

doi: 10.16104/j.issn.1673-1891.2022.03.015

基于 4 类峰值位图块的可逆信息隐藏

夏婷婷^{1,2}, 刘 奇^{1,2}

(1. 武夷学院数学与计算机学院, 福建 武夷山 354300;

2. 福建省茶产业大数据应用与智能化重点实验室, 福建 武夷山 354300.)

摘要: AMBTC 压缩技术被广泛地应用于信息隐藏领域, 这类方法主要是通过对高、低量化值的判断, 选择用于信息隐藏的位图块。因此, 如何简单有效地选用更多位图块, 这对方案的性能优劣起到关键性作用。基于此, 提出了一种基于 4 类峰值位图块的可逆信息隐藏方案。方案采取 2 种策略对秘密信息进行隐藏, 一类藏于 4 类峰值位图块内, 一类藏于高、低量化值相等的位图块中。前者, 采用 (7, 4) 汉明码的信息隐藏策略; 后者, 采用秘密信息直接替换位图的方法。同时, 将采用对高量化值的最低比特位进行标识、记录, 用来区分 4 类峰值位图的第 1 个位图和其他位图。实验结果表明, 该方案的效果优于 1 种峰值位图和 2 种峰值位图方案, 在保持较高图像质量的同时, 大大提高了信息隐藏量。

关键词: 灰度图像; 可逆信息隐藏; 4 类峰值位图; AMBTC

中图分类号: TP309.2 **文献标志码:** **文章编号:** 1673-1891(2022)03-0079-06

A Reversible Data Hiding Method Based on Four Types of Peak Bitmaps

XIA Tingting^{1,2}, LIU Qi^{1,2}

(1. School of Mathematics and Computer, Wuyi University, Wuyishan, Fujian 354300, China;

2. Fujian Key Laboratory of Big Data Application and Intellectualization for Tea Industry, Fujian 354300, China.)

Abstract: AMBTC compression technology is widely used in the field of information hiding. This method mainly selects the bitmap blocks for information hiding by judging high and low quantization values. Therefore, how to use more blocks simply and effectively plays a key role in the performance of the scheme. A reversible information hiding scheme based on four kinds of peak bitmap blocks is proposed in this paper. The scheme adopts two strategies to hide secret information: one is to hide secret information in four types of peak bitmap blocks, and the other is to hide secret information in bitmap blocks with equal high and low quantization values. The former adopts (7, 4) hamming code information hiding strategy, and the latter uses secret information to replace bitmaps directly. At the same time, the lowest bit of the high quantization value will be identified and recorded to distinguish the first bitmap of the four types of peak bitmaps from other bitmaps. Experimental results show that the proposed scheme is superior to one peak bitmap scheme and two peak bitmap schemes, and greatly improves the amount of information hiding while maintaining high image quality.

Keywords: grayscale image; reversible data hiding; four types of peak bitmaps; AMBTC

0 引言

随着在线信息交换需求的不断增长, 用于安全信息交换的数据隐藏方法也有了更长足的发展^[1]。数据隐藏方法主要是通过视觉质量和数据隐藏能力 2 个参数来评价, 通常二者之间存在此消彼长的制衡关系。

目前, 数据隐藏技术(隐写)主要基于 3 类, 分别

为空域、频域和压缩域^[2-3]。空域隐藏方法, 主要是将秘密信息通过直接更改原始像素值来完成信息隐藏。LSB 是空域信息隐藏方法里的最典型策略, 它是通过更改原始像素最低一位像素来完成。该方法适用于二值图像、灰度图像以及彩色图像。基于频域的信息隐藏方法, 主要是将覆盖图像通过预先确定的变换函数如离散余弦变换(DCT)、离散傅里叶变换(DFT)、离散小波变换(DWT)转为系数, 然

