

GIF 图像掩密算法的统计安全性比较^{*}

丁玮¹, 孔祥维¹, 尤新刚^{1,2}, 胡 岚²

(1. 大连理工大学信息安全研究中心, 辽宁 大连 116024;

2. 北京电子技术应用研究所, 北京 100091)

摘 要: 利用一维统计特征的检测方法, 高阶统计特征的检测方法对几种常用的信息隐藏方法的安全性进行了比较研究。实验结果表明, 只改动图像索引值的信息隐藏方法的安全性优于同时改动调色板及索引值的信息隐藏方法。Fridrich 信息隐藏算法由于采用了最优奇偶值分配的方法, 其安全性优于 EZ Stego。

关键词: 信息隐藏; GIF 图像; 调色板; 索引值; 安全性

中图分类号: TP391; TP309 **文献标识码:** A **文章编号:** 0529-6579 (2004) S2-0152-04

信息隐藏^[1]技术的目的是把秘密信息隐藏在不受怀疑的载体中以实现秘密通信。许多掩密软件用 GIF 做为掩密载体。GIF 图像一般由两部分组成: 调色板和图像数据; 调色板定义了 N 种颜色索引对 (i, c_i) , 它为每一个颜色向量 c_i 指配一个索引 i ; 图像数据保存了每一个像素的调色板索引值, 而不是保存实际的颜色值。GIF 的调色板中最多包含 256 种颜色, 并且 GIF 使用了 LZW 算法对像素索引值进行压缩, 因而图像文件比较小, 得到了广泛的应用。

1 GIF 图像信息隐藏算法安全性比较

关于 GIF 图像的信息隐藏算法出现较早, 但种类不多, 其主要隐藏的方法可以归类为 3 种情况: 改动调色板、同时改动调色板和索引值、改动索引值。改动调色板典型的掩密算法是免费软件 Gif-shuffle^[2]; S Tools^[3] 和 Hide & Seek^[4] 软件应用的是同时改动调色板和索引值的原理; 只改动索引值的原理应用比较广泛, 典型的方法如 EZ Stego^[5] 以及 Fridrich 的信息隐藏算法^[6-8]。

在本节中我们利用一维统计特征检测方法和高阶统计特征检测方法分别对 Hide & Seek, S Tools, EZ Stego 及 Fridrich 的掩密算法进行了检测实验, 分析了这些掩密算法的安全性^[9]。

1.1 Hide & Seek 和 S Tools

这类信息隐藏方法首先利用抖动的方法将图像调色板的颜色数减少到一定的基色, 然后将基色通过轻微修改后扩展到整个调色板中, 最后根据要入的信息位相应的修改索引值。对于这类信息隐藏

的检测可以从两方面入手——检测调色板信息和检测索引值信息。

用 S Tools 算法隐藏信息后按亮度排序的调色板中包含了较多的相似色彩。对于信息隐藏的检测而言, 这是一个明显的视觉特征。除了视觉特征可以作为判断依据以外, 由于图像调色板色彩方差在用 S Tools 进行信息隐藏前后发生很大变化, 因此调色板色彩的方差可以作为检测的统计特征来判断图像是否隐藏了信息。

我们不仅可以通过检测调色板来判断图像是否隐藏了信息, 还可以通过像素的索引值来检测。隐藏信息后的图像的各颜色分量的直方图呈现孤立成对的特征, 隐藏信息量越大, 孤立颜色对的出现频率越大。这一特征为检测图像中隐藏了多少信息量提供了检测手段, 检测公式如下:

$$P = \frac{D_a^3}{D_h^2 \cdot (D_e + D_o)} \quad (1)$$

其中, D_a 为各颜色分量直方图中的孤立成对数, D_h 为各颜色分量直方图中的所有成对数, D_e 为各颜色分量直方图中坐标为偶数的孤立非成对数, D_o 为各颜色分量直方图中坐标为奇数的孤立非成对数。

Hide & Seek 算法将调色板分成从 0 到 252 的三元组, 其增长步长为 4, 并且图像中的“最白”的颜色为 (252, 252, 252) 而不是 (255, 255, 255)。为了减小视觉上的失真, Hide & Seek 4.0 的补丁程序将调色板的颜色从 256 色减少为 128 基色, 并通过扩展产生一个新的调色板, 即在新调色板中, 颜色 0 与颜色 1 相同, 颜色 2 与颜色 3 相同

