对 OPA 密写的检测和增强安全性的调色板图像密写方案

张新鹏, 王朔中
(上海大学通信与信息工程学院, 上海 200072)

摘 要: 在以调色板图像为载体的密写方法中, 最佳奇偶分配(OPA) 是一种失真小, 因而隐蔽性较好的方法. 本文首先指出 OPA 方法仍存在安全漏洞, 因为经 OPA 密写的图像中存在一些特殊的颜色, 这些颜色只能被修改为其它颜色, 其它颜色却不能改为这些颜色, 分析者可根据这些特殊颜色察觉秘密信息的存在. 本文提出一种新的调色板图像密写方法, 既保留了 OPA 失真小的优点, 又不存在可被密写分析者利用的奇异颜色. 另外, 本文还考虑了调色板图像密写的另一个安全漏洞, 即由调色板图像生成二值图像并检测其混乱程度可暴露其中是否含有密写信息. 根据原始图像局部特性进行自适应嵌入可进一步提高系统的安全性, 使这种基于二值图像的分析方法也不能奏效.

关键词: 信息隐藏; 密写; 调色板图像; 密写分析
中图分类号: TP309.7 文献标识码: A 文章编号: 0372-2112 (2004) 10-1702-04

Detection of OPA Stego-Data and Secure Steganography in Palette Images

ZHANG Xin-peng, WANG Shuo-zhong
(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

Abstract: Although distortion caused by the optimal parity assignment (OPA) steganography is very small, it is not secure enough due to the existence of some peculiar colors. No other colors can be modified into these colors in data embedding. This paper proposes a new steganographic scheme that, while keeping the advantage of low distortion of the OPA, avoids the above-mentioned peculiar colors so that steganalysis based on the exploration of these colors is defeated. In addition, existence of stego-information in a palette image may often be revealed through observation of a generated binary image that, due to the randomness of the embedded data, exhibits no correlation with the host image. Making use of the local properties in the host image so that the chaotic degree is reduced, this security flaw in many steganographic techniques using palette images is removed.

Key words: information hiding; steganography; palette image; steganalysis

1 引言

密写(steganography) 是信息隐藏(information hiding) 的一个重要分支, 其目的是在不引起第三方或监控者怀疑的情况下将信息秘密、安全地发送出去, 也就是将"正在通信"这一事实隐藏起来[1]. 近年来已出现了针对不同载体类型的多种密写方法, 例如应用于未压缩图像的 LSB 方法[2]、PVD 法[3] 和应用于 JPEG 图像的 F5 方法[4]. 另一方面, 如果监控者对载体数据进行统计分析时察觉出额外信息的存在, 便可以阻止这种隐蔽通信. 这种对密写信息存在性的检测称为密写分析(steganalysis)[5~7]. 为了提高密写技术的安全性, 研究者又提出了许多具有反分析性能的密写新方法, 例如本文作者针对 RS 分析和 V^2 分析本文本文提出的安全 LSB 密写[8] 和用于彩色图像载体的 ARQP 方法[9] 等.

调色板图像在 Internet 上很常见(如 GIF 格式), 可以作为密写载体. 与普通彩色图像不同, 调色板图像用很少的颜色种类(如 256 色) 显示出可接受的彩色视觉效果. 在调色板图像中, 为每一种出现的颜色分配一个索引值, 每个像素便对应一个颜色索引值. 如果一幅调色板图像中仅出现 256 种颜色, 那么一个颜色索引只需 8 比特, 图像所需的存储空间便大大减小.

利用调色板图像的密写法可分为两大类: 基于调色板的方法和基于像素的方法. 第一种方法通过改变调色板中颜色的排列顺序来嵌入秘密信息[10], 其优点是信息隐藏不会改变图像的显示效果, 缺点是嵌入量并不会随载体图像尺寸的增大而增大. 调色板的杂乱无章也会引起监控者的怀疑, 而且许多图像处理软件可以根据亮度、出现频率对调色板进行重排, 这样就会删除已嵌入的秘密信息. 另一种方法并不改变调色板顺序, 而是将秘密信息隐藏在像素中. EZ Stego 方法将颜色按照亮度排序, 用奇数位置的颜色代表秘密信息 0; 偶数位置的颜色代表秘密信息 1, 如果像素原始颜色代表的信息与欲嵌入的秘密信息不同, 则将该像素的索引值改变为相邻位置

