doi:10.11959/j.issn.1000-436x.2015166

基于重复压缩的密文图像可逆数据隐藏方法

刘九芬[1,2,3]，韩涛[1,2]，田雨果[1,2]，刘文彬[1,2]

（1. 信息工程大学，河南 郑州 450001；2. 数学工程与先进计算国家重点实验室，河南 郑州 450001；
3. 中国科学院 信息工程研究 所信息安全国家重点实验室，北京 100093）

摘 要：为了提高密文图像上的可逆数据隐藏方法的性能，在加密前处理图像以获得数据嵌入空间。首先，分析了在解压缩分块上进行可逆嵌入的可行性；其次，计算了成功恢复解压缩分块的理论概率；最后，提出了基于解压缩分块的密文图像可逆数据隐藏算法，主要过程包括预处理、加密、数据嵌入、数据提取与图像恢复。从图像恢复错误率、嵌入容量和 *PSNR* 这 3 个方面与 3 种已有方法进行对比，表明所提方法实现了数据提取和图像解密在操作上的完全分离，且图像恢复错误率更低，嵌入容量更大，*PSNR* 更高。

关键词：重复压缩；可逆数据隐藏；图像加密；隐私保护；差分扩展

中图分类号：TP309 **文献标识码**：A

Reversible data hiding in encrypted images using recompression

LIU Jiu-fen[1,2,3], HAN Tao[1,2], TIAN Yu-guo[1,2], LIU Wen-bin[1,2]

(1. Information Engineering University, Zhengzhou 450001, China;
2. State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China;
3. State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China)

Abstract: To improve the performance of RDH (reversible data hiding) method in encrypted images, the embedding room was reserved before encryption and realizes the real separation of the data extraction from image decryption in operation. First, the feasibility of RDH in JPEG decompressed image blocks was researched. Then the theoretical probability of successful recovery of decompressed image blocks was calculated. Finally, a method of RDH in encrypted images was proposed and the main steps of the method include pretreatment, encryption, data embedding, data extraction and image recovery. The performance of the proposed method was compared with three existing RDH methods in encrypted images. The results demonstrate the proposed method has less error in image recovery, and for given embedding rates, the *PSNR* of decrypted image containing the embedded data are significantly improved.

Key words: recompression; RDH; image encryption; privacy protection; difference expansion

1 引言

可逆数据隐藏又称可逆数字水印，是数字水印领域的一个重要分支。它不仅能够完整地提取嵌入秘密消息，还能无损地恢复载体数据，被广泛应用于军事、医学、法律等对原始载体数据要求较高的领域。可逆数据隐藏的主要目标是在达到较大嵌入容量的同时保持较小的失真。在医疗诊断、法庭取证等实际应用中，可逆数据隐藏具有重要使用价值。 近年来，作为信息隐藏技术的研究热点之一，可逆数据隐藏得到了快速的发展。

在理论方面，Kalker[1]建立了可逆数据隐藏的率失真模型，并对于无记忆载体信号给出了可逆数据隐藏的率失真界，最后提出了一种递归码构造方法，但该构造达不到率失真界。Zhang[2,3]改进了二元载体信号的递归码构造方法，并证明了在压缩算法能达到熵的情况下该构造能达到率失真界，这就建立了数据压缩与二元载体信号的可逆数据隐藏

收稿日期：2014-08-16；**修回日期**：2014-10-15
基金项目：国家自然科学基金资助项目(61379151)
Foundation Item: The National Natural Science Foundation of China (61379151)

之间的等价性。

在应用方面，近年来，已有大量的相关研究成果涌现出来。Fridrich[4]提出了可逆数据隐藏的一般性框架：首先提取原始载体信号的可压缩特征，然后将特征进行无损压缩，最后将剩余空间用于承载额外的秘密消息。Tian[5]提出了一种基于差值扩展的可逆数据隐藏方法，主要思路是对每个像素分组的差值进行扩展（比如乘以2），这样差值的LSB(least significant bit)位就会全变为0，即可用于承载秘密消息。Ni[6]提出了一种基于直方图平移的可逆数据隐藏方法，该方法主要通过平移图像直方图来获得数据嵌入的空间。已有多种可逆数据隐藏方法[7~9]将差值扩展或直方图平移与图像残差（例如预测误差）结合以获得更好的隐藏性能。

关于图像的隐私保护，加密是一种有效且流行的方法，可将原始有意义的明文都变换为毫无意义的密文[10]。而基于加密图像的可逆数据隐藏具有特定的应用背景和前景。比如，云服务的提供者不能在给加密数据进行数据着色(将数据嵌入到载体中的一种方式)时引入永久失真，所以其更倾向于选择基于加密数据的可逆数据着色技术。假定一幅医学图像存储在数据中心，数据中心的服务器在加密的医学图像上使用可逆数据隐藏技术嵌入一些标记，这样就可以在不知道图像内容的情况下管理图像或者验证其完整性，病人的隐私即得到了保护。另一方面，拥有密钥的医生经过解密和无损恢复得到原始图像，以便进行进一步的诊断。Hwang[11]提出了一种基于信誉的信任管理算法，该算法主要使用数据着色和软件水印以提高安全性，数据加密和数据着色技术可保护数据拥有者的隐私和数据完整性。

目前，基于密文图像的可逆数据隐藏方法得到了一些研究者的关注。Zhang[12]提出了一种基于密文图像的可逆数据隐藏方法，首先将密文图像分为互不相交的像素块，然后通过翻转每块中一半像素的3层LSB位在该块嵌入1 bit数据。在接收端，通过寻找解密图像上每个像素块中的翻转部分进行数据提取和图像恢复，该过程的实现主要借助于解密图像的空域相关性。Hong[13]通过使用一个不同的平滑度计算方法和边缘匹配技术改进了Zhang[12]在接收端的方法，更加充分地利用了空域相关性，减小了载体恢复的错误率。以上提到的2种方法在提取数据时都依赖于原始图像的空域相关性，在提取数据前需要先解密加密图像。此时图像内容无法对数据提取者保密。

为了将数据提取过程和图像解密过程分开，Zhang[14]根据压缩加密图像[15,16]的思想来获取数据嵌入的空间。Johnson[15]指出，加密数据的压缩可形式化为接收端上有边信息的信源编码，其中，一种典型的方法就是通过使用信道编码中奇偶校验矩阵的伴随式来无损地压缩数据。Zhang[14]通过寻找奇偶校验矩阵的伴随式来压缩加密的LSB平面，以获得秘密消息的嵌入空间，在接收端使用的边信息仍然是解密图像的空域相关性。Zhang[14]虽然实现了数据提取与图像解密的分离，但是仅能在加密图像中提取数据。

