

有效的调色板图像水印算法^{*}

杨 成, 杨义先, 蔡满春

(北京邮电大学信息安全中心, 北京 100876)

摘 要: 调色板图像与真彩图像相比具有色彩量化、独立调色板等特点。调色板图像作为被广泛使用的媒体类型, 直接针对它的数字水印研究却非常少。通过分析目前调色板图像水印的主要方法, 指出存在有效性低、失真严重、不能抵抗调色板分析攻击等缺点。针对以上问题, 提出了一种有效的调色板图像水印算法。算法通过聚类分析来减少失真, 使用梯度调整策略来增加嵌入容量, 并采用频率对比方式完成对水印的嵌入和提取。在保证高的保真度前提下, 算法有效性达到 100%, 并能抵抗调色板分析攻击、格式转换攻击和全选拷贝攻击等。

关键词: 调色板图像; 数字水印; 量化; 有效性

中图分类号: TP391 **文献标识码:** A **文章编号:** 0529-6579 (2004) S2-0128-04

调色板图像由于其空间占用少、网络传输带宽低, 如 GIF、调色板 BMP, 在实际中被广泛使用, 需要发展针对调色板图像的数字水印技术^[1], 以便对其进行版权保护、使用控制、盗版跟踪等。目前, 关于图像数字水印算法的研究都集中在真彩图像上, 而对调色板图像的水印算法研究很少, 方法上主要是间接利用真彩图像的水印算法或通过调色板调整来嵌入信息^[2]。这主要是由于在调色板图像中嵌入信息存在难度。第一, 调色板图像中由于色彩量化, 颜色数量最多 256 种, 像素之间的相关性高, 对其进行轻微修改都可能产生较大的失真, 而间接使用真彩图像算法往往会由于较大的量化失真而降低了算法的有效性和鲁棒性; 第二, 调色板图像由调色板和图像索引矩阵构成, 对调色板的任意排列不影响图像, 使得无法直接利用调色板来嵌入信息; 第三, 对调色板中的颜色进行修改往往会产生某种特殊模式, 容易受到调色板分析的攻击, 如排序等。

针对以上问题, 本文从移动网络中大量调色板图像需要版权保护的需求出发, 考虑到终端用户的处理能力有限, 更注重水印算法有效性等因素, 提出了一种有效的调色板图像水印算法。该算法针对调色板图像的特点, 将信息隐藏在图像索引矩阵中, 而不改变调色板, 避免了调色板分析。为了最大限度的利用像素及其颜色之间的相关性, 我们采用聚类分析技术对调色板进行聚类分组, 并引入了

JND 限^[3], 使得每个聚类组中任两个元素的距离小于事先确定的 JND 限。为了最大化嵌入容量, 利用聚类的结果, 采用梯度调整策略对像素间的相关性进行了调整, 降低相关性, 增加可嵌入的信息, 并以相邻像素的频率关系嵌入信息。

实验表明, 该算法的有效性达到 100%, 保真度高, 并能抵抗调色板分析攻击、格式转换攻击和全选拷贝攻击。

1 调色板图像算法存在的问题

目前调色板图像数字水印算法策略大都可以分为两类, 一类就是将调色板图像转换为真彩图像, 间接利用真彩图像的算法进行信息的嵌入和提取; 另一类就是直接利用调色板图像的特点对调色板进行修改调整来设计算法。

图 1 中给出了间接利用真彩算法的策略所经历主要步骤。 X 、 XW 为调色板图像, Y 、 YW 为真彩图像, $H(^{\circ})$ 和 $F(^{\circ})$ 为调色板图像与真彩图像之间转换的函数, M 为嵌入的信息。

但是在一般情况下, $F(^{\circ})$ 不等于 $H(^{\circ})$ 的逆函数, 因此。提取信息时就必须要考虑 $F(^{\circ})$ 带来的影响。 $F(^{\circ})$ 的目标是在尽量减少失真的情况下, 将 YW 中用到的颜色数量减少到 256 以下, 主要通过量化来完成。