收稿日期: 2022-03-20

基金项目: 福建省中青年教育科研项目(科技类)(JAT200671), 福建省自然科学基金项目(2021J011145)。

作者简介: 夏婷婷(1985—), 女, 黑龙江鸡西人, 讲师, 硕士, 研究方向: 信息隐藏、光纤通信。

后将秘密信息嵌入系数里,最后通过修改过的系数来重构隐写图像,该方法与基于空域的方法比,后者信息隐藏能力更强,但是对秘密信息的保护力和抗攻击力却是前者更强。压缩域隐藏方法,主要是将秘密信息嵌入到各种压缩版本的媒体数据中。

基于压缩域的信息隐藏方案在需要实时数字隐蔽通信的场景中非常有用,为了优化资源利用,联合摄影专家组(JPEG)、Huffman 编码、基于字典的编码等数据压缩技术起到了关键作用,可以显著减小信息的大小。通常,压缩技术分为 2 类:无损压缩技术、有损压缩技术。无损压缩技术,如算术编码、Huffman 编码、基于字典的编码等,允许原始媒体实现原始媒体的完美重构,而有损压缩技术可以从其压缩形式获得近似重构^[4-6]。有损压缩技术,如矢量量化(VQ)、块截断编码(BTC)、绝对矩块截断编码(AMBTC)等具有较高的压缩率。在有损压缩技术中,BTC、AMBTC 是用于实时通信的最简单的图像压缩技术之一,可提供最佳图像质量和高压缩比。虽然 VQ、JPEG、BTC、AMBTC 等方法都被用于信息隐藏,但基于 AMBTC 的信息隐藏策略在低复杂度方面具有明显的性能优势^[7-8]。因此,本文提出的算法是基于 AMBTC 压缩领域的可逆信息隐藏。

1 AMBTC 压缩

AMBTC 的编码/解码流程如下:

首先,将像素大小 $M \times N$ 的图像分割成非重叠块,每个非重叠块像素大小为 w , $w = n \times n$ 。因此,序列 $\{x_i\}, i = 1, 2, 3, \dots, w$, 定义为一个块内的所有像素。块内平均像素值 \bar{x} 为:

$$\bar{x} = \frac{\sum_{i=1}^w x_i}{w} \tag{1}$$

根据 \bar{x} , 可以得到一个位平面 B :

$$B_i = \begin{cases} 0, & x_i \leq \bar{x} \\ 1, & x_i > \bar{x} \end{cases} \tag{2}$$

在此基础上,进一步可以得到高量化值 H 、低量化值 L , 公式如下:

$$H = \left\lfloor \frac{1}{t} \sum_{x_i > \bar{x}} x_i \right\rfloor \tag{3}$$

$$L = \left\lfloor \frac{1}{w - t} \sum_{x_i \leq \bar{x}} x_i \right\rfloor \tag{4}$$

式中: t 为大于 \bar{x} 的像素的数量, AMBTC 的压缩三元组为 $[H, L, B]$ 。

AMBTC 的解码过程如下所示:

$$y_i = \begin{cases} H, & B_i = 1 \\ L, & B_i = 0 \end{cases} \tag{5}$$

式中: y_i 是重建的块的像素值。对于 8 位灰度图像, AMBTC 的压缩比为 $8n^2/(2 \times 8 + n^2)$, 当 $n=4$ 时, AMBTC 压缩比为 4。

以下通过一个具体例子来展示 AMBTC 压缩过程:

假设一个像素块 $X = [58, 45, 68, 75; 45, 96, 85, 88; 66, 69, 50, 77; 35, 49, 82, 46]$, 通过式(1)计算可得, 平均像素值 $\bar{x} = 64.625$, 根据式(3)、式(4), 可得高量化值 $H = 78$, 低量化值 $L = 47$ 。然后, 原像素块被

$$(78, 47, B = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}) \text{ 取代, 因此重构的像素块}$$