^{*} 收稿日期: 2004-09-11

基金项目: 国家“863”高技术研究发展计划资助项目 (2002AA054010); 国家自然科学基金资助项目 (60372083)

作者简介: 丁 玮 (1980 年生), 女, 硕士研究生; E-mail: weiding_dlut@hotmail.com

…。由于 Hide & Seek 将调色板各颜色分量分成增长步长为 4 的三元组, 因此可利用公式 (2) 计算以 4 为步长的孤立的色度级的个数在各颜色分量直方图所占的比例来判断图像是否利用 Hide & Seek 隐藏了信息。

$$P=\frac{D_a^2}{D_h\circ D_i}$$

(2)

其中, D_a 为各颜色分量直方图中的以 4 为步长的孤立的色度级的个数, D_h 为各颜色分量直方图中的以 4 为步长的所有色度级的个数, D_i 为各颜色分量直方图中的以 4 为步长的非孤立的色度级的个数。

表 1 列出了使用上述检测原理检测数码相机的图片 S Tools 和 Hide & Seek 隐藏方法的实验结果。

表 1 一维统计特征检测 S Tools 和 Hide & Seek 隐藏方法的实验结果

Tah 1 Detecting results of S-Tools and Hide & Seek using one-dimension statistical character

隐藏方法	图像分辨率	嵌入信息大小/kB	数量 (幅)	检测正确率/%
S Tools	648×480	37.5	65	98.46
S Tools	720×480	42	20	100
Hide & Seek	648×480	37.5	20	100

实验结果表明, 使用一维统计特征能以较高的检测正确率检测 S Tools 和 Hide & Seek 隐藏方法, 因此这 2 种掩密方法不抵抗一维统计特征检测。这主要是由于此类掩密算法对图像进行了较大的改动, 隐藏后图像的统计特征表现明显, 攻击者可以从调色板和索引值两方面着手, 进行统计检测, 检测正确率较高。

1.2 EZ Stego

修改索引值的隐藏算法分别运用不同的方法为图像的索引值分配奇偶值, 使相互替换的索引值所代表的颜色在视觉上近似, 以减少信息嵌入所引入

的噪声, 典型的算法如 EZ Stego, 下面将介绍针对它的检测原理。

1.2.1 一维统计特征检测 EZ Stego 首先对调色板重新排序, 通过相邻的索引值之间的替换来嵌入信息。如果要隐藏的信息 (0 和 1) 在概率分布上大致相同, 那么同组 (或称索引值对) 的索引值的出现频率也应是大致相同的, 因此我们可以通过理论上索引值频率分布与随机选取的图像的索引值频率分布的比较来检测调色板图像中是否隐藏了信息。其关键是如何获得理论上的频率分布。理论频率分布必须独立于随机选取的样本图像, 因为在大多数检测情况下无法得到原始图像来与掩密图像进行比较。用 EZ Stego 进行信息嵌入只是改动索引值的最低有效位, 造成单个索引值的出现频率发生改变, 并没有改变这组索引值总的出现频率, 因此可以用同组索引值出现频率的算术均值作为理论上的频率分布。由以上分析可知, 随机选取的图像的频率分布和理论上的频率分布 (即同组索引值出现频率的算术均值分布) 之间的相似程度可以作为判断图像是否隐藏信息的一种方法。Westfeld 提出用 χ^2 检验方法^[10] 计算相似度对 EZ Stego 进行检测, 表 2 显示了检测结果, 嵌入信息概率域值设为 60%。

EZ Stego 没有改动原图像调色板, 虽然改变索引值对图像的直方图造成改变, 但由实验结果可知, 提取一维统计特征对其进行检测, 检测正确率较低。对比表 1 可以看出, 在一维统计特征检测的条件下, EZ Stego 的抗攻击性优于 Hide & Seek 与 S-Tools 这两种软件。

1.2.2 高阶统计特征检测 为了提高检测的可靠性, Fridrich^[11] 提出了针对 EZ Stego 的高阶统计特征分析方法。

我们分别对不同来源的图像用 Fridrich 提出的高阶统计特征检测方法进行了实验, 图像嵌入容量为 0.5 bpp, 横坐标代表 GIF 图像, 纵坐标代表嵌入信息容量, 实验结果如图 1 所示。