常用量化方法的有均匀量化和最小方差量化, 这两种量化都引入了量化噪声。均匀量化采用一致

^{*} 收稿日期: 2004-09-11

基金项目: 国家 973 基础研究资助项目 (TG1999035804); 国家自然科学基金资助项目 (4042022); 教育部优秀青年教师资助项目

作者简介: 杨 成 (1974 年生) 男, 博士研究生; E-mail: cafeeyang@163.com

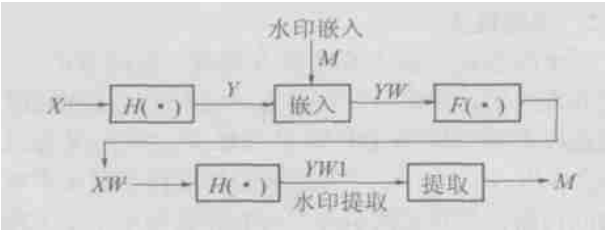


图 1 间接采用真彩图像算法的
调色板图像水印一般过程

Fig 1 General procession of palette image indirectly
using algorithms for true color images

的量化步长 q ，最小方差量化对 YW 中每一种颜色 x_i 根据其使用频率采用不同的量化步长 q_i 。根据 Widraw 的证明^[3]可以得出，若量化频率至少是 YW 颜色空间最高采样频率的两倍，则量化噪声对 YW 无影响。但实际上，由于 YW 颜色空间的分布是很难实现估计的，因此无法确定最大量化步长及其分布，也就无法去除量化噪声对 YW 的影响。量化噪声的存在降低了算法鲁棒性，无法达到 100% 的有效性。此外，间接利用真彩图像算法的调色板图像水印由于增加了 $H(·)$ 和 $F(·)$ ，使得算法效率也下降很多，因此在要求有效性高和效率高的场景并不适用。

利用调色板图像特点调整调色板来嵌入信息是目前研究相对较多的方法，主要有 Liu Hongmei 的调色板冗余法^[2]、Niimi. M. 的 BPCS^[4]和颜色量化法^[5]等。这几种方法都在调色板中产生了新的颜色或改变了调色板，使得无法抵抗包括调色板排序、颜色调整在内的调色板分析攻击、格式转换攻击，任何对调色板的轻微修改都会破坏嵌入的信息而无法正确提取。

此外，调色板冗余法的基本假设是调色板颜色的冗余和颜色使用频率的不平衡，对于正好具有 256 种颜色的调色板图像或颜色使用频率大致相同的情况，该算法无能为力。如果将嵌入隐藏信息的调色板图像 (GIF 和索引 BMP) 进行格式转换到真彩图像，再转换为调色板图像，信息也将会被破坏。

2 有效的调色板图像水印算法

2.1 嵌入算法

2.1.1 调色板聚类 传统的调色板处理如排序、增加冗余等在调色板中产生了某种特殊模式，容易受到调色板分析攻击，而且处理后的调色板颜色往往呈线性排列，由此嵌入信息往往会对图像产生较大失真。通过调色板聚类可以得到较好的颜色分

组，组内各颜色在所选择的 JND 下相同，而且不会产生线性排列，比单纯的排序要有更好的保真效果和更大的隐藏容量。由于 JND 衡量了保真性，因此当调色板聚类产生错误的时候，图像也已经失真严重，没有保护的价值了。

设调色板图像 Y 中使用到的所有颜色为集合 $\{x_1, x_2, \dots, x_n\}$ ， $x_i = (r_i, g_i, b_i)$ ， n 为颜色总数。定义任意两种颜色 x_i 与 x_j 之间的相似性测度 D_{ij} 和聚类准则函数 J ：

$$D_{ij} = (|r_i - r_j| + |g_i - g_j| + |b_i - b_j|) / 756$$

$$J = \sum_{l=1}^k \sum_{x \in C_l} \|x - m_l\|^2 \tag{1}$$

若某一个聚类组内存在某两个元素之间 $D_{ij} > JND$ ，则对该聚类组进行进一步的 K-均值聚类，直到每一个聚类组内各元素之间 $D_{ij} < JND$ 。具体流程如下。