文献[12~14]所述3种方法都是从加密图像上腾出数据嵌入空间。然而，当加密图像的熵最大化时，这3种方法都只能达到很小的嵌入率，或在大嵌入率情况下会生成图像质量很差的直接解密图像(包含嵌入数据的解密图像)，同时在数据提取或图像恢复时还会有一定的错误率。Ma[17]提出并使用“加密前预留数据嵌入空间”的嵌入模型，有效地克服了文献[12~14]的缺点。

本文沿用Ma[17]提出的嵌入模型，提出一种基于密文图像的可逆数据隐藏方法。在图像加密之前对图像进行预处理，预留出数据嵌入空间，使用一种传统的可逆数据隐藏方法将用于恢复图像的信息嵌入到其他像素中，然后对图像进行加密，使用加密图像中的预留嵌入空间来承载秘密消息。最终实现图像接收端在加密图像和解密图像上都能提取消息。同文献[12~14]方法相比，本文方法能够在如下2个方面达到良好的性能：1)降低了数据提取和图像恢复的错误率；2)对于给定的嵌入率，提高了含有嵌入数据的解密图像的PSNR (peak signal to noise ratio)，同时对于可接受的PSNR，扩大了嵌入率范围。

2 相关工作

文献[12~14]中的方法都是在图像加密后腾出数据嵌入空间的。首先使用带加密密钥的标准密码算法来加密原始图像，然后将加密图像发给数据嵌入者(比如数据库管理员)，数据嵌入者根据隐写密钥以无损的方式腾出加密图像上的嵌入空间来承载秘密消息。接收者既可能是图像拥有者也可能是可信第三方，可根据数据隐写密钥提取嵌入数据，同时根据加密密钥恢复原始图像。

3 种方法中的加密图像均是使用流密码算法对原始灰度图像的每个比特平面进行加密得到的。

Zhang[12]方法首先将加密图像分为大小为 $a\times a$ 的互不重叠的像素块，每个像素块承载 1bit 数据。然后根据消息嵌入密钥将每个像素块伪随机地分为 2 个集合 S_1 和 S_2，若待嵌比特为 0，则翻转 S_1 中每个加密像素的最低 3 个 LSB 平面；否则翻转 S_2 中每个加密像素的最低 3 个 LSB 平面。为了提取秘密消息和恢复原始图像，接收者首先解密图像，翻转 S_1 中像素的最低 3 个 LSB 平面来形成一个 $a\times a$ 的像素块，记为 S_D^1，然后翻转 S_2 中像素 3 层 LSB 位来形成另外一个 $a\times a$ 的像素块，记为 S_D^2，S_D^1 和 S_D^2 中将会有一个为原始像素块。由于自然图像的空域相关性，原始像素块将会更加平滑，所以选择 S_D^1 和 S_D^2 中纹理更为平滑的像素块作为原始像素块，并相应地提取出嵌入比特。而当像素分块比较小或者纹理比较复杂时，数据提取和图像恢复可能会出现错误。

通过完全利用像素来计算每个像素块的平滑程度以及使用边缘匹配技术，Hong[13]降低了 Zhang[12]的错误率，首先计算 2 个候选像素块的平滑程度，然后根据两者的差分绝对值的降序排列来实现该像素块上的数据提取和载体恢复，已恢复块还可用于度量未恢复块的平滑程度，即边缘匹配。

Zhang[12] 和 Hong[13]方法必须在解密图像上提取消息，而且每个分块上的数据提取与图像恢复是关联在一起的，有可能同时出错，并且 a 越小，出错率越高。对于大小为 N 的图像，近似嵌入率为 $\frac{1}{a^2}$，计算复杂度为 $O(N)$。

Zhang[14]方法首先对加密图像进行置乱，并将置乱后的加密图像分为一些大小为 L 的像素分组，然后使用一个奇偶校验矩阵压缩每个像素分组的最低 k 个 LSB 平面，将 kL bit 数据压缩为 $(kL-S)$bit，将 S bit 数据接在压缩数据后面，替换原始的最低 k 个 LSB 平面，实现数据嵌入。比如，设一个加密图像像素分组为 $x_1,x_2,\cdots,x_L$，其最低 k 个 LSB 平面为 c，长度为 kL bit。数据嵌入者生成一个大小为 $(kL-S)(kL)$ 的奇偶校验矩阵 $\boldsymbol{G}$，压缩 c 为其伴随式 s，即 $s=\boldsymbol{G}c$，由于 s 的长度为 $(kL-S)$，所以将 S bit 可以用于承载秘密消息。在接收端，直接解密图像可以得到每个像素的前 $(8-k)$ 个 MSB (multiple significant bit)平面，然后根据邻域像素的这些 MSB 平面来估计 $x_i\ (1\leqslant i\leqslant L)$，得到 c 的估计值 c'。另一方面，接收者测试伴随式 s 的陪集 $\Omega(s)$ 中每个向量，其中，$\Omega(s)=\{u\,|\,\boldsymbol{G}u=s\}$，可从 $\Omega(s)$ 的每个向量都得到一个 c 的恢复值，选择一个最接近估计值 c' 的恢复值作为恢复的 LSB 平面。

Zhang[14]方法可以准确地在加密图像上提取消息，但恢复原始图像时可能会出现错误。尤其当 L、k 较小或 S 较大时更加容易出错。对于大小为 N 的图像，近似嵌入率为 $\frac{S}{L}$，计算复杂度为 $O(N2^S)$。

3　可恢复块上实现可逆嵌入可行性

本节介绍本文的理论基础。需要说明的是，本文中所指的分块大小均为 8×8。

3.1　重复压缩与可恢复块

记 C 为原始空域图像，对 C 进行质量因子为 q 的 JPEG 压缩，得到 JPEG 图像 J_C。对 J_C 进行相同质量因子的 JPEG 解压缩，得到空域图像 I_C，称 I_C 为 J_C 的解压缩图像。记质量因子 q 对应的量化矩阵为 $\boldsymbol{Q}$，且 $\boldsymbol{Q}$ 的系数为 $\boldsymbol{Q}(i)(1\leqslant i\leqslant 64)$。

定义 1　(重复压缩) I_C 为 J_C 的解压缩图像，I_S 为 I_C 对应的载密图像(无嵌入数据时 $I_S=I_C$)，对 I_S 再次进行质量因子为 q 的 JPEG 压缩，再解压缩生成非压缩格式图像 I_C'，将这个过程称为 JPEG 重复压缩，简称为重复压缩。

记 X 为解压缩分块，$\mathrm{DCT}(X)$ 为对 X 进行 DCT 变换后的分块，$d_X(i)$ 为其系数，$\boldsymbol{D}(X)$ 为对 $\mathrm{DCT}(X)$ 使用量化矩阵 $\boldsymbol{Q}$ 量化后的系数矩阵，$D_X(i)$ 为其系数，$RC(X)$ 为对 X 进行重复压缩得到的分块，其中，$1\leqslant i\leqslant 64$。

