JPEG 图像中的安全密写方案

张新鹏 王朔中
(上海大学通信与信息工程学院 上海 200072)

摘 要：JPEG 是一种常见的图像格式，在 JPEG 图像中进行密写有着重要的实际意义。但以往的密写技术往往会改变载体图像的 DCT 系数直方图或增大分块效应，分析者可以据此察觉秘密信息的存在。为了提高密写安全性，该文提出一种新的 JPEG 密写方案。该方案在将秘密信息嵌入非 0 的非直流 DCT 系数时，既不改变原始图像的 DCT 系数直方图，也不增加载体图像的分块效应，因此可以抵抗基于直方图和基于分块特性的密写分析。实验表明，当载体图像的质量因子大于 35 时，秘密信息的安全性可以得到有效保证。

关键词：信息隐藏，密写，JPEG，密写分析

中图分类号：TP309 **文献标识码**：A **文章编号**：1009-5896(2005)11-1813-05

Secure Steganographic Algorithm in JPEG Images

Zhang Xin-peng Wang Shuo-zhong
(*School of Communication and Information Engineering, Shanghai University, Shanghai* 200072, *China*)

Abstract Although it is very meaningful to embed secret messages into a cover image in JPEG format, an analyst can detect the presence of hidden data according to a change of DCT coefficient histogram or an increase of blockiness parameter. In order to improve the security of embedded information, a secure steganographic technique is proposed, in which AC DCT coefficients with non-zero values are used to carry the secret bits when the DCT coefficient histogram and blockiness property are reserved. This way, the histogram-based and blockiness-based steganalyses are effectively resisted. Simulation results show that the proposed method is secure for cover images with quality factor more than 35.

Key words Information hiding, Steganography, JPEG, Steganalysis

1 引言

密写(steganography)是信息隐藏(information hiding)的一个重要分支，其目的是将信息秘密，安全地发送出去而不引起第三方的怀疑[1]。作为密写的对抗技术，密写分析(steganalysis)的目的则是检测多媒体载体中额外秘密信息的存在性[2]。迄今为止，已出现了针对不同载体类型，不同密写算法的多种分析方法[3]。密写与密写分析是相互促进，共同发展的两方面。当新的密写分析能够对某种密写方案进行有效检测时，密写方往往会相应地改进嵌入算法以弥补漏洞，对抗密写分析，促使密写技术进一步发展[4－6]。

JPEG 是一种使用非常广泛的图像格式，以 JPEG 图像作为密写载体有着重要的应用价值。JPEG 图像是由分块 DCT 变换后的系数按照一定的量化表量化而成，量化后的系数是量化表中对应量化步长的整数倍。目前已出现了多种以JPEG图像为载体的密写方法。有的方法先修改量化表中对应中高频分量的量化步长，然后将秘密信息嵌入在图像的中高频系数上[7]。但修改后的中高频量化步长会小于低频量化步长，这种异常会暴露秘密信息的存在，因此安全性不高。大多数密写方法并不改变原始图像的量化表，而是根据一定的规则直接将秘密信息嵌入在量化后的DCT系数上，如Jsteg，OutGuess[8]，F5[9]等。但这些方法会改变载体图像的一些统计特性，Fridrich提出了两种有效的密写分析方法，可以觉察DCT系数直方图[10]和分块特性[11]的变化，据此检测出秘密信息的存在。

本文提出一种以 JPEG 图像为载体的新密写方案，这种密写方案既不改变原始图像的量化表，也不改变载体图像的 DCT 系数直方图和分块特性，使安全性得到有效提高。

2004-04-28 收到, 2004-12-28 改回
国家自然科学基金(60372090)，上海市科委基础研究重点项目(04JC14037)和上海市教委青年基金(04AC93)资助课题

2 对 JPEG 图像的密写分析

JPEG 压缩过程如下：首先将图像分成 8×8 的小块，在每个小块内进行 DCT 变换，然后根据 DCT 系数的位置按照一定的步长进行量化，最后对量化的 DCT 系数编码。在本文的讨论中，原始图像的量化表(即对应于不同频率位置的量化步长)始终不变，DCT 系数也始终是对应量化步长的整数倍，所以本文将 DCT 系数看作整数，处理时也作为整数来处理。