数值为 $y = [47, 47, 78, 78; 47, 78, 78, 78; 78, 78, 47, 78; 47, 47, 78, 47]$ 。

2 (7,4) 汉明码在信息隐藏中的应用

2.1 (7,4) 汉明码容错原理

(7, 4) 汉明码是一类线性分组码, 其作为差错控制码在数字通信系统中得到了广泛的应用。(7, 4) 汉明码的码长为 7, 最小距离为 3, 纠错能力为 1 比特, 检错能力为 2 比特。(7, 4) 汉明码是由 4 个信息位和 3 个奇偶校验位组成。4 个信息位为 d_1, d_2, d_3, d_4 , 3 个校验位 p_1, p_2, p_3 , 因此, 构成的码字为 $[p_1 p_2 d_1 p_3 d_2 d_3 d_4]$, 其中:

$$\begin{cases} p_1 = d_1 + d_2 + d_4 \\ p_2 = d_1 + d_3 + d_4 \\ p_3 = d_2 + d_3 + d_4 \end{cases} \tag{6}$$

假设: H 为 (7, 4) 汉明码的校验矩阵, $H = \begin{bmatrix} 0001111 \\ 0110011 \\ 1010101 \end{bmatrix}$, RCW 为接收方接收的汉明码, Z 为校验子。则

$$Z = H \times RCW^T = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} p_1 + d_1 + d_2 + d_4 \\ p_2 + d_1 + d_3 + d_4 \\ p_3 + d_2 + d_3 + d_4 \end{bmatrix} \tag{7}$$

如果 RCW 码字在传输过程中没有发生错误的話, Z 等于 0; 当 Z 不等于 0 时, 校验子 Z 与误码位置 V 具有一对一的对应关系, 如表 1 所示。因此, 可根据误码位置对接收到的码字在对应位置上纠错。

表 1 校验子与错误位置间的关系

Z	V
000	0000000
001	1000000
010	0100000
011	0010000
100	0001000
101	0000100
110	0000010
111	0000001

$$Z = (H \times Q^T)^T \oplus S \quad (8)$$

$$F = Q \oplus V \quad (9)$$

$$S = (H \times F^T)^T \quad (10)$$

举例说明: 假设要嵌入的秘密信息 S 为“100”, 秘密信息所要嵌入的位图 B 的 16 比特为“1000110011000011”, 则先取前 7 比特 Q 为“1000110”, 根据式(7), 可以得到 Z 为“010”。查表 1 可以找到 V 为“010 0000”, 根据式(9)可得 F 为“1100110”, 则完成秘密信息的嵌入。提取过程, 接收方将 F 代入式(10)即可以得到 S 为“100”。一个 4×4 位图, 最多可以截取出 3 个 Q , 因此最多可以嵌入 9 比特信息。

2.2 (7,4) 汉明码在信息隐藏中的应用

利用(7,4)汉明码和 AMBTC 的位图 B , 可以实现信息的隐藏和提取。假设 S 为秘密信息比特流, Q 为位图 B 中选取的 7 比特, Z 为校验子, V 为位置序列, F 为接收的码字, 秘密信息藏入如式(8)和(9)所示, 秘密信息的提取如式(10)所示。

2.3 提出的方案

本文提出的信息隐藏方案主要包括 AMBTC 图像压缩、4 类峰值位图的选取、信息隐藏策略、标识信息、附加信息嵌入、信息提取、图像恢复等内容。整体方案设计如图 1 所示。

信息的嵌入具体方案如下:

