Y,HE H,et al. Multi-MSB compression based reversible data hiding scheme in encrypted images[J]. IEEE transaction on circuits and system for video technology,2021,31(3):905-916.
[20] YIN Z X,SHE X M,TANG J,et al. Reversible data hiding in encrypted images based on pixel prediction and multi-MSB planes rearrangement[J]. Signal processing,2021,187(10):10814.
[21] YIN Z,XIANG Y,ZHANG X. Reversible data hiding in encrypted images based on multi-MSB prediction and Huffman coding[J]. IEEE transaction on multimedia,2020,22(4):874-884.
[22] GAO G Y,ZHANG L P,LIN Y,et al. High-performance reversible data hiding in encrypted images with adaptive Huffman code[J]. Digital signal processing,2023,133:103870.

[责任编辑:陆炳新]