用JPEG图像的DCT系数来负载秘密信息难免要对DCT系数有所改动，尽管改动后的值仍然是整数(即对应量化步长的整数倍)，但DCT系数的直方图有可能发生变化。另一方面，由于相邻块上的信息嵌入操作是不相关的，这种行为会加剧图像的分块效应。Fridrich等[10,11]提出的两种密写分析方法就是先估计出原始图像的DCT系数直方图和分块特性，然后与待检测图像的DCT系数直方图和分块特性进行比较，如果差异较大，则说明待检测图像是经过密写处理的。

虽然分析者无法得到真正的原始图像，但可以构造出一幅统计特性相近的参考图像。具体方法如下：将待检测的图像删除左侧 4 行或上面 4 列，然后重新分块并对DCT系数按相同的量化表进行量化(量化表可以从待检测图像的文件头中读出或根据待检测图像的DCT系数值求出[12])，便得到参考图像。参考图像与原始图像有相近的内容，并且用相同的量化表进行量化，所以可将参考图像的DCT系数直方图和分块特性作为原始直方图和原始分块特性的估计。分块特性可以用下式定量表示：

$$B=\sum_{i=1}^{\lfloor (M-1)/8\rfloor}\sum_{j=1}^{N}\left|g_{8i,j}-g_{8i+1,j}\right| +\sum_{i=1}^{M}\sum_{j=1}^{\lfloor (N-1)/8\rfloor}\left|g_{i,8j}-g_{i,8j+1}\right| \tag{1}$$

其中M，N是图像的行数和列数，$g_{u,v}$表示像素灰度值，函数$\lfloor\cdot\rfloor$表示下取整。式(1)的含义是相邻块的相邻像素灰度值之差的绝对值总和。用式(1)计算参考图像的分块特性B_E和待检测图像的分块特性B_1，如果B_1明显大于B_E或者参考图像与待检测图像的DCT系数直方图存在明显差异，则认为待检测图像是经过密写的。这样的分析方法不但能够有效察觉以Jsteg，OutGuess，F5 等方法嵌入的秘密信息，而且可以进一步估计出嵌入的信息量。

3 安全密写方案

为了提高 JPEG 图像中密写信息的安全性，我们提出一种新的密写方案，这种新方案在嵌入秘密信息的同时几乎不改变原始图像的 DCT 系数直方图和分块特性。新方案遵循如下两条原则：

(1) 由于人的视觉对低频分量比较敏感，所以直流 DCT 系数不用于负载秘密信息；另外，JPEG 图像中有很大一部分(甚至绝大部分)系数为 0，如果在密写时修改这些系数会大大减少 0 的数量，引起分析者的怀疑，所以值为 0 的 DCT 系数不用于负载秘密信息。也就是说，秘密信息仅嵌入在非 0 的非直流 DCT 系数上，并且每个系数用来负载 1bit 秘密数据。

(2) 量化后的DCT系数都是整数，用正奇数，负偶数(…, –6, –4, –2, 1, 3, 5, …)代表秘密信息1，用负奇数，正偶数(…, –5, –3, –1, 2, 4, 6, …)代表秘密信息 0。

具体密写方法如下：

首先统计原始图像的DCT系数直方图和分块特性。对位于不同频率位置的DCT系数分别统计直方图，每个小块中的非直流DCT系数共有 63 个，记这 63 个直方图为$\boldsymbol{H}_1$，$\boldsymbol{H}_2$，…，$\boldsymbol{H}_{63}$，第k个频率位置上值为m的DCT系数的个数记为$h_k(m)$。并根据式(1)计算原始图像的分块特性，记为B_0。这些统计数据在密写后应该几乎没有变化。

将秘密信息逐比特地对应于非 0 的非直流DCT系数，如果原始系数代表的信息与欲嵌入的秘密比特相同，那么就不需作任何改动；如果不同，则需要修改DCT系数。设第k个频率位置上值为m的DCT系数中需要改动的个数为$w_k(m)$，由于秘密信息往往经过加密，可以看作 0，1 随机分布的比特流，所以DCT系数需要改动的概率为 1/2，即