- (1) 令聚类组 $C_0 = \{x_1, x_2, \dots, x_n\}$ ， $C = C_0$ ；
- (2) 设 $K = 2$ ，迭代次数 $i = 1$ ；
- (3) 统计空间 C 中所有调色板项目对应的像素个数，选取个数最多的两个项目作为初始聚类中心；
- (4) 使用相似性测度 M 和聚类准则函数 J ，调用 K 均值聚类算法^[3]，将聚类组 C 划分为两个聚类组 $C_{i,1}$ 、 $C_{i,2}$ ，并使准则函数 J 达到最小值。此时有 $C = C_{i,1} \cup C_{i,2}$ 且 $C_{i,1} \cap C_{i,2} = \phi$ ；
- (5) 按给定的 JND 限，对产生的聚类组 $C_{i,1}$ 、 $C_{i,2}$ ，分别考察是否满足任两个元素 x_i 和 x_j 的 $D_{ij} < JND$ 。若 $C_{i,1}$ 不满足，则令 $C = C_{i,1}$ ， i 增加 1，转 3；若 $C_{i,2}$ 不满足，则令 $C = C_{i,2}$ ， i 增加 1，转 3；若 $C_{i,1}$ 、 $C_{i,2}$ 均满足，说明空间 C 聚类成功；
- (6) 重复 3~5 直到所有空间均聚类成功，最终产生的所有聚类组为 $C_1, C_2, \dots, C_k$ 。

2.1.2 梯度调整 设聚类后得到 K 个聚类组 $P_1, P_2, \dots, P_K$ ，则调色板中每一种颜色 c 属于且只属于一组，定义 $G(c) = k, k < K$ ，令 $D_{\max}(k)$ 为第 P_k 个聚类组中相似性测度的最大值，即：

$$D_{\max}(k) = \max\{D_{i,j} | i \neq j, x_i \in P_k, x_j \in P_k\} \tag{2}$$

定义调色板图像 Y 的图像矩阵中相邻像素 (i, j) 和 $(i, j+1)$ 的梯度为：

$$D = D_{mm}$$

其中

$$x_m = RGB(i, j), x_n = RGB(i, j + 1) \tag{3}$$

则相邻像素之间存在如下关系：

$$D = 0, \text{ if } x_m = x_n$$
$$D \leq D_{\max}[G(x_m)],$$
$$\text{if } x_m \neq x_n \text{ and } G(x_m) = G(x_n)$$
$$D > D_{\max}[G(x_m)],$$
$$\text{if } x_m \neq x_n \text{ and } G(x_m) \neq G(x_n) \quad (4)$$

考虑到嵌入的有效性和保真性，一般只有 (4) 式中第二种关系可以用来嵌入信息，但是可以利用前面的聚类分组，通过对像素值增加轻微的扰动来改变相邻像素的梯度，使其满足 (4) 式中第二种关系。方法是在 $G(x_m)$ 中查找与 x_m 相似性测度最小的 x_l ，且 $l \neq m$ ，令 $\text{RGB}(i, j+1) = x_l$ ，则有 $0 < D' \leq D_{\max}[G(x_m)]$ 。

但是第 3 种关系经过梯度调整后产生的变化会有较大失真，即 $D_{nl} > \text{JND}$ ，需作进一步的调整。文献 [6] 指出在的点阵中如果中间点位于色彩边缘和视觉不敏感区域，则其可以被修改。为了满足算法需要，结合聚类结果，我们对他进行了简化和调整，对于相邻像素，只有以下一种情况可以用来调整梯度，即：

$$\begin{array}{ccccccc} & O' & & O & & X & & X' \\ \hline \end{array}$$

其中，
 $O = \text{RGB}(i, j), X = \text{RGB}(i, j+1),$
 $G(O) = G(O'), G(X) = G(X'), G(O) \neq G(X)$
在这种情况下，修改 O 或 X 不会引起视觉上的失真。对所有满足上述关系的相邻像素，按照前面的方法调整梯度，以满足式 (4) 中第 2 种关系关系。