记 A 是 C 上分块，B 是 I_C 上相同位置上的分块，S 为 B 对应的载密分块，$\boldsymbol{N}$ 为对 B 进行数据嵌入所引入的加性噪声矩阵，即

$$S=B+\boldsymbol{N}\text{。}$$

下面给出对 8×8 载密分块 S 的重复压缩过程。

1) DCT 变换。由 DCT 变换的可加性知

$$\mathrm{DCT}(S)=\mathrm{DCT}(B)+\mathrm{DCT}(N)$$

2) 量化。有

$$D_S(i)=\mathrm{round}\left(\frac{d_S(i)}{Q(i)}\right)=\mathrm{round}\left(\frac{d_B(i)+d_N(i)}{Q(i)}\right)$$

其中，$1\leqslant i\leqslant 64$，round(·) 表示四舍五入运算。

3) 以非压缩形式保存图像数据。则

$$RC(S)=\mathrm{int}_8\left(\mathrm{DCT}^{-1}\left(QD(S)\right)\right)$$

其中，$QD(S)$ 的系数为 $Q(i)D_X(i)$，$1\leqslant i\leqslant 64$；$\mathrm{DCT}^{-1}(\cdot)$ 表示图像逆 DCT 变换；$\mathrm{int}_8(\cdot)$ 表示 8 bit 取整处理，即将小于 0 的值变为 0，将大于 255 的值变为 255，并对 0 与 255 之间的值四舍五入，前 2 种处理为溢出处理，后一种为舍入处理。

定义 2 (不可恢复块)对解压缩载体分块 B，若 $RC(B)=B$，则称分块 B 为可恢复块；否则，称其为不可恢复块。可恢复块的数量与其量化矩阵和截断误差有关，与图像的纹理复杂度无关。量化步长越短，或截断误差出现次数越多，可恢复块数量就会越少。采用 NRCS 图像库[18]的 3 000 幅灰度图像，将其转换成质量因子为 $q(q=50,51,\cdots,98)$ 的 JPEG 解压缩图像，统计在质量因子 q 下可恢复块出现的频率，结果如图 1 所示。可以看出：1)当质量因子小于等于 92 时，可恢复块出现的频率很接近 1，并且会随着质量因子的增加而略微增加，这是因为截断误差会随着质量因子的增加而减少[19]，导致可恢复块的数量增多；2)当质量因子大于等于 93 时，随着质量因子的增加，其可恢复块出现的频率会骤降，这是因为此时量化矩阵左上角某些频率位置上的量化步长等于 1，量化步长 1 将无法去除在 $[-0.5, 0.5)$ 以外的噪声而导致该块不可恢复，并且质量因子越大，量化步长 1 的个数越多，从而不可恢复块出现的概率就越大。

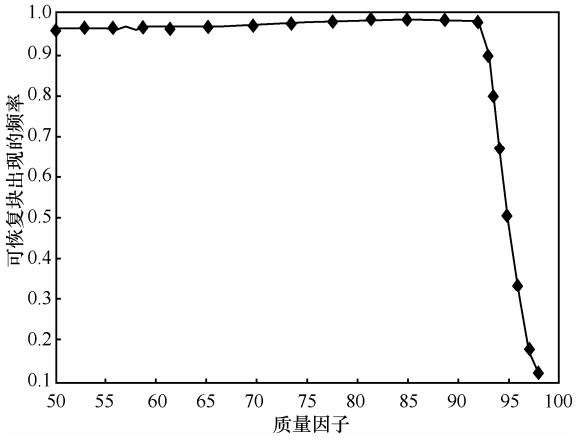


图 1 不同质量因子下可恢复块出现的频率

3.2 可恢复块上实现可逆嵌入的充分条件

对于可恢复块 B 及其对应的载密分块 S，如果有 $RC(S)=B$，则通过对载密分块 S 进行重复压缩就能得到载体分块 B，实现载体恢复。显然，是否有 $RC(S)=B$ 是可逆嵌入成功的关键，这与载密图像 I_S 的生成算法有关。

定理 1 设某一 8×8 分块 B 为可恢复块，对 B 进行数据嵌入后得到 S，若 $D(S)=D(B)$，则有 $RC(S)=B$。

证明
$$\begin{aligned}RC(S)&=\mathrm{int}_8\left(\mathrm{DCT}^{-1}\left(QD(S)\right)\right)\\&=\mathrm{int}_8\left(\mathrm{DCT}^{-1}\left(QD(B)\right)\right)\\&=RC(B)=B\end{aligned}$$

证毕。

定理 1 显示，对于可恢复块 B 和其对应的载密分块 S，只要保证 $D(S)=D(B)$，那么该载密分块 S 就可通过重复压缩得到原始载体分块 B，实现数据的可逆嵌入。下面定理给出了 $D(S)=D(B)$ 的充要条件。

定理 2 条件同定理 1，空域嵌入噪声矩阵为 $\boldsymbol{N}$，使得 $D(S)=D(B)$ 成立的充要条件为

$$L_B^q(i)\leqslant d_N(i)<H_B^q(i)$$

其中，$H_B^q(i)$ 由下式表示

$$\begin{aligned}L_B^q(i)&=\frac{-Q(i)}{2}-\varepsilon_B(i)\\H_B^q(i)&=\frac{Q(i)}{2}-\varepsilon_B(i)\end{aligned}\tag{1}$$

$\varepsilon_B(i)$ 为 $\mathrm{int}_8(\cdot)$ 操作所引入的频域加性噪声

$$\varepsilon_B(i)=d_B(i)-D_B(i)Q(i)\tag{2}$$

其中，$1\leqslant i\leqslant 64$。

证明 由重复压缩的量化过程知，$D(S)=D(B)$ 成立的充要条件为

$$D_B(i)=D_S(i)=D_B(i)+\mathrm{round}\left(\frac{\varepsilon_B(i)+d_N(i)}{Q(i)}\right)$$

由于 round(·) 为四舍五入取整运算，从而使上式成立的充要条件为

$$-\frac{1}{2}\leqslant\frac{\varepsilon_B(i)+d_N(i)}{Q(i)}<\frac{1}{2}\tag{3}$$

其中，$1\leqslant i\leqslant 64$。

结合式(1)和式(3)，可知定理 2 成立。

证毕。

由定理 2 知，只要数据嵌入过程引入的每一个频域加性噪声 $d_N(i)$ 都在 $L_B^q(i)$ 与 $H_B^q(i)$ 之间 $(1\leqslant i\leqslant 64)$，嵌入就可以成功。$L_B^q(i)$ 和 $H_B^q(i)$ 可根据图像分块 B 和质量因子 q 由式(1)和式(2)直接得到，此时若能确切知道频域加性噪声 $d_N(i)$ 的概率分布以及相应的概率密度函数，即可计算出 $d_N(i)$ 出现在 $L_B^q(i)$ 与 $H_B^q(i)$ 之间的概率，即 $D(S)=D(B)$ 的概率。