收稿日期: 2003-06-16; 修回日期: 2004-04-21
基金项目: 国家自然科学基金(No. 60372090)

的索引值, 使其能够对应相应的秘密信息[11]. 但是, 亮度接近的不同颜色之间的差异可能很大, 所以 EZ Stego 会引起较大的失真. Fridrich 提出了一种改进的方法, 令每种颜色代表的秘密信息为该颜色三个分量之和除以 2 的余数, 嵌入时将像素索引值调整为代表秘密信息的并且与原始颜色最接近的颜色索引值[12]. 随后 Fridrich 又提出了性能更好的方法, 当像素的原始颜色与欲嵌入的秘密信息不同时, 将其改为最接近的颜色即可. 也就是说, 一种颜色和与其最接近的颜色必然代表不同的秘密信息, 这种方法称作最佳奇偶分配(optimal parity assignment) 密写[13]. 与上一种方法相比, OPA 密写引起的失真更小. 由于调色板图像中出现的颜色与索引值一一对应, 在下面的讨论中, 我们将不区分这两个名词, 而且下文提到的/ 颜色0 都是指调色板图像中出现的颜色. OPA 密写方法简述如下:

(1) 计算每两个不同颜色之间的距离, $d_{ij}=|c_i-c_j|$, 这里 c_i、c_j 表示不同的颜色, |#| 表示取欧几里得范数, 即对三个分量之差的平方和开根号. 初始的集合 C 为空集.

(2) 将所有的 d 按从小到大的顺序排列.

(3) 循环第四步直至集合 C 包含所有的颜色.

(4) 按顺序选择下一个 d_{kl}(初始时选择第一个), 其中 c_k、c_l 至少有一个不属于 C, 若有多个相等的 d_{kl} 同时满足要求, 则随机选择一个. 如果不存在这样的 d_{kl}, 说明集合 C 已经包含所有的颜色, 结束循环. 如果 c_k、c_l 都不属于 C, 则令 $P_k=0$、$P_l=1$; 如果 $c_k\mid C$ 且 $c_l\in C$, 令 $P_k=1-P_l$, 同样若 $c_k\in C$ 且 $c_l\mid C$, 则令 $P_l=1-P_k$. 更新 C= CG{c_k} G{c_l}.

至此, 每一个颜色 c_i 都有一个或 0 或 1 的值 P_i. 颜色可以看作三维空间中的一个点, d_{kl} 可以看作联接两点的边, 设在第四步中选出的 d_{kl} 构成集合 D, 那么点集 C 与边集 D 组成了一个图. 显然在这个图中没有孤立点, 也没有回路, 但这个图并不一定是全连通的, 即这个图是森林. 每条边两端颜色对应的 P 值必然是不同的, 而且不难证明下面一个重要性质[13]: 若给定颜色 c_k, c_l 为不同于 c_k 的其它颜色, 那么所有的 d_{kl} 中最小的一个必然属于 D, 如果 d_{kl} 中有多条边同时达到最小值, 那么至少有一个属于 D.

(5) 嵌入秘密信息时, 首先将秘密信息的每一比特用伪随机游走的方式对应于图像中的一个像素. 如果像素颜色对应的 P 值和欲嵌入的比特相同, 就不作改动; 如果不同, 则将该像素的颜色改变为与原始颜色最接近的颜色, 因为连接这两种颜色的边属于 D, 所以修改后的颜色对应的 P 值必然与欲嵌入的比特相同. 由此可见, 对像素最大的改动也只是调整到最相近的颜色上, 所以这种奇偶分配颜色的方法被称为/ 最佳的0. 提取秘密信息时将像素颜色对应的 P 值取出即可.

本文首先指出 OPA 密写并不安全, 并针对其中的缺陷给出分析方法, 然后在此基础上提出更为安全的密写方案.

2　对 OPA 方法的分析

尽管 OPA 密写法引入的失真很小, 但以下的分析表明它仍然存在安全漏洞, 嵌入信息的存在性可通过分析某些异常颜色检测出来.