2.1.3 频率对比 根据前面的调色板聚类 and 梯度调整，使得图像中满足式 (4) 中第 2 种关系的相邻像素数量达到最大，为水印嵌入提供充足空间。为了使嵌入的水印信息与调色板数据无关，计算每种颜色的统计频率， $f(x_i) = n(x_i)/n$ ，并通过改变相邻像素颜色的频率对比关系嵌入信息，其中 n 为像素总数， $n(x_i)$ 为颜色为的 x_i 像素数。为简单起见，对图像进行行扫描，对图像矩阵中每 2 个相邻像素 (i, j) 和 $(i, j+1)$ 进行如下处理：

当 $b = 1, x_m \neq x_n, G(x_m) = G(x_n)$ 时，若 $f(x_m) < f(x_n)$ ，则 (i, j) 与 $(i, j+1)$ 交换；若 $f(x_m) > f(x_n)$ ，则不变。
当 $b = 0$ 时， $x_m \neq x_n, G(x_m) = G(x_n)$ ，若 $f(x_m) < f(x_n)$ ，则不变；若 $f(x_m) > f(x_n)$ ，则 (i, j) 与 $(i, j+1)$ 交换。其中 $x_m = \text{RGB}(i, j), x_n = \text{RGB}(i, j+1)$ ， b 为待嵌入的信息位。

2.2 提取算法
水印信息经嵌入算法嵌入到调色板图像中后，提取方法相对简单。首先采用与嵌入算法中相同的调色板聚类算法和 JND 限进行聚类，得到聚类组集合 $\{P_1, P_2, \dots, P_K\}$ 。然后对图像进行与嵌入式相同的扫描，并对图像矩阵中所有相邻像素 (i, j) 和 $(i, j+1)$ 进行如下处理：

令 $x_m = \text{RGB}(i, j), x_n = \text{RGB}(i, j+1)$ 当 $x_m \neq x_n, G(x_m) = G(x_n)$ 时，若 $f(x_m) < f(x_n)$ ，则 $b = 0$ ；若 $f(x_m) > f(x_n)$ ，则 $b = 1$ 。 b 即为提取出的信息位，所有的 b 即可构成完整的信息。

2.3 实验结果
实验中采用了 101×80 的调色板 BMP 图像进行测试，嵌入的信息为 512 bit 随机位， $\text{JND} = 0.1$ 。图 2 中给出了嵌入信息前后的直观比较。可以看出，嵌入信息前后图像在视觉效果良好。

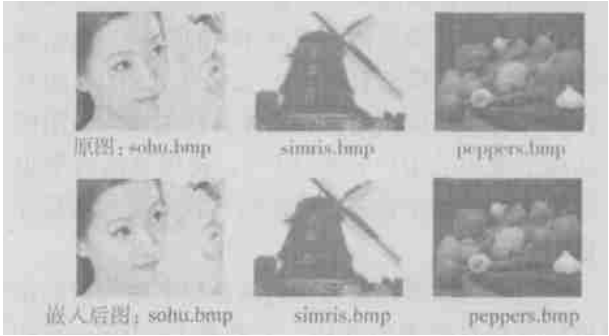


图 2 嵌入信息前后图像对比
Fig. 2 Original and watermarked images

表 1 中给出了嵌入信息后上述 3 个图像的信噪比和有效性，其中有效性均为 100%，信噪比都在 38 dB 以上。

表 1 峰值信噪比以及有效性

Ta b 1 PSNR and validity

图像	信息长度 bits	峰值信噪比			有效性/%
		R	G	B (PSNR)	
sohu	512	42.96	40.51	38.12	100
simris	512	51.72	46.83	47.34	100
peppers	512	51.17	42.97	47.27	100