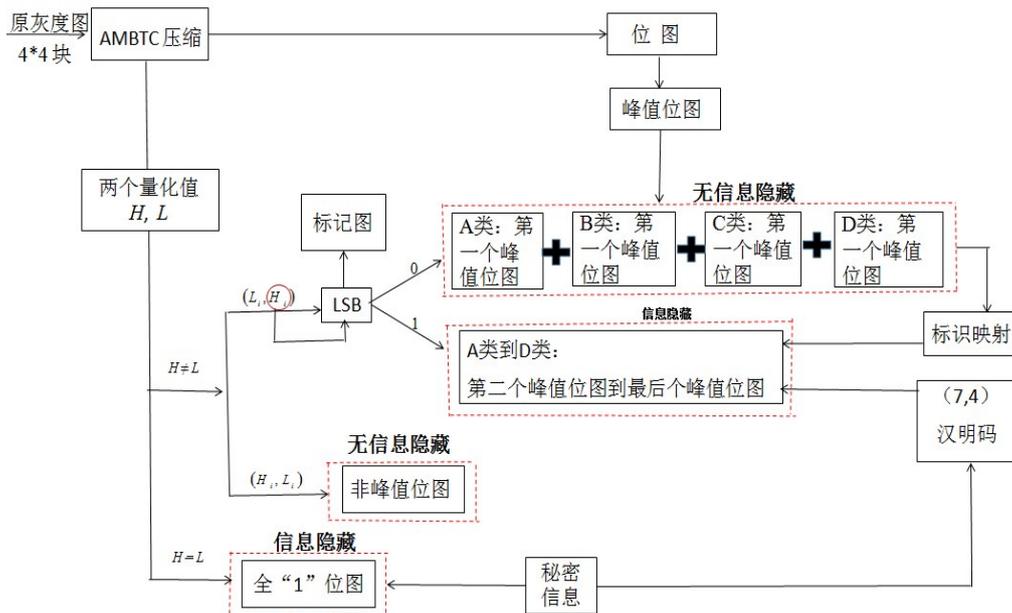


图 1 整体的方案图

1) 4 类峰值位图的选取

具体做法: 按从上到下, 从左到右, 取出一个位图块中 16 bits, 作为二进制数值, 并将二进制数值转换成十进制数值, 16 384 个位图块都进行这个操作, 最后统计频数出现最多的前 4 类十进制数, 其十进制数对应的位图被称为 4 类峰值位图。这 4 类峰值位图按出现频率由高到低的顺序标为 A 类、B 类、C 类、D 类。

2) 信息隐藏策略

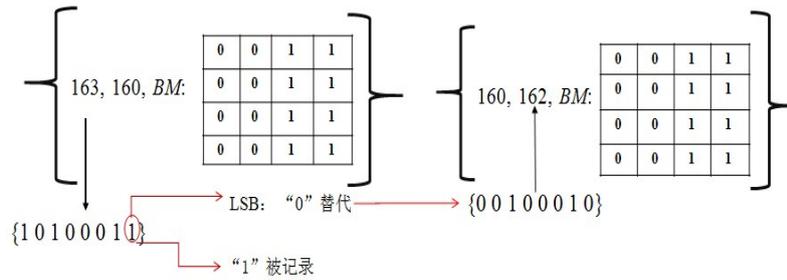
信息隐藏策略包括 2 部分, 一类将 16 bits 秘密信息直接嵌入高、低量化值相等的位图块, 另一类是将秘密信息以 3 bits 为一组通过(7,4)汉明码信息隐藏方法嵌入 4 类峰值位图中。因为(7,4)汉明码信息藏入原则为“改 1 藏 3”。因此每个峰值位图最多可以有效藏入 6 bits 信息。

3) 标识信息、附加信息嵌入

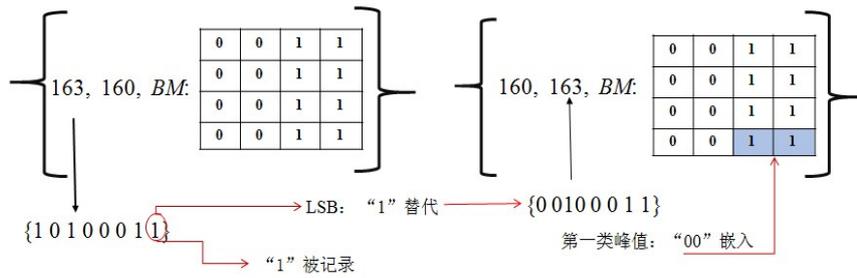
为了区分位图块是否是峰值位图块, 我们采取

高、低量化值对调策略。如果是峰值位图则将 3 元组中的高低量化值进行前后对调,如果不是峰值位图,则保持先是高量化值,后是低量化值的顺序。为了进一步分区,4类峰值位图里,每一类的第一个位图和其余位图,我们将高量化值的 LSB(最低位比特)进行表示,如果是前者表示为“0”,如果是后者