图 1 表示最佳奇偶分配后的情况, 七种颜色中 c_1、c_3、c_5、c_6 用于代表秘密信息 0(图中为空心点), c_2、c_4、c_7 用于代表秘密信息 1(图中为实心点). 如果距颜色 A 最近的颜色为 B, 便划一个箭头由 A 指向 B, 例如距 c_1、c_3 最近的颜色为 c_2, 而距 c_2、c_4 最近的颜色为 c_3. 箭头表示如果颜色 A 与秘密信息不匹配, 将会改为颜色 B. 在图 1 中, c_1、c_5、c_7 指向了其它颜色, 却没有其它颜色指向它们, 定义这样的颜色为/ 奇异颜色0.

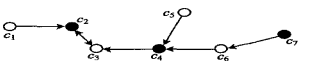


图 1　最佳奇偶分配示意

设一幅图像中有 M 个奇异颜色 $s_1, s_2, \ldots, s_M$, 与它们最接近的颜色分别为 $t_1, t_2, \ldots, t_M$, 原始图像中这些颜色出现的次数为 $h_{s,m}$、$h_{t,m}$. OPA 密写时, 奇异颜色有可能改变为相近的颜色, 却不会有其它颜色改变为奇异颜色. 设密写后上述颜色出现的次数为 $hc_{s,m}$、$hc_{t,m}$, 并设密写率为 A(秘密比特数与像素数的比值), 则有:

$$E(hc_{s,m})=\left(1-\frac{A}{2}\right)\#h_{s,m},\quad m=1,2,\ldots,M \tag{1}$$

$$E(hc_{s,m})=\left(1-\frac{A}{2}\right)\#h_{t,m}+\frac{A}{2}\#h_{s,m}+B,\quad m=1,2,\ldots,M \tag{2}$$

式(2)中的 B 表示 s_m 以外其它颜色变为 t_m 的像素数的期望, 若除 s_m 以外没有别的颜色指向 t_m, 那么 B 为 0. 如果一幅图像没有经过密写, $h_{s,m}$ 与 $h_{t,m}$ 的大小关系会比较随机; 如果密写率为 100%(每个像素上都含有秘密信息), 就会有接近一半的像素由 s_m 改为 t_m, 从式(1,2)中可以看出, 每一个 $h_{t,m}$ 的期望都大于等于相应的 $h_{s,m}$ 的期望. 因此, 分析者可以根据 $h_{t,m}$ 大于 $h_{s,m}$ 的比例来判断一幅图像是否被密写过.

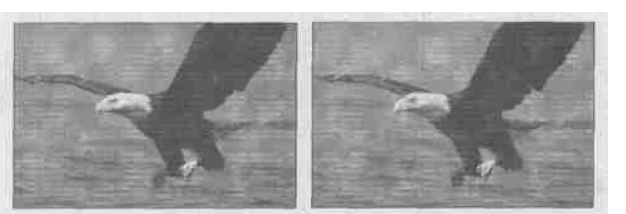


图 2　原始图像　　图 3　全部像素都含有 OPA 密写信息的结果

表 1　原始图像和密写图像中 $h_{t,m}$ 与 $h_{s,m}$ 的比较

密写率(%)	$h_{t,m}>h_{s,m}$	$h_{t,m}=h_{s,m}$	$h_{t,m}<h_{s,m}$	总数
0(原始图像)	32	4	29	65
40	49	2	14	65
70	57	3	5	65
100	60	4	1	65

图 2 为大小为 256@384、包含 256 种颜色的彩色图像, 100% 密写率 OPA 密写后的结果如图 3, 峰值信噪比为 36l 6 dB, 几乎没有视觉影响. 在 256 种颜色中, 共有 65 种奇异颜色. 表 1 比较了原始图像和密写图像中 65 对 $h_{t,m}$ 与 $h_{s,m}$ 的大

小, 原始图像中 $h_{t,m}$与 $h_{s,m}$的大小关系没有明显规律, 而密写率越高, $h_{t,m}$大于$h_{s,m}$的比例就越大. 当这个比例明显大于 1/2 时, 分析者便可以认为载体图像中含有秘密信息.