3.3 数据嵌入成功的理论概率

对于解压缩分块 B，本文利用 k（k=1,2）层 LSB 替换算法进行嵌入，得到相应载密分块 S。若 $RC(S)=B$，则数据嵌入成功。若 B 为不可恢复块，对相应的 S 进行重复压缩得到的 $RC(S)$ 一般不等于 B。所以，仅选择可恢复块进行数据嵌入。由 3.1 节可知，大部分情况下，可恢复块出现的频率较高，保证了可恢复块能提供较高的嵌入容量。下面讨论对于可恢复块 B 和相应载密分块 S，$RC(S)=B$ 成立的理论概率，即数据嵌入成功的理论概率。

定理 3 给出频域加性噪声 $d_N(i)$ 的概率分布以及相应的概率密度函数。

定理 3　对解压缩图像 I_C 的每一个可恢复分块 B，随机选择 $l(1\leqslant l\leqslant 64)$ 个像素利用 k（k=1, 2）层 LSB 替换嵌入数据，空域噪声矩阵为 $\boldsymbol{N}$，则 DCT($\boldsymbol{N}$) 的每个频率位置上的系数 $d_N(i)$ 都近似地服从如下高斯分布，即

k=1 时，

$$d_N(i)\sim N\left(0,\frac{l}{128}\right)$$

k=2 时，

$$d_N(i)\sim N\left(0,\frac{5l}{128}\right)$$

其中，$1\leqslant i\leqslant 64$。

证明　以 k=2 为例，此时 $\boldsymbol{N}$ 中每一位置取值集合为 $\{3,2,1,0,-1,-2,-3\}$，且具体概率如下

$$p(0)=1-\frac{3l}{256},\ p(1)=p(-1)=\frac{3l}{1\,024}$$

$$p(2)=p(-2)=\frac{l}{512},\ p(3)=p(-3)=\frac{l}{1\,024}$$

由于 $\boldsymbol{N}$ 的系数两两相互独立，由中心极限定理，每一个 $d_N(i)$ 都近似地服从高斯分布[20]，在此记为 $N(\mu,\sigma^2)$。下面计算 μ 和 σ^2，从而确定具体的分布。

根据 DCT 变换，有

$$d_N(i)=\sum_{j=1}^{64}a_i(j)N(j)$$

其中，$1\leqslant i\leqslant 64,1\leqslant j\leqslant 64$，$a_i(j)$ 为 DCT 变换矩阵的系数

$$a_i(j)=\frac{1}{4}w\left(\left\lfloor\frac{j}{8}\right\rfloor\right)w(j\bmod 8)\cdot\cos\left(\frac{\pi}{16}\left(2\left\lfloor\frac{i}{8}\right\rfloor+1\right)\left\lfloor\frac{j}{8}\right\rfloor\right)\cdot\cos\left(\frac{\pi}{16}(2(i\bmod 8)+1)\left\lfloor\frac{j}{8}\right\rfloor\right)$$

其中，$\lfloor\cdot\rfloor$ 表示向下取整。若 t=0，$w(t)$=$\frac{\sqrt{2}}{2}$；若 $t\neq 0$，$w(t)$=1。

从而，$d_N(i)$ 的数学期望 μ 为

$$\mu=\mathrm{E}\left(d_N(i)\right)=\sum_{j=0}^{63}a_i(j)\mathrm{E}\left(N(j)\right)\tag{4}$$

由于

$$\mathrm{E}\left(N(i)\right)=0p(0)+\sum_{j=1}^{3}jp(j)+\sum_{j=1}^{3}-jp(j)=0\tag{5}$$

将式(5)代入式(4)，即得 $\mu=0$。而 $d_N(i)$ 的方差 σ^2 为

$$\begin{aligned}\sigma^2&=\mathrm{D}\left(d_N(i)\right)=\mathrm{E}\left(d_N{}^2(i)\right)-\mu^2\\&=\sum_{j=1}^{64}a_i{}^2(j)\mathrm{E}\left(N^2(j)\right)+\\&\quad\sum_{j=1}^{64}\sum_{k=1,k\neq j}^{64}a_i(j)a_i(k)\mathrm{E}\left(N(j)N(k)\right)-\mu^2\end{aligned}\tag{6}$$

由于 $N(j)$ 与 $N(k),k\neq j$ 两两相互独立以及 $\mu=0$，从而

$$\mathrm{E}\left(N(j)N(k)\right)=\mathrm{E}\left(N(j)\right)\mathrm{E}\left(N(k)\right)=0$$

进而有

$$\sum_{j=1}^{64}\sum_{k=1,k\neq j}^{64}a_i(j)a_i(k)\mathrm{E}\left(N(j)N(k)\right)=0\tag{7}$$

根据 DCT 变换矩阵系数 $a_i(j)$ 的性质[21]，知

$$\sum_{j=1}^{64}a_i{}^2(j)=1\tag{8}$$

而

$$\mathrm{E}\left(N^2(i)\right)=0^2\left(1-\frac{3l}{256}\right)+\sum_{j=1}^{3}j^2p(j)+\sum_{j=1}^{3}(-j)^2p(j)=\frac{5l}{128}\tag{9}$$

其中，$1\leqslant i\leqslant 64$。将式(7)~式(9)和 $\mu=0$ 代入式(6)，即得 $\sigma^2=\frac{5l}{128}$。

对于 k=1，N 中每一位置取值集合为 $\{-1,0,1\}$，且具体概率如下

$$p(0)=1-\frac{l}{128},\ p(1)=p(-1)=\frac{l}{256}$$

同理可得 $\mu=0$，$\sigma^2=\frac{l}{128}$。

证毕。

应用定理 2 与定理 3，得到数据嵌入成功的理论概率。

定理 4 B 为可恢复块，且重复压缩使用的质量因子为 q，从 B 中随机选择 $l(1\leqslant l\leqslant 64)$ 个位置应用 $k(k=1,2)$ 层 LSB 替换嵌入数据，得到载密分块 S，则 $RC(S)=B$ 成立(即数据嵌入成功)的理论概率 $P_B(l,k,q)$ 可由下式计算

$$P_B(l,k,q)=\prod_{i=1}^{64}\int_{L_B^q(i)}^{H_B^q(i)} f(l,k,x)\mathrm{d}x \tag{10}$$