$$w_k(m)\approx\frac{1}{2}\cdot h_k(m),\quad m\neq 0,\ k=1,2,\cdots,63 \tag{2}$$

修改DCT系数时将其加 1 或减 1 都可以完成嵌入秘密比特的操作，例外情况是原始系数为 1 时要修改到–1 或＋2，原始系数为–1 时要修改到–2 或＋1。设$w_k(m)$个需要改动DCT系数中有$u_k(m)$个向负向调整，$v_k(m)$个向正向调整，由于调整而新产生的第k个频率位置上值为m的DCT系数的个数记为$w_k'(m)$，则有

$$w_k(m)=u_k(m)+v_k(m),\quad m\neq 0,\quad k=1,2,\cdots,63 \tag{3}$$

$$w'_k(m)=u_k(m+1)+v_k(m-1),\quad m\neq -1,0,1,\quad k=1,2,\cdots,63 \tag{4}$$

$$w_k'(-1)=u_k(1)+v_k(-2),\quad k=1,2,\cdots,63 \tag{5}$$

$$w_k'(1)=u_k(2)+v_k(-1),\quad k=1,2,\cdots,63 \tag{6}$$

我们希望密写后DCT系数直方图没有变化，也就是说对每个频率位置k，找到合适的$u_k(m)$，$v_k(m)$使$w_k(m)$与$w_k'(m)$尽可能相等。设第k个频率位置上DCT系数的最小值为$\min_k$，令

$$u_k(\min_k)=0,\quad v_k(\min_k)=w_k(\min_k) \tag{7}$$

即$w_k(\min_k)$个DCT系数全部向正向调整，然后m从小到大逐步迭代计算：

$$u_k(m+1)=\begin{cases}0, & w_k(m)-v_k(m-1)<0\\ w_k(m)-v_k(m-1), & 0\le w_k(m)-v_k(m-1)\le w_k(m+1)\\ w_k(m+1), & w_k(m)-v_k(m-1)>w_k(m+1)\end{cases} \tag{8}$$

$$v_k(m+1)=w_k(m+1)-u_k(m+1) \tag{9}$$

注意这里已经考虑到了$u_k(m)$，$v_k(m)$的取值范围必须在[0 $w_k(m)$]之间，另外，0 值DCT系数不用于负载秘密信息，所以从小到大计算$u_k(m)$，$v_k(m)$时应跳过m=0 的情况，具体的解析表达式在这里略过。例如，当{w(−4)=1, w(−3)=4, w(−2)=10, w(−1)=17, w(1)=16, w(2)=11, w(3)=6, w(4)=2}时，根据上述方法得到的u = {0, 1, 3, 7, 10, 6, 5, 1}，v = {1, 3, 7, 10, 6, 5, 1, 1}，新生成的{w'(−4)=1, w'(−3)=4, w'(−2)=10, w'(−1)=17, w'(1)=16, w'(2)=11, w'(3)=6, w'(4)=1, w'(5)=1}，可见新直方图与原直方图仅有微小变化。综上所述，我们可以计算出在保持原始DCT系数直方图的条件下，不同频率位置需改动的DCT系数中应向负向调整和应向正向调整的个数。

逐个小块逐个系数进行密写。当原始系数代表的信息与欲嵌入的秘密比特相同时，不需作任何处理；如果不同，按下面的方法进行调整：

(1) 每完成一个DCT系数的调整，都要根据式(1)测算当前图像的分块特性，记为B_{temp}，并用$s_k(m)$，$t_k(m)$分别表示第k个频率位置上原始值为m的DCT系数中已经向负向调整，正向调整的个数。

(2) 设即将处理的 DCT 系数处在第 k 个频率位置上，原始值为 m，并且需要调整。如果

$$B_{\text{temp}}\le B_0 \tag{10}$$

这里B_0是原始图像的分块特性，且

$$\frac{u_k(m)}{u_k(m)+v_k(m)}-0.05\le\frac{s_k(m)}{s_k(m)+t_k(m)}\le\frac{u_k(m)}{u_k(m)+v_k(m)}+0.05 \tag{11}$$

则将该 DCT 系数以概率 $\dfrac{u_k(m)}{u_k(m)+v_k(m)}$ 向负向调整，以概率 $\dfrac{v_k(m)}{u_k(m)+v_k(m)}$ 向正向调整。