从互联网上获取 20 幅调色板图像, 大小从 512@512 到 100@100 不等, 颜色数也从 256 到 50 不等. 根据统计, 奇异颜色数与总颜色数之比的均值为 0l 27. 一般说来, 可以从载体图像的调色板中得到足够多用于分析的奇异颜色. 图 4 给出了在这些图像中密写率不同时 $h_{t,m}$大于等于$h_{s,m}$的比例(仅包含$h_{t,m}$等于$h_{s,m}$比例的一半). 可以看出, 经过密写后 $h_{t,m}$大于等于 $h_{s,m}$的比例明显升高. 因此, OPA 密写并不安全.

类似地, 也可以对文献[12]中的密写方法进行分析. 当像素颜色三个分量之和除以 2 的余数与秘密信息不符时, 文献[12] 中的密写方法便将像素颜色改为符合该条件的最接近的颜色. 同样, 可能存在一些奇异颜色, 在密写中奇异颜色像素只能改为其它颜色, 而不会有像素改为奇异颜色. 密写分析时, 求出待检测图像中奇异颜色像素数、分量之和与奇异颜色不同余并与奇异颜色最接近的颜色像素数, 如果后者大于前者的比例明显地大于 1/2, 便可以认为该图像经过密写.

3 抗奇异颜色分析的安全密写方案

为了对抗上述基于奇异颜色的密写分析, 我们提出如下的安全密写方案. 首先, 与 OPA 方法的前四步相同, 用若干条较短的边将所有的颜色 C 连接起来, 组成了一个既没有孤立点、也没有回路的图 G. 将颜色点集 C 分成两个子集, 并使图中相邻的点(即同一条边两端的点) 分属不同子集.

在图 G 中, 可能有些点只有一条边. 设这样的点为 c_S, 在与 c_S 属于不同子集的点中寻找与 c_S 距离最近的点, 将它与 c_S 相连. 这样, 所有的点都有至少两条边, 即至少有两个相邻点, 而且每条边两端的点依然分属不同子集.

用这两个子点集分别代表秘密信息 0 和 1, 并将秘密信息的每一比特用伪随机游走的方式对应于图像中的一个像素, 游走方式可以作为密钥. 若像素颜色所属子集与欲嵌入的比特相符, 就不作改动; 若不符, 则将原始颜色随机地改为与之相邻的一种颜色. 尽管这种修改会略大于 OPA 方法, 但在这种方法中不会存在奇异颜色, 因为每种颜色都可以改为两种以上的其它颜色, 也会有两种以上的其它颜色改为这种颜色. 因此, 基于奇异颜色的密写分析就无法进行.

同样以图 2 为原始图像, 由颜色集合与边集合构成的图如图 5 所示, 空心点和实心点表示不同子集, 分别代表秘密信息 0 和 1. 由图 5 可见, 不存在用于密写分析的奇异颜色. 利用本节方法以密写率 100% 得到的密写图像如图 6, 因密写引起的峰值信噪比为 351 2 dB. 与 OPA 相比虽然略微增加了失真, 但视觉效果依然很好, 与原图几乎没有差异, 尤为重要的是无法利用奇异颜色进行密写分析.