表 2 用一维统计量检测 EZ Stego 的实验结果

Tah 2 Detecting results of EZ Stego using one-dimension statistical character

图片来源	图像分辨率	嵌入信息大小/kB	数量 (幅)	检测正确率/%
柯达数码相机	720×480	42	78	8.97
型号未知数码相机	640×480	37.5	123	22.76
扫描图片	最小: 286×154 最大: 569×816	根据图像大小不同选择大小不同的文本, 对每幅图都达到满嵌;	130	50.77
网上随机搜索图片	最小: 640×427 最大: 700×511	根据图像大小不同选择大小不同的文本, 每幅图都达到满嵌;	9	22.22
Fabein 图像库下载的标准图像	最小: 572×392 最大: 768×512	根据图像大小不同选择大小不同的文本, 每幅图都达到满嵌;	12	44.44

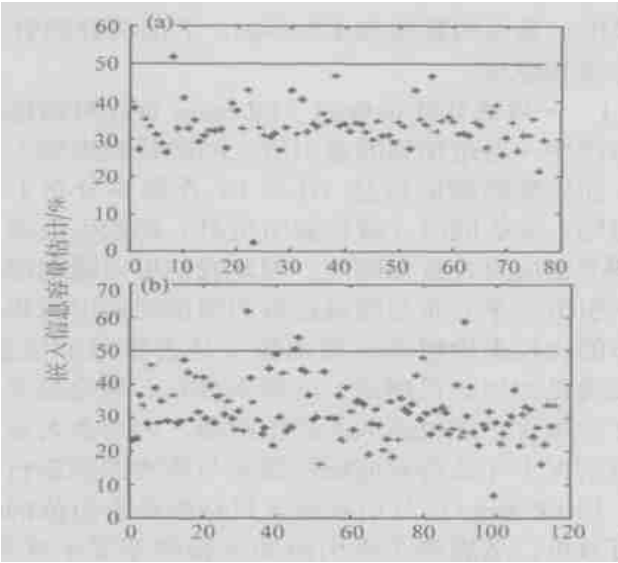


图 1 EZ Stego 高阶统计特征检测实验结果

Fig 1 Detecting results of multidimensional statistical character of EZ Stego

- (a) 77 幅数码相机拍摄的图像转变为 GIF 格式
- (b) 116 幅扫描图像转变为 GIF 格式

实验结果表明, 对于 EZ Stego 而言, 虽然 Fridrich 的高阶统计量检测方法对于嵌入信息长度的估计有一定的误差, 但这种方法对于隐藏信息存在性的估计比一阶统计特征检测方法可靠。

1.3 Fridrich 信息隐藏方法

本小节通过高阶统计特征检测实验证明 Fridrich 隐藏方法^[8]的安全性优于 EZ Stego。

1.3.1 高阶统计特征检测 图 2 显示了利用高阶统计特征检测 Fridrich 隐藏方法的实验结果, 嵌入容量同样选择 0.5 bpp, 实验结果表明这种检测方法不仅不能准确估计掩密图像的信息容量, 而且对于掩密通信存在性的检测正确率仅有 45% (估计嵌入容量为零说明未成功检测掩密通信存在)。这个实验结果与上一节检测 EZ Stego 的实验结果 (图 1) 对比证明 Fridrich 隐藏方法抗高阶统计特征检测性能优于 EZ Stego。

综合各种实验结果可知, 信息隐藏过程使图像统计特征发生改变, 利用统计特征可以实现信息隐藏分析。Hide & Seek 与 S-Tools 可以被一阶统计特征检测方法可靠检测, 其安全性低于只修改索引值的信息隐藏方法。使用 EZ Stego 进行掩密通信, 其存在性可被 Fridrich 的高阶统计特征检测方法准确的检测。用 Fridrich 的高阶统计特征检测方法不能有效地检测 Fridrich 提出的信息隐藏方法。因而 Fridrich 信息隐藏方法是目前比较安全的一种基于 GIF 的信息隐藏算法。