其中，$f(l,1,x)$ 为 $N\left(0,\frac{l}{128}\right)$ 的概率密度函数，$f(l,2,x)$ 为 $N\left(0,\frac{5l}{128}\right)$ 的概率密度函数。

证明 应用定理 2，有

$$P_B(l,k,q)=P\left\{L_B^q(i)\leqslant d_N(i)<H_B^q(i),1\leqslant i\leqslant 64\right\}$$

应用定理 3，每个 $d_N(i)$ 服从高斯分布，从而有

$$P\left\{L_B^q(i)\leqslant d_N(i)<H_B^q(i)\right\}=\int_{L_B^q(i)}^{H_B^q(i)} f(k,l,x)\mathrm{d}x \tag{11}$$

由于 DCT 变换是一种去相关性变换，具有良好的去相关性，从而可视 $d_N(i)$ 与 $d_N(k)$ 两两相互独立$(i\neq k)$，于是有

$$P_B(l,k,q)=\prod_{i=1}^{64}P\left\{L_B^q(i)\leqslant d_N(i)<H_B^q(i)\right\} \tag{12}$$

将式(11)代入式(12)即得结论，其中，$1\leqslant i\leqslant 64$。

证毕。

由定理 4 知，嵌入成功的概率与分块载密像素个数 l、质量因子 q 以及嵌入数据的层数 k 有关。分块载密像素个数 l 越大或者 k 越大，高斯分布的方差就越大，从而式(10)右边部分的积分值越小，即嵌入成功的概率 $P_B(l,k,q)$ 负相关于参数 l、k；质量因子 q 越大，量化步长越小，由式(1)和式(2)知 $H_B^q(i)-L_B^q(i)$ 越小，从而式(10)右边部分的积分值越小，即嵌入成功的概率 $P_B(l,k,q)$ 负相关于质量因子 q。

4 密文图像的可逆数据隐藏算法

上节描述了在解压缩图像可恢复块上进行可逆嵌入的可行性，本节讨论更一般的情况：如何在原始空域图像的加密图像上实现可逆数据隐藏算法。算法的基本流程如图 2 所示。

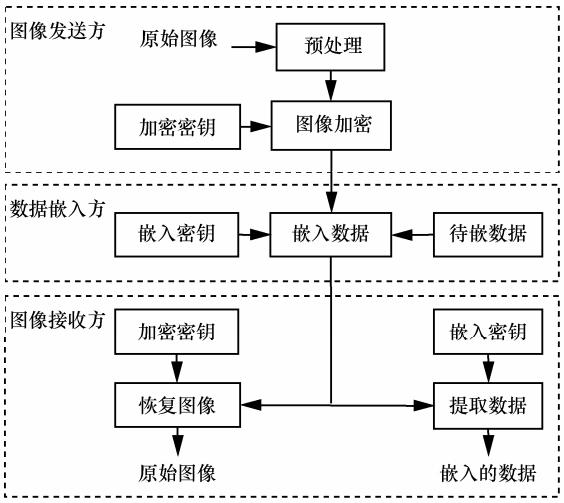


图 2 本文算法流程

图像发送方对图像进行预处理，预留出消息嵌入位置，使用加密密钥对图像加密后发送给数据嵌入方。数据嵌入方使用嵌入密钥在加密图像上预留的嵌入位置上嵌入消息。接收方根据不同的实际情况，可以在加密图像上提取消息，也可以在解密图像上提取消息，并且能以较低的错误率恢复原始图像。

对于大小为 $M\times N$ 的原始空域图像 C，按照以下步骤实现密文图像的可逆隐藏算法。

1) 预处理

预处理过程分为 2 个步骤：图像整合分区和自嵌入。图像整合分区指将图像分为 PA 和 PB 2 个部分，PA 部分中每个分块都是解压缩可恢复块，其 $k(k=1,2)$层 LSB 平面是预留出的消息嵌入空间，每个分块选取 l 个像素值进行嵌入。自嵌入指将恢复原始图像所需的关键信息可逆地嵌入到 PB 部分，这个过程需要用到传统的可逆数据隐藏算法。本文使用文献[5]中的基于差分扩展的可逆数据隐藏算法。

① 图像整合分区

首先将 C 分为大小为 8×8 的互不重叠的分块，记为 $A_{i,j}$，$1\leqslant i\leqslant\left\lfloor\frac{M}{8}\right\rfloor$，$1\leqslant j\leqslant\left\lfloor\frac{N}{8}\right\rfloor$。不失一般性，可设 $M \bmod 8=0$ 和 $N \bmod 8=0$。对 C 中每个分块进行质量因子为 q 的 JPEG 压缩和解压缩，得到解压缩空域图像 I_C，记 C 与 I_C 的像素差值为 $E=C-I_C=\{e_{i,j}\}$，$1\leqslant i\leqslant M$，$1\leqslant j\leqslant N$。显然，由 E 和 I_C 可以准确地恢复出原始图像 C。

从图像 I_C 中选择用来嵌入数据的图像块。首先排除 I_C 上的不可恢复块和不满足条件

$\max\left\{\left|e_{i,j}\right|\right\}\leqslant t$ 的分块，其中，(i,j) 取遍该分块内像素索引。其余块可作为用于嵌入数据的待选图像块。由定理 4 知，每个可恢复块上数据嵌入成功的概率一般不同，计算每个选出的待选图像块的嵌入成功的概率，并根据对嵌入容量的需求选择具有较大嵌入成功概率的 m 个待选图像块，作为嵌入数据的可用图像块。其中，参数 t 的选择和 m 的计算会在下文给出。

对图像 I_C 中图像块进行位置编排。使用一个长度等于分块个数的 01 串标记每个图像分块。将选出的 m 个可用图像块标记为“1”，剩余的其他图像块标记为“0”。按照从上到下、从左到右的顺序记录下连续的 01 串 $P=\{p_i\}$，其中，p_i =0 或 1，且 $i=1,2,\cdots,\left\lfloor\frac{MN}{64}\right\rfloor$。同时，将按照相同顺序遇到的第 i 个标记为“1”的图像块移到整幅图像的第 i 个块的位置上，其他图像块保持顺序依次后移，其中，$1\leqslant i\leqslant m$。至此，图像 I_C 被分为 2 部分：在图像头部由标记为“1”的图像块组成用来嵌入消息的部分 PA 和在图像尾部由标记为“0”的图像块组成的部分 PB。位置重新排列后，将 PB 部分中各个解压缩分块替换为原始图像上对应的分块。此时图像 I_C 变为新的图像 I'。至此图像的整合分区已经完成。