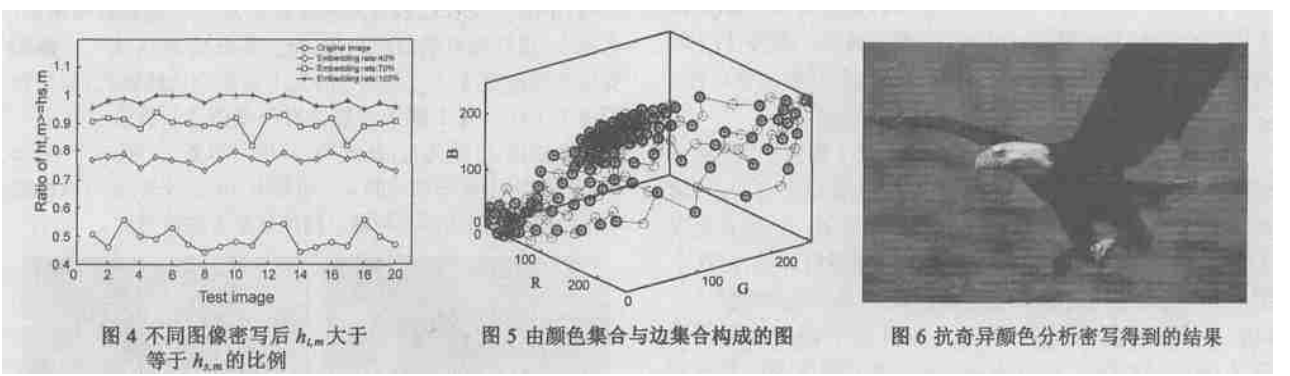


图 4 不同图像密写后 $h_{t,m}$大于等于 $h_{s,m}$的比例

图 5 由颜色集合与边集合构成的图

图 6 抗奇异颜色分析密写得到的结果

4 结合图像局部特性进一步提高密写安全性

尽管上节所述方案可以抵抗基于奇异颜色的密写分析, 但分析者仍然可以利用类似文献[14] 中的方法察觉秘密信息的存在. 假设分析者已经知道嵌入方法如本文第 3 节所述, 便可以得到两个子集的分配情况(即图 5). 将待检测图像中颜色属于第一个子集的像素用 0 代替, 颜色属于第二个子集的像素用 1 代替, 生成一个二值图像. 因为原始调色板图像中往往存在较多颜色单一的区域, 如果待分析的图像没有被密写, 从这个二值图像中便能够大致看出一些轮廓. 而如果这个调色板图像被密写过, 因为秘密信息往往被加密, 并用伪随机游走的方式对应于载体图像的像素, 得到的二值图像将是十分混乱的. 因此分析者可以根据二值图像的混乱程度来判别一幅图像是否含有秘密信息.

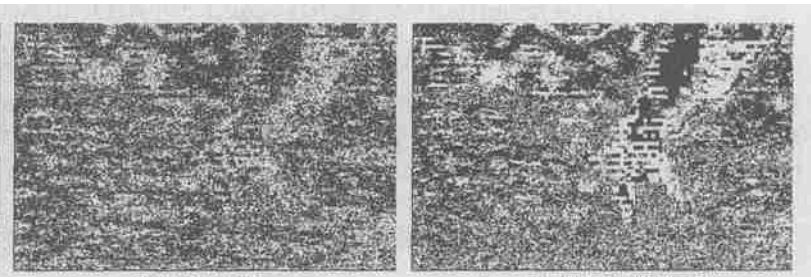
为了进一步增强密写系统的安全性, 我们对上节中提出的密写方案作如下改进: 将原始图像分为许多 2@2 的小块, 伪随机游走各小块进行密写. 嵌入时, 首先求出当前小块中四个像素的颜色平均值, 再计算四个颜色到平均值的距离和, 如果此值小于某个给定的阈值, 便不在这个小块进行嵌入; 反之, 可以进行嵌入. 在小块中按上节方法嵌入 4 比特信息后, 重新计算四个颜色到平均值的距离和, 如果该值小于给定的阈值, 那么在这个小块中的嵌入是无效的, 要在下一个小块中重新嵌入.

在接收方, 以同样的方法游走各小块、并对每个小块计算四个像素颜色到颜色平均值的距离和, 仅仅在该值大于阈值的小块中提取秘密信息即可.

因为人的视觉在图像的平滑部分较为敏感, 而这种密写方法保证了在平滑区域不作改动, 所以具有很好的隐蔽性. 而且, 分析者得到的二值图像中的大致轮廓往往来自平滑部分的边缘, 所以这种密写方法使得二值图像中的大致轮廓依然存在. 在这种方法中, 阈值取得越高, 嵌入量越小, 二值图像中的轮廓就越明显.

仍旧用图 2 作原始图像, 代表秘密信息 0 和 1 的两个子集如图 5, 图 7 即结合局部特性密写后的图像. 在这里, 密写率为 5116%, 密写引起的峰值信噪比为 3713dB. 同样密写率条件下 OPA 密写引起的峰值信噪比为 3910dB, 但对含密图像进行分析得到的二值图像比较混乱(图 8(a)), 很容易检测出秘密信息的存在. 而在由图 7 得到的二值图像中仍能看出原图像的大致轮廓(图 8(b)), 因此很难判断图 7 中是否含有秘密信息, 安全性更好.