(3) 如果不满足式(11)，若

$$\frac{s_k(m)}{s_k(m)+t_k(m)}>\frac{u_k(m)}{u_k(m)+v_k(m)}+0.05,$$

将当前的 DCT 系数向正向调整；若

$$\frac{s_k(m)}{s_k(m)+t_k(m)}<\frac{u_k(m)}{u_k(m)+v_k(m)}-0.05,$$

将当前的 DCT 系数向负向调整。

(4) 如果满足式(11)，但不满足式(10)，分别计算向正向调整和向负向调整后本块周边像素与相邻块相邻像素的灰度之差的绝对值总和，即本块在调整后对图像分块特性的贡献，比较两者，将当前的 DCT 系数向那个使调整后的分块贡献较小的方向调整。

(5) 转向步骤(1)，更新B_{temp}，$s_k(m)$，$t_k(m)$，以便对下一个DCT系数进行调整。直至嵌入所有秘密信息或用尽所有非 0 的非直流DCT系数。

在上述过程中，步骤(2)的目的是随机调整以嵌入数据，步骤(3)的目的是维持原有的 DCT 系数直方图，步骤(4)的目的是维持原有的分块特性。这里，维持直方图不变比维持分块特性有更高的优先级。即：只要已作负向调整的系数个数与已作正向调整的系数个数的比例偏离了既定的范围，就要强迫设定下一个系数的调整方向(步骤(3))。而当这个比例在既定范围内时，还要确保分块特性不超出初始值(步骤(4))。

尽管信息嵌入过程略为复杂，但提取过程非常简单。将含密图像的非 0 非直流 DCT 系数取出，如果除以量化步长后是正奇数或负偶数，则代表秘密比特 1；如果除以量化步长后是负奇数或正偶数，则代表秘密比特 0。

4　模拟实验

以大小为 512×512，质量因子为 70 的JPEG灰度图像Man为原始载体，按照本文方法进行密写，可以嵌入 4.5×10^4个秘密比特，秘密信息长度与密写后文件长度的比值为 0.15，即嵌入率为 15%(本方案的嵌入率与载体图像的质量因子和特性有关)。密写造成的峰值信噪比(PSNR)为 35.1dB，密写图像与原始图像几乎没有视觉差异，如图 1 所示。

根据式(1)计算的原始图像分块特性值B_0为 5.40×10^5，含秘密信息的图像的分块特性值依然为 5.40×10^5，利用文献[10]的方法估计出的分块特性值B_E为 5.51×10^5。如果不采用本文

(a) 原始 JPEG 图像　　(b) 密写后的图像

图 1

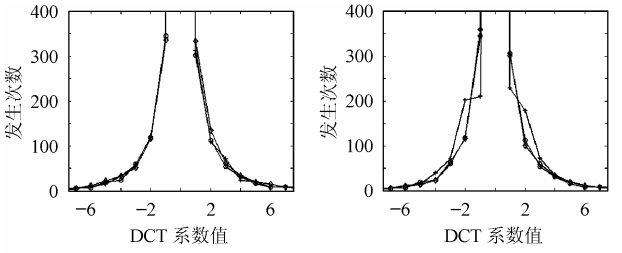
提出的安全密写方法，而是随机地将DCT系数值向负向或正向调整，密写后的分块特性值变为 6.56×10^5，与原始图像相

比有明显增加，因此密写分析可以检测出秘密信息的存在。而采用本文方案密写前后的分块特性几乎不变，所以能有效地抵抗基于分块特性的密写分析。

另一方面，原始图像的DCT系数直方图也能得到很好的保持。将原始图像每个小块中(3,3)位置上的DCT系数取出，统计其直方图，记为$h_O(m)$；分别统计利用安全密写方案和随机调整方案得到的含密图像中的(3,3)DCT系数直方图，记为$h_S(m)$，$h_R(m)$；再用文献[9]中的方法由密写图像生成参考图像，并将参考图像的直方图作为原始直方图的估计，由不同密写图像得到的原始直方图估计是不同的，分别记为$h_{SE}(m)$，$h_{RE}(m)$，结果如图 2。可以看出参考图像的直方图$h_{SE}(m)$，$h_{RE}(m)$和安全密写方案得到的直方图$h_S(m)$比较接近原始直方图$h_O(m)$，而随机调整方案得到的直方图$h_R(m)$与原始直方图$h_O(m)$差距较大，因此随机调整方案是不安全的。用下式衡量其它 4 个直方图与原始直方图的差异程度：

$$\rho=\frac{\sum_{m\neq 0}\left|h(m)-h_O(m)\right|}{\sum_{m\neq 0}h_O(m)} \tag{12}$$