标识为“1”。此时,原始的高、低量化值的 LSB 值将被记录。同时,每一类峰值位图中非第一个位图,其余位图的后 2 位也将进行标记,4类位图的标记码分别为“00”“01”“10”“11”。在前面的过程中,所有被单独记录下来的比特信息将作为附加信息,接在秘密信息前,一同被藏入,具体如图 2 所示。



a. 每类峰值位图的第一个位图标记示意图



b. 每类峰值位图非第一个位图标记示意图

图 2 信息标记与记录

具体信息提取与图像恢复方案如下:

1)接收方收到藏有秘密信息的 AMBTC 压缩码后,对所有 3 元组进行搜索,找到所有高低量化值顺序颠倒的 3 元组,这就找到了峰值位图块,其中高量化值最低位为“1”的位图块则藏有秘密信息,(7,4) 汉明码提密规律如 2.2 所示,可提取出秘密信息。同时,高量化值最低位为“0”的位图块则表明是每一类峰值位图的第 1 个位图块,通过藏秘的位图块后 2 位标识可以找到对应的位图块,因此可以用每一类峰值位图的第 1 个位图块来还原藏有秘密的位图块。

2)找到所有高低量化值相等的位图块,直接提取其中的 16 比特秘密信息,并用全“1”比特对位图块进行还原。

3)当所有秘密信息都找到后,根据峰值位图块的个数取出相应个数的附加信息,来还原峰值位图高量化值的最低位。

4)把此时已经提取过的秘密信息,并进行过恢

复的 AMBTC 压缩码进行解码,即可得到原始无损的 AMBTC 压缩图像。

3 实验部分

3.1 实验参数

我们采用 MATLAB 编程语言实现算法,用来测试其性能。实验选取了 6 幅像素大小为 512×512 的 256 阶灰度图进行测试,如图 3 所示。

为了定量说明算法的性能,使用 2 个指标来评价:峰值信噪比 PSNR (Peak Signal to Noise Ratio)和信息隐藏量 HC(Hiding Capacity),并使用嵌入率 ER (Embedding Rate),即平均每像素所嵌入的比特数作为客观评判有效载荷的重要指标。

$$PSNR = 10 \log_{10} \left(\frac{(P - 1)^2}{e^2} \right) \quad (11)$$

式中: P 是图像灰度级总数; e 表示图像均方误差,具体定义为:



图 3 6 幅灰度图

$$e = \sqrt{\frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - f'(x, y)]^2} \quad (12)$$

式中:图像大小为 $M \times N$, M 为长, N 为宽。

3.2 实验结果与分析

由表 2 可知:6 张标准灰度图被压缩后得到的 PSNR 平均值为 32.23 dB,本方案得到的隐写图像的平均 PSNR 为 30.33 dB,较前者平均下降了 1.90 dB。

方案得到信息隐藏能力平均值为 4 768 bits,第 A 类峰值位图到第 D 类峰值位图平均个数分别为 297、241、146、110。峰值位图的个数直接影响了信息隐藏量大小。当高、低量化值相等时,相应的位图可一次性藏入 16 bits 信息,通过观察表 2,可见“Bird”具有最多有 8 个此类位图,可嵌入信息量为 $8 \times 16 = 128$ bits。

表 2 方案的实验结果

参 数	Lena	Barbara	Goldhill	Bird	Zelda	Boat	平均值
压缩图像 PSNR/(dB)	33.23	29.39	32.86	30.01	36.72	31.16	32.23
隐写图像 PSNR/(dB)	30.88	28.03	30.70	28.86	34.99	28.50	30.33
下降的 PSNR/(dB)	2.35	1.36	2.16	1.15	1.73	2.66	1.90
第 A 类峰值位图个数	447	312	304	169	307	243	297
第 B 类峰值位图个数	364	217	234	144	287	200	241
第 C 类峰值位图个数	111	103	223	112	130	196	146
第 D 类峰值位图个数	77	103	151	108	87	131	110
$H=L$,位图个数	1	0	3	8	0	0	2
信息隐藏能力/(bits)	5 986	4 386	5 496	3 302	4 842	4 596	4 768