② 自嵌入

首先，生成需要自嵌入的信息。由于图像 I' 中 PA 部分都是解压缩分块，恢复图像时需要知道每个块与对应原始块在每个位置上的差值 $E_{\mathrm{PA}}=\{e_{i,j}\}$，其中，$(i,j)$ 取遍 PA 中所有像素索引。E_{PA} 和 P 就是需要可逆地嵌入到 PB 部分的标记信息。为了减少标记信息长度，在嵌入之前使用 LZW 编码压缩标记信息，记压缩后的信息为 M_1。P 压缩后长度固定，而像素差值压缩后的长度与具体差值 $e_{i,j}$ 有关，$\left|e_{i,j}\right|$ 越大，压缩后的长度也越大，所以选择可用图像块时限制条件 $\max\limits_{(i,j)\in B}\left\{\left|e_{i,j}\right|\right\}\leqslant t$。实验中，$t$ 取经验值 7。这样除符号外，每个差值可用至多 3 bit 数据表示。

下面简要介绍如何利用文献[5]方法实现自嵌入过程。

沿用文献[5]中定义。记灰度像素值对 (x,y) 的整数均值 ave 和差分 dif 分别为

$$ave=\left\lfloor\frac{x+y}{2}\right\rfloor,dif=x-y$$

其逆运算为

$$x=ave+\left\lfloor\frac{dif+1}{2}\right\rfloor,y=ave-\left\lfloor\frac{dif}{2}\right\rfloor \tag{13}$$

记该像素组对应的待嵌数据为 b(b=0 或 1)，使用差分扩展方法或者差分更改方法将 b 嵌入到差分值的 LSB 位。对于差分扩展，新的差分 $dif'=2dif+b$，对于差分更改，新的差分 $dif'=2\left\lfloor\frac{dif}{2}\right\rfloor+b$。由式(13)，计算新的像素值对 (x',y')

$$x'=ave+\left\lfloor\frac{dif'+1}{2}\right\rfloor,y'=ave-\left\lfloor\frac{dif'}{2}\right\rfloor \tag{14}$$

差分扩展方法对原始像素值的修改幅度大于差分更改方法，但由 (x',y') 恢复为 (x,y) 时不需要额外的辅助信息。对于由差分更改方法得到的 (x',y')，恢复为 (x,y) 时需要知道原始差分值 h 的 LSB。

为处理差分扩展中可能出现的像素值溢出问题，定义满足条件

$$\left|2dif+b\right|\leqslant\min(2(255-ave),2ave+1)$$

的差分 dif 为可扩展的。定义满足条件

$$\left|2\left\lfloor\frac{dif}{2}\right\rfloor+b\right|\leqslant\min(2(255-ave),2ave+1)$$

的差分 dif 为可改变的。显然，可扩展的差分必是可改变的。

将 PB 部分像素值两两分组，每行相邻 2 个位置$(i, 2j)$和$(i, 2j+1)$上的像素值分为一组。假定共有 n 组，计算每组的像素差分，记为一维向量 $\{dif_1,dif_2,\cdots,dif_n\}$。将这些差分按照取值的不同分为 4 个不相交的集合 EZ、EN、CN 和 NC。其中，EZ 包含所有值为 0 或 –1，且可扩展的差分；EN 包含除 EZ 之外的所有可扩展的差分；CN 包含除 EZ 和 EN 之外的所有可改变的差分；NC 包含所有其他的差分。

选择 EZ 中全部差分和 EN 中部分差分使用差分扩展方法，记 EN 中用来差分扩展的部分为 $EN1$，剩余部分为 $EN2$。为了使嵌入前后像素值的均方误差最小，定义

$$\begin{cases}EN1=\left\{dif\in EN\,\middle|\,|dif|\leqslant T\right\}\\ EN2=\left\{dif\in EN\,\middle|\,|dif|>T\right\}\end{cases} \tag{15}$$

T 的取值会在下文给出。

生成长度为 n 的01串 $A=\{a_i\}$，用来标记每个像素组是否使用差分扩展方法，其中，a_i 取0或1，且 $i=1,2,\cdots,n$。并记录 $EN2$ 和 CN 中差分值($dif=1, dif=-2$ 除外)的LSB平面 $M'=\{m'_i\}$，其中，m'_i 取0或1，$i=1,2,\cdots,L'$，L' 为 M' 的长度。将生成的01串 A 和记录的LSB接在一起(中间增加一个分隔标记)，同样使用LZW编码压缩，压缩后记为 M_2。将 M_2 接在 M_1 之后，同样在 M_1 和 M_2 中间增加一个分隔标记，并在尾部接一个结束标记，组成全部需要在PB部分嵌入的数据 $M=M_1\cup M_2=\{m_i\}$，其中，$i=1,2,\cdots L$，L 为 M 的长度。下面给出具体嵌入过程。

ⓐ 令 $i=1, j=1$。

ⓑ 判断 dif_i 所属集合。若 $dif_i\in EZ\cup EN1$，则 $dif'_i=2dif_i+m_j$，根据式(14)计算该组新的像素值，$i=i+1$，$j=j+1$；若 $dif_i\in EN2\cup CN$，则 $dif'_i=2\left\lfloor\frac{dif_i}{2}\right\rfloor+m_j$，根据式(14)计算该组新的像素值，$i=i+1$，$j=j+1$；若 $dif_i\in NC$，$i=i+1$。

ⓒ 重复步骤ⓑ直至 $j=L$。

下面给出可以不记录 $EN2$ 和 CN 中差分值为1或–2的差分LSB的原因。对于 $EN2$ 和 CN 中差分，嵌入前后的差分值要么相同，要么仅在LSB位不同。若嵌入后 CN 中差分 $0\leqslant dif\leqslant 1$，且没有进行差分扩展(标记值为0)，则嵌入前差分一定是1，否则嵌入前此差分是可扩展的，会被标记为1。同理，对于嵌入后 $-2\leqslant dif\leqslant -1$ 且标记为0的差分，嵌入前差分一定是–2。对于这2种情况，不需要记录其LSB就能恢复出原始像素值。

记 $CH=EZ\cup EN\cup CN$，由于嵌入过程全部在 CH 中进行，所以嵌入前后 CH 的范围不会发生改变。相应地给出提取与恢复过程。

ⓐ 找出嵌入后的差分值集合 CH，并按照从前至后的顺序提取每个差分值的LSB，直至遇到结束标记，得到 M，经过LZW解码可得到 E_{PA}、P、A 和 M'。

ⓑ 令 $i=1, j=1$。

ⓒ 若 $dif_j\in CH$，转入步骤ⓓ；若 $dif_i\in NC$，$j=j+1$，转入步骤ⓕ。

ⓓ 若 $a_j=1$，则 $dif'_j=\left\lfloor\frac{dif_j}{2}\right\rfloor$，$j=j+1$；若 $a_j=0$ 且 $0\leqslant dif_j\leqslant 1$，则 $dif'_j=1$，$j=j+1$；若 $a_j=0$ 且 $-2\leqslant dif_j\leqslant -1$，则 $dif'_j=-2$，$j=j+1$；若前3个条件都不满足，则 $dif'_j=2\left\lfloor\frac{dif_j}{2}\right\rfloor+m'_i$，$i=i+1$。