图 7 结合局部特性密写得到的结果



(a) OPA 密写　　(b) 结合局部特性密写

图 8 对两种密写结果(密写率 51.6%)进行分析得到的二值图像

5 结论

本文首先指出失真较小的 OPA 密写并不是安全的, 该方法中存在一些颜色只能改变为相近的其它颜色, 其它颜色却不可以改为这些颜色, 通过分析这些颜色的直方图可以检测是否存在秘密信息. 为了对付这种分析, 本文随后提出了一种新的调色板密写方法, 其中没有可以用于密写分析的特殊颜色, 一方面弥补 OPA 方法的上述漏洞, 同时保留了失真较小的优点.

针对由调色板图像生成二值图像并检测其混乱程度的分析方法, 本文考虑了原始图像局部特性, 进一步改进了嵌入方法, 使之具有更为优良的隐蔽性和安全性.

密写与密写分析一向是在对抗中不断发展的, 对本文提出的密写方法进行有效分析将是下一步的研究内容.

参考文献:

[1] F A P Petitcolas, R J Anderson, M G Kuhn. Information Hiding) A Survey[J]. Proceedings of IEEE, 1999, 87(7): 1062- 1078.

[2] W Bender, D Gruhl, N Morimoto, A Lu Techniques for Data Hiding [J]. IBMSystem Journal, 1996, 35(3, 4): 313- 336.

[3] D C Wu, W H Tsai. A Steganographic Method for Images by Pixel2Value Differencing[J]. Pattern Recognition Letters, 2003, 24(9, 10): 1613 - 1626.

[4] A Westfeld. F5) A Steganographic Algorithm[A]. In 4th International Workshop on Information Hiding, Lecture Notes in Computer Science 2137[C]. Heidelberg, Berlin, Springer2Verlag, 2001, 289- 302.

[5] J Fridrich, M Goljan. Practical Steganalysis of Digital Images) State of the Art[A]. In: Security and Watermarking of Multimedia Contents IV, Proceedings of SPIE 4675[C]. San Jose, USA: SPIE, 2002, 1- 13.

[6] H Wang, S Wang. Cyber Warfare) Steganography vs. Steganalysis[J]. Communication of the ACM(to appear)

[7] X Zhang, S Wang. Vulnerability of Pixel2Value Differencing Steganography to Histogram Analysis and Modification for Enhanced Security [J]. Pattern Recognition Letters, 2004, 25(3): 331- 339.

[8] X Zhang, S Wang, K Zhang. Steganography with Least Histogram Abnormality[A]. In Computer Network Security, Lecture Notes in Computer Science 2776[C]. Heidelberg, Berlin: Springer2Verlag, 2003, 395- 406.

[9] S Wang, X Zhang, K Zhang. Steganographic Technique Capable of Withstanding RQP Analysis[J]. Journal of Shanghai University, 2002, 6(4): 273- 277.

[10] http://www.darksidel.com.au/gifshuffle/

[11] R Machado. EZ Stego. http://www.stegol.com/

[12] J Fridrich, M Goljan. A New Steganographic Method for Palette Images [A]. In Proceedings of IS&T PICS[C]. Savannah, Georgia, USA: IS&T, 1999, 285- 289.

[13] J Fridrich, D Rui. Secure Steganographic Methods for Palette Images [A]. In The 3rd Information Hiding Workshop, Lecture Notes in Computer Science 1768[C]. New York: Springer2Verlag, 2000, 47- 60.

[14] A Westfeld, A Pfitzmann. Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Stegnos, and S2Tools) and Some Lessons Learned[A]. In The 3rd Information Hiding Workshop, Lecture Notes in Computer Science 1768[C]. New York, Springer2Verlag, 2000, 61- 76.

作者简介:

张新鹏　男, 1975 年 9 月出生于黑龙江密山, 1995 年毕业于吉林大学数学系, 2004 年获上海大学工学博士学位. 现为上海大学通信与信息工程学院教师, 主要研究领域: 数字水印、密写与密写分析、数字图像处理、ATM 交换等. 已发表论文四十余篇.

王朔中　男, 1943 年 9 月出生于重庆, 1966 年毕业于北京大学, 1982 年获英国伯明翰大学博士学位, 现任上海大学通信与信息工程学院教授、博士生导师. 研究领域: 图像处理, 音频信号处理, 信息隐藏.