从表 3 可知,(7,4)汉明码信息隐藏策略应用于 1 类峰值位图、2 类峰值^[9]、4 类峰值的平均信息隐藏

量分别为 1 808 bits、3 248 bits、4 768 bits,平均嵌入率分别为 0.003 4、0.006 2、0.009 1,可见本文提出的

方案较前 2 种方案具有较大的提升,分别是前二者的 2.68 倍、1.42 倍。同时,与前 2 种方案比,本方案所得图像品质平均值分别下降 1.67、1.67 dB,但依然保持较高的视觉效果。图像信息隐藏策略中,存

在图像品质跟信息隐藏量互相抑衡的现象,通常信息隐藏量增加,图像品质则会下降,信息隐藏量减少,图像品质提升。

表 3 采用不同方案的数据对比

方 案	参 数	Lena	Barbara	Goldhill	Bird	Zelda	Boat	平均值
1 类 峰值 位图	HC	2 692	1 866	1 866	1 136	1 836	1 452	1 808
	PSNR	32.88	29.19	32.70	29.87	36.41	30.93	32.00
	ER	0.005 1	0.003 6	0.003 6	0.002 2	0.003 5	0.002 8	0.003 4
2 类 峰值 位图	HC	4 870	3 162	3 264	1 994	3 552	2 646	3 248
	PSNR	32.88	29.19	32.70	29.88	36.40	30.93	32.00
	ER	0.009 2	0.006 0	0.006 2	0.003 8	0.006 8	0.005 0	0.006 2
4 类 峰值 位图	HC	5 986	4 386	5 496	3 302	4 842	4 596	4 768
	PSNR	30.88	28.03	30.70	28.86	34.99	28.50	30.33
	ER	0.011 4	0.008 4	0.010 5	0.006 3	0.009 2	0.008 8	0.009 1

5 结论

本文采用 AMBTC 对秘密图像进行预处理,对位图块进行特征统计,找到峰值位图。将秘密信息嵌入 4 种峰值位图中,同时,将峰值位图的高量化值的末尾值用“0”“1”做标记,区分每一类峰值位图的第 1 个位图块和其他位图块。4 类峰值块的被嵌入

秘密信息的位图块的末两位,用“00”“01”“10”“11”4 种标识来区别,具有减少附加信息嵌入量的目的。本文有效地利用了(7,4)汉明码的嵌入方法和“改 1 藏 3”特性,尽可能地降低对图像品质的伤害,本文提出的信息隐藏策略不仅实现了秘密图像的安全传输,而且在完整提取出秘密信息的同时,可以恢复出原始 AMBTC 图像。

参考文献:

- [1] 项煜东,吴桂兴.一种基于像素预测的图像可逆信息隐藏策略[J].计算机科学,2018,45(2):189-196.
- [2] 夏婷婷,江玲.基于 AMBTC 压缩位图和候选字符串表的可逆信息隐藏方法[J].辽宁工业大学学报,2021,41(4):250-256.
- [3] MALIK A, SIKKA G, VERMA H K. A high capacity data hiding scheme using modified AMBTC compression technique[J]. International Arab Journal of Information Technology, 2018, 16(1):148-155.
- [4] 吴友情,郭玉堂,汤进,等.基于自适应哈夫曼编码的密文可逆信息隐藏算法[J].计算机学报,2020,44(4):846-858.
- [5] XIA T T, LIN J, CHANG C C, et al. A novel adjustable reversible data hiding method for AMBTC-compressed codes using hamming distance[J]. International Journal of Embedded Systems, 2021, 14(4):313-323.
- [6] LIN J, LIN C C, CHANG C C. Reversible steganographic scheme for AMBTC compressed image based on (7,4) hamming code[J]. Symmetry, 2019, 11(10):1236.
- [7] 李亚峰,郑紫微,杨任尔,等.一种新型汉明码混合变换及在图像信息隐藏中的应用[J].移动通信,2018,42(4):77-82.
- [8] 武晓帅,徐明,乔通,等.图像空域可逆信息隐藏研究进展[J].中国图象图形学报,2022,27(1):125-149.
- [9] 夏婷婷.基于矩阵编码和两类峰值位图的可逆信息隐藏[J].闽江学院学报,2021,42(2):56-65.