ⓔ 根据式(14)计算原始像素值。

ⓕ 重复步骤ⓒ，直至 $j=n$。

上述嵌入过程后，可将嵌入后的图像当做初始图像进行二次嵌入甚至多次嵌入。发生多次嵌入时，需要在嵌入数据中增加特殊标记，告知此次是第几次嵌入。本文仅使用一次嵌入。

由嵌入过程知，单次嵌入时的最大嵌入容量为 $C_{\mathrm{up}}=|EZ|+|EN|+|CN|$，当 $L\leqslant C_{\mathrm{up}}$ 时才能保证嵌入成功。式(15)中 T 值的选取不会影响 C_{up}，但会影响 L' 和载密图像的质量。T 值越大，集合 $EN1$ 越大，L' 越小，容许 M_1 的长度越大，但嵌入后的图像质量越差。根据数据长度 L 的大小，先取较小的 T，后逐渐增大直到满足 $L\leqslant C_{\mathrm{up}}$。以标准图像大小为512×512的Lena.bmp为例，当取图像中一半的分块作为PA部分，一半的分块作为PB部分时，取 T=2，经计算，此时 L=10 598，$C_{\mathrm{up}}=36\,956$。单次嵌入就能满足需要。

虽然图像整合过程使图像丧失了块与块之间的相关性，但块内像素间的相关性仍然存在，由于每个分块大小都是8×8，应用差分扩展算法的每个分组内的2个像素值不会来自不同的分块，所以不会降低文献[5]方法的性能。但是，由于分块间相关性较弱，并不是每一个可逆隐藏算法都能用到自嵌入过程。选择和使用其他可逆隐藏算法时，要注意避免利用块间的强相关性。记自嵌入后得到的图像为 I''。若知道PA部分的大小，就能在 I'' 中找到PB部分，从而可由图像 I'' 直接准确恢复出原始图像 C。

2) 图像加密

使用流密码将图像 I'' 加密，得到的加密图像记为 E。例如：灰度值 $X_{i,j}$ 可以用8 bit表示，$X_{i,j}(0), X_{i,j}(1),\cdots, X_{i,j}(7)$

$$X_{i,j}(h)=\left\lfloor\frac{X_{i,j}}{2^h}\right\rfloor \bmod 2 \tag{16}$$

其中，$h=0,1,\cdots,7$，使用异或操作计算加密比特 $E_{i,j}(h)=X_{i,j}(h)\oplus r_{i,j}(h)$，其中，$r_{i,j}(h)$ 由一个标准流密码算法生成，加密密钥为 K_1。

最后，将PA部分包含分块的个数嵌入在加密图像的起始位置，告知数据嵌入者可用的嵌入范

围。具体地，把加密图像上 PA 部分第一个分块的 LSB 平面的首 16 bit 直接替换为 PA 部分大小。由于第一个分块经过重复压缩就能去除嵌入影响，所以无需保留嵌入位置的原始 LSB 平面。图像加密后，数据嵌入者或者第三方在没有加密密钥的情况下将不能访问原始图像的内容，因此，图像内容得到了保护。

3) 加密图像的数据嵌入

得到加密图像 E 后，数据嵌入者不需要获得原始图像就可以在 E 上进行消息嵌入。从 E 的第一个分块提取数据，可以获知 PA 部分具体大小。数据嵌入者使用 k $(k=1,2)$层 LSB 替换的方法将消息嵌入在除去第一个分块的PA 部分，隐写密钥为 K_2，保证每个分块有 l 个像素值承载 kl bit 数据。设嵌入数据后的加密图像为 E'。如此，图像管理者就能通过在加密图像上嵌入一些符号来标记和管理这些图像，包含图像拥有者、管理者的身份、时间戳等信息。为了提取方便，可以在信息尾部接上一个固定的结束标记。没有隐写密钥 K_2 的第三方将不能提取嵌入的数据。

4) 数据提取与图像恢复

由于数据提取和图像解密是完全独立的，按照数据提取的不同顺序可实现以下 2 种不同的应用。

应用 1　从加密图像中提取数据

① 数据提取

图像管理者为了管理加密图像和更新加密图像中的嵌入数据，需要对嵌入数据进行提取。图像所有者和使用者得到加密图像 E' 后，如果要追踪数据来源，也需要对嵌入数据进行提取。提取时，使用密钥 K_2 在图像 E' 的 PA 部分提取 k 层 LSB 平面，直至遇到结束标记。图像的数据进行更新时，数据库管理员只需使用相同密钥对新数据进行再次嵌入。整个过程都在加密图像上进行操作的，避免了图像内容的泄露。

② 图像恢复

除了提取更新嵌入数据，图像接收者还可以恢复出原始图像。下面给出具体过程。

ⓐ 图像接收者提取图像第一个分块的首16 bit 数据，确定 PA 部分和 PB 部分的准确位置。

ⓑ 使用加密密钥对图像进行解密。例如需要解密加密像素值 $E'_{i,j}$，计算 $X'_{i,j}(h)=E'_{i,j}(h)\oplus r_{i,j}(h)$，$X'_{i,j}=\sum_{h=0}^{7}X'_{i,j}(h)2^h$，其中，$E'_{i,j}(h)$ 是 $E'_{i,j}$ 通过式(16)得到的二元比特，$r_{i,j}(h)$ 同样由 K_1 控制生成，$X'_{i,j}$ 即为解密后的像素值。

ⓒ 利用自嵌入算法中的提取和恢复算法在PB 部分提取出 $E_{\mathrm{PA}}=\{e_{i,j}\}$ 和 P，并将 PB 部分恢复至自嵌入前的状态。

ⓓ 将 PA 部分中每个分块进行重复压缩，对重复压缩后的 PA 部分中的每个像素值 $X''_{i,j}$，计算 $C'_{i,j}=X''_{i,j}+e_{i,j}$，并将 $C'_{i,j}$ 作为恢复后的像素值。

ⓔ 根据 P，重新调整分块顺序，得到恢复出的原始图像。

应用 2　从解密图像中提取数据

用户有时会想先解密图像然后从解密图像提取消息。下面给出一个具体的应用场景：假设 Alice 将图像上传到云服务器，为了保护图像内容，图像已经进行了加密，云服务器需要通过在加密图像上嵌入一些符号来标记和管理这些图像，包含图像拥有者、云服务器的身份、时间戳等信息。授权用户 Bob 拥有加密密钥和消息嵌入密钥，可以下载并解密这些图像，Bob 还希望能够获得带嵌入信息的解密图像，即解密后图像中包含可用于追踪数据来源的信息。先图像解密后数据提取完全适合于这种情况。