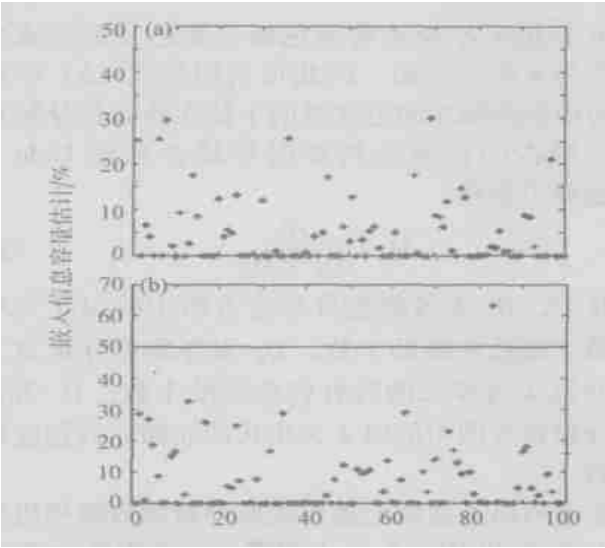


图 2 Fridrich 高阶统计特征检测实验结果

Fig 2 Detecting Results of multidimensional statistical character of Fridrich

- (a) 100 幅数码相机拍摄图像转变为 GIF 格式
- (b) 100 幅扫描图像转变为 GIF 格式

2 结 论

本文利用基于一维统计特征的检测方法, 基于高阶统计特征的检测方法进行了实验, 对比了几种常用的信息隐藏算法的安全性。实验结果表明, 只改动索引值的信息隐藏方法的安全性优于同时改动调色板及索引值的信息隐藏方法。Fridrich 信息隐藏算法由于采用了最优奇偶值分配的方法, 其安全性优于 EZ Stego。

参考文献:

[1] 尤新刚, 周琳娜, 郭云彪. 信息隐藏学科的主要分支及术语 // 信息隐藏(CIHW2001)[C]. 西安: 西安电子科技大学出版社, 2001: 43—50.

[2] KWAN M. Invariant watermarking using feature points[J]. IEEE Trans, 2003. Gifshuffle 2.0. Available from <http://www.darkside.com.au/gifshfle>.

[3] Available from www.jjct.com/steganography/toolmatrix.htm

[4] Available from www.jjct.com/steganography/toolmatrix.htm

[5] MACHADO R. EZ Stego, Stego Online, Stego. Available from <http://www.stego.com>. 1997.

[6] FRIDRICH J. Applications of data hiding in digital images. Tutorial for The ESSPA-99, Brisbane, Australia, 1999.

[7] FRIDRICH J. A new steganographic method for palette-based images. IS&T PICS, Savannah, Georgia, 1999: 285—289.

[8] FRIDRICH J, DU J. Secure steganographic methods for palette images // Proc the Third Inform. Hiding Workshop LNCS, vol. 1768. New York: Springer-Verlag, 2000. 47—embedding method. Proc. SPIE 2615, 226—259.

[9] 周琳娜, 白金茹, 尤新刚, 等. 数理统计特性与信息隐

藏系统的安全性. 信息隐藏(CIHW2002), 北京: 机械工业出版社, 2002: 28— 33.

[10] WESTFELD A, PFITZMANN A. Attacks on Steganographic Systems[J] . Lecture Notes in Computer Science, 2000, 1768: 61— 75.

[11] JESSICA F, MIROSLAV G, DAVID S. Higher order statistical steganalysis of palette images[C] . Proc EI SPIE Santa Clara, 2003: 178— 190.

Statistical Security Comparison of Steganographic Algorithm on GIF Images

DING Wei¹, KONG Xiang wei¹, YOU Xin gang^{1,2}, HU Lan²

(1. Dalian Univeraity of Technology, Dalian 116024, China;

2. Beijing Institute of Electronic Technology and Application, Beijing 100091, China)

Abstract: The principle of steganography on GIF images is introduced in this paper. Some familiar steganographic methods based on GIF images are compared by the detecting methods on one dimension statistical character and multiple dimension statistical character. The results of experiments indicate that data hiding methods on modifying indices of images have better security than methods on modifying indices and palette of images together. The steganographic method presented by Fridrich is more secure than EZ Stego because it uses the optimal parity assignment.

Key words: steganography; GIF images; palette; index; security

(上接第 151 页)

Fibonacci Transformation and Its Applications in Digital Image Watermark

SHANG Yan hong, LI Nan, ZOU Jian cheng

(CAD Research Center, North China University of Technology, Beijing 100041, China)

Abstract: A digital image scrambling method based on Fibonacci transformation is proposed. By considering its favorable property of uniformity, this paper tries to present a definition of the image scrambling effects and to compare the scrambling effects of Fibonacci transformation with those of the Arnold transformation. The experimental results show that Fibonacci transformation is better than Arnold transformation. The application of Fibonacci transformation in image watermark is also considered. The watermarking embedding algorithm is robust to some anti attack tests.

Key words: Fibonacci transformation; image scrambling; digital watermark