将$h_S(m)$，$h_{SE}(m)$，$h_R(m)$，$h_{RE}(m)$分别代入式(12)中的$h(m)$，得到$\rho_S(m)$，$\rho_{SE}(m)$，$\rho_R(m)$，$\rho_{RE}(m)$分别为 0.02，0.10，0.31，0.09。可见安全密写方案得到的直方图比估计的原始直方图更接近原始直方图，所以利用直方图估计的密写分析同样无法察觉秘密信息的存在；而随机调整方案得到的直方图远远偏离了原始直方图和估计直方图，所以并不安全。



(a) (3,3)位置DCT系数的原始直方图$h_O(m)$，使用安全方案密写后的直方图$h_S(m)$和估计直方图$h_{SE}(m)$

(b) (3,3)位置DCT系数的原始直方图$h_O(m)$，使用随机方案密写后的直方图$h_R(m)$和估计直方图$h_{RE}(m)$

图 2

用数字照相机采集 100 幅包含人物与风景的图片，大小为 300×400，灰度值为 8bit，以质量因子 70 压缩后作为原始图像集。然后分别以安全方案和随机方案对原始图像进行密写，得到安全方案含密图像集与随机方案含密图像集。对这 3 个图像集进行基于分块特性的密写分析，即首先计算待检测图像的分块特性B_1，然后由待检测图像生成参考图像，并计算参考图像的分块特性B_E，记

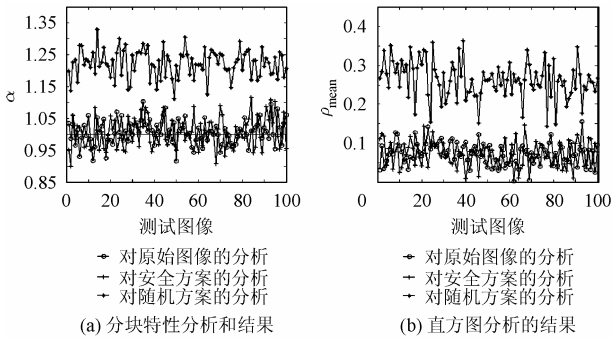
$$\alpha=B_1/B_E \tag{13}$$

图 3(a)给出了对 3 个图像集的分析结果。对于随机调整方法，分块特性在密写时显著提高，因此α值明显大于 1，不能抵抗分块特性分析。而对于原始图像和安全密写方案，α值始终在 1 附近并混迭在一起，因此分块特性分析对安全密写方案是无效的。事实上，使用安全密写方案密写后的分块特性与原始分块特性非常接近，α值在 1 附近波动是因为对原始分块特性的估计并不特别准确。

类似地，对这 3 个图像集进行直方图分析。设待检测图像第k个频率位置上的DCT系数直方图为$h_k^R(m)$，由待检测图像得到的参考图像中第k个频率位置上的DCT系数直方图为$h_k^E(m)$。计算

$$\rho_k=\frac{\sum_{m\neq 0}\left|h_k^R(m)-h_k^E(m)\right|}{\sum_{m\neq 0}h_k^E(m)} \tag{14}$$

表示直方图之间的差异，并对每一幅图像求出ρ_k的平均值ρ_{mean}。图 3(b)显示了对 3 个图像集进行直方图分析得到的ρ_{mean}。由随机方案得到的ρ_{mean}较大，说明密写引起了较明显的直方图失真。而由原始图像和安全密写方案得到的ρ_{mean}较接近于 0 并混迭在一起，表示待检测图像的DCT系数直方图与参考图像的DCT系数直方图相差不大，无法用直方图分析的方法检测出秘密信息的存在。



(a) 分块特性分析和结果

(b) 直方图分析的结果

图 3 对原始图像，安全方案和随机方案产生的密写图像分析的结果

当载体图像的压缩因子较高时，本文所述的密写方案有较好的特性。载体图像的质量因子越小，量化 DCT 系数时采用的步长就越大，在密写的同时既不改变直方图又不增大图像分块特性就比较困难。因为维持直方图不变有较高的优先级，所以载体图像的 DCT 系数直方图可以不受压缩因子影响而得到保持，但分块特性有可能因为密写而增加。以不同质量因子的图像 Man 为原始载体，对原始图像及密写后的图像进行分块特性分析，得到的α值如图 4。当载体图像质量因子大于 35 时，无法根据α值分辨密写图像；但当载体图像