从前面的算法描述可以看出，该算法可以抵抗对调色板图像的全选拷贝攻击、调色板排序攻击和对调色板的轻微修改、格式转换攻击。这里以 sohu.bmp 为例，表 2 给出了攻击后提取信息的误码率 BER，即错误提取的信息位数与总位数的比值。其中排序前后的调色板如表 3 所示。

表 2 不同攻击下的误码率

Tab 2 BER under attacks

攻击操作	BER
“全选”、“拷贝”、“粘贴”	0
调色板排序	0
调色板修改：将第 116 项改为 (50, 33, 30)	0
格式转换	0

表 3 sohu.bmp 排序前后的调色板

Tah 3 The palette and sorted palette of sohu.bmp

索引	嵌入图调色板 (R, G, B)	排序后嵌入图调色板 (R, G, B)
.....		
111	(151, 179, 84)	(240, 240, 240)
112	(243, 207, 197)	(240, 240, 237)
113	(245, 234, 226)	(240, 198, 170)
114	(27, 23, 21)	(240, 167, 153)
115	(246, 231, 227)	(239, 246, 237)
116	(47, 35, 30)	(239, 238, 238)
.....		

3 结 论

调色板图像作为网络等实际应用中大量存在的媒体，直接针对它的数字水印研究却非常少。目前在调色板图像中添加水印的思路主要有两个，即间接利用真彩图像水印算法和对调色板进行修改。他们存在着有效性低、不能抵抗简单的调色板分析攻

击等缺点。本文提出的调色板图像水印算法通过将信息直接嵌入到调色板的图像矩阵中，避免这些问题，在满足较高的保真度前提下，使有效性达到 100%，并可以抵抗调色板分析攻击、格式转换攻击、全选拷贝攻击等。另外，该算法也为调色板图像水印的研究提供了新的思路和探索。

参考文献:

[1] COX I J, MILLER M L, BLOOM J A. Digital watermarking [M] . USA: Elsevier Science, 2002; 51— 62.

[2] NIIMI M, NODE H, KAWAGUCHI E. High capacity and secure digital steganography to palette-based images[J] . Proceedings of International Conference on Image Processing, 2002, 2: 917—920.

[3] THEODORIDIS S, KOUTROUMBAS K. Pattem recognition [M] . 2nd edition. USA: Elsevier Science, 2003; 212— 230.

[4] NIIMI M, NODE H, KAWAGUCHI E. A study on the steganography using bit plane complexity based region segmentation method[J] . Proceedings of IVCNZ 98, 1998; 74— 77.

[5] WIDROW B, KOLLAR I. Statistical theory of quantization [J] . IEEE Transactions on Instrumentation and Measurement, 1996, 45(2): 353—361.

[6] PAN G, WU Z H, PAN Y H. A data hiding method for few - color images[J] . Proceedings of International Conference on Acoustics, Speech, and Signal Processing, 2002, 4: 3469— 3472.

Valid Digital Watermarking of Palette based Image

YANG Cheng, YANG Yi xian, CAI Man chun

(Information Security Centre, Beijing University of Posts and Telecomms, Beijing 100876, China)

Abstract: The palette based image is widely used in practical, such as in internet and in mobile network. But there is little research on the palette based image in digital watermark field, since in contrast with the true color image, the palette—based image has quantized color cube and independent color table. By analyzing the present digital watermark methods of palette based image, disadvantages of them are pointed out in this paper. They are low validity, severe visual distortion and resistlessness to the palette analysis. A valid digital watermark method of palette based image is proposed. The cluster analysis is used to reduce the visual distortion, and the embedding capacity is increased by using the grads modulation scheme which is presented. Then the raw message is embedded into or extracted from the palette based image by adjusting the color frequencies of nearest pixels in image. On the premise of guaranteeing the image fidelity, the proposed method is perfect valid which means the bit error rate is almost zero. Furthermore, the method can resists attacks of the palette analysis, the format transformation and the copy paste operation.

Key words: palette based image; digital watermark; quantization; validation