质量因子小于 35 时，密写图像的分块特性有比较明显的增加。可见，以质量因子大于 35 的 JPEG 图像作为载体是该密写方案安全性的一个重要保证。

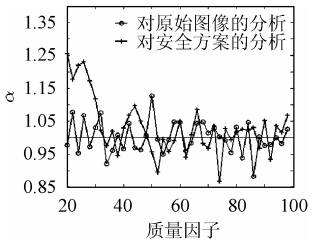


图 4　以不同质量因子 JPEG 图像为载体时的分块特性分析

5　结束语

本文以 JPEG 图像为载体设计了一种新的密写方案。该方案在将秘密信息嵌入非 0 的非直流 DCT 系数时，既不改变原始图像的 DCT 系数直方图，也不增加载体图像的分块效应，因此可以有效抵抗基于直方图失真或分块特性的密写分析。

在本方案中，维持直方图不变比维持分块特性有更高的优先级。实际上，也可以按照采用维持分块特性具有较高优先级的原则设计嵌入方法。无论采用哪种优先级排序，当载体图像的质量因子过低时，直方图和分块特性较难兼顾。实验表明，质量因子大于 35 的图像作为载体时，安全性可以得到有效的保证。

参 考 文 献

[1] Petitcolas F A P, Anderson R J, Kuhn M G. Information hiding — A survey. *Proc. IEEE*, 1999, 87(7): 1062 − 1078.

[2] Wang H, Wang S. Cyber warfare — Steganography vs. steganalysis. *Communication of ACM*, 2004, 47(10): 76 − 82.

[3] Fridrich J, Goljan M. Practical steganalysis of digital images — State of the art. in Security and Watermarking of Multimedia Contents IV, Proceedings of SPIE, 2002, 4675: 1 − 13.

[4] Zhang X, Wang S, Zhang K. Steganography with the least histogram abnormality. Computer Network Security, Lecture Notes in Computer Science, Springer-Verlag, 2003, 2776: 395 − 406.

[5] Wang S, Zhang X, Zhang K. Steganographic technique capable of withstanding RQP analysis. *Journal of Shanghai University*, 2002, 6(4): 273 − 277.

[6] Zhang X, Wang S. Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security. *Pattern Recognition Letters*, 2004, 25(3): 331 − 339.

[7] Chang C C, Chen T S, Chung L Z. A steganographic method based upon JPEG and quantization table modification. *Information Sciences*, 2002, 141(1-2): 123 − 138.

[8] Provos N. Defending against statistical steganalysis. in 10th USENIX Security Symposium, Washington, DC, 2001: 323 − 335.

[9] Westfeld A. F5 — A steganographic algorithm. in 4th International Workshop on Information Hiding, Pittsburgh, PA, USA, Lecture Notes in Computer Science, Springer-Verlag, 2001, 2137: 289 − 302.

[10] Fridrich J, Goljan M, Hogea D. Steganalysis of JPEG image: Breaking the F5 algorithm. in 5th International Workshop on Information Hiding, Noordwijkerhout, The Netherlands, Lecture Notes in Computer Science, Springer-Verlag, 2002, 2578: 310 − 323.

[11] Fridrich J, Goljan M, and Hogea D. Attacking the OutGuess. in Proc. of the ACM Workshop on Multimedia and Security 2002, Juan-les-Pins, France, December 6, 2002. http://www.ws.binghamton.edu/fridrich/Research/acm_outguess.pdf

[12] Fridrich J, Goljan M, Du R. Steganalysis based on JPEG compatibility. in Special Session on Theoretical and Practical Issues in Digital Watermarking and Data Hiding, SPIE Multimedia Systems and Applications IV, Proceedings of SPIE, Denver, CO, August 20-24, 2001, 4518: 275 − 280.

张新鹏：　男，1975年生，博士，讲师，研究方向为信息隐藏、数字图像处理.

王朔中：　男，1943年生，教授，博士生导师，研究方向为图像处理、音频信号处理、信息隐藏.