接下来，描述如何生成一幅含标记的解密图像。

①生成含标记的解密图像

按照下述过程，图像拥有者可以获得含标记的解密图像。

ⓐ 类似于应用 1，图像拥有者使用加密密钥对图像进行解密，但为了保留标记，解密时跳过 PA 部分的 k 层 LSB 平面。可以计算出包含嵌入数据的部分解密图像，记此时得到的图像为 I'''。

ⓑ 同样提取图像第一个分块的首 16 bit 数据，确定 PA 部分和 PB 部分的准确位置。利用自嵌入算法中的提取算法在 PB 部分提取出 P，据此重新排列图像中所有分块，将 PA 部分和 PB 部分都恢复到原始位置，此时就可以获得包含嵌入数据的解密图像，即含标记的解密图像，记此时图像为 C'。

C' 与原始图像 C 相比，保持着良好的视觉不可见性。具体来讲，失真来源于 3 个方面：一是对 PA 部分的压缩与解压缩，二是数据嵌入过程对 PA 的 k 层 LSB 平面的修改，三是自嵌入过程对 PB 部分的修改。第一部分失真虽不可避免，但使用较大的质量因子可减少该部分失真，第二部分失真可以通过调整参数 k 和 l 来控制；第三部分可以使用合适的最新的可逆数据隐藏技术来获得良好的性能。

②数据提取和图像恢复

在生成含嵌入数据的解密图像C'后，图像拥有者可以进一步提取消息和恢复原始图像。该过程与传统的可逆数据隐藏方法类似，下面给出具体过程。

ⓐ 根据P，将C'调整为I'''。I'''与I''仅在PA部分的k层LSB平面上有部分差异。

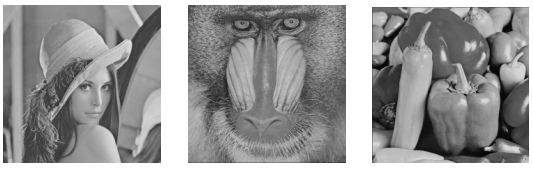
ⓑ 根据K_2，在I'''的PA部分提取出嵌入数据。

ⓒ 此时可重复应用1中图像恢复的实验©~ⓔ，得到恢复出的原始图像C''。

由于可恢复块的可恢复概率不全为1，对I'''中PA部分的分块重复压缩后，可能不能完全去除嵌入过程造成的影响，导致最终恢复的图像与原始图像存在细微差异，但这些误差在允许范围内。算法中选择可恢复概率较高的可恢复块组成PA部分可使这种细微差异最小化。

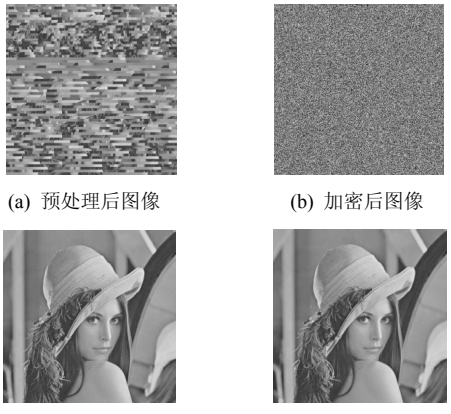
5 实验与分析

为了测试本文算法的性能，选择大小为512×512的3幅标准灰度BMP图像[22]进行测试，如图3所示，从左至右依次为Lena.bmp、Baboon.bmp和Peppers.bmp。以Lena.bmp图像为例，设置参数$q=75,k=2,l=64$，图4展示了此时算法过程中的4幅关键图像。观察可知，图4(c)和图4(d)与原始图像相比，并无明显的视觉差异。



(a) Lena.bmp (b) Baboon.bmp (c) Peppers.bmp

图3 3幅典型灰度图像



(a) 预处理后图像 (b) 加密后图像

(c) 含标记信息的解密图像 (d) 恢复后图像

图4 实验过程中的4幅图像

使用错误率度量本文算法恢复图像的性能，使用PSNR度量含标记的解密图像的质量。下面分别从这2个方面进行实验与分析。

1) 错误率

错误率为恢复错误的像素值个数与图像大小的比值。实验中，令PA部分与PB部分大小相同，选择不同的q、k、l，计算图3所示的3幅图像的错误率均值，实验结果如表1所示。

由表1可知，l较小时($l\leqslant 16$)，单层嵌入和双层嵌入都能实现零错误；l较大时($l\geqslant 24$)，单层嵌入拥有比双层嵌入更小的错误率。由此可确定预处理过程中参数m的选择。优先选择较小的k(k=1)和l($l\leqslant 16$)，然后根据所需的嵌入容量计算m的取值。因为PB部分需要承担自嵌入任务，m不能太大。若m大于块个数的一半，可取k=2并增大l，重新计算m。

由定理4，可恢复载密块通过重复压缩实现去除嵌入影响的成功率负相关于q、k、l。由表1可以看出，相同条件下，q、k和l越大，错误率越高。恰好验证了定理4的正确性。为了限制错误率，本文实验中限制$l\leqslant l_{\max}$，其中，$l_{\max}=32$。

表1 不同参数下的错误率

q	k=1				k=2			
	l=8	l=16	l=24	l=32	l=8	l=16	l=24	l=32
65	0	0	0	0	0	0	0	0.001 0
75	0	0	0	0.001 0	0	0	0.002 4	0.005 2
85	0	0	0.001 8	0.004 6	0	0	0.057 8	0.113 0

将本文算法与文献[12~14]方法进行比较。本文方法中质量因子取75，逐渐增加l的大小来控制嵌入率的变化。对于文献[12,13]方法，逐渐增加分块的大小控制嵌入率的变化。对于文献[14]方法，令分组长度为200，取2层LSB平面，逐渐增加s的大小来控制嵌入率的变化。计算不同嵌入率下恢复原始载体过程出现的错误率均值，实验结果如图5所示。

从图5可以看出，本文方法在相同嵌入率下有更小的错误率，并且单层嵌入的错误率明显低于双层嵌入。尤其选择单层嵌入时，错误率几乎为0。因为双层嵌入对可恢复块上像素值的修改幅度更大，嵌入成功的概率会更小。而且本文方法在嵌入率较大的情况下仍能保持很小的错误率，当嵌入率达到0.5 bpp时，错误率仍然在0.06以内。而文献[12~14]方法，仅能在嵌入率非常小的情况下保持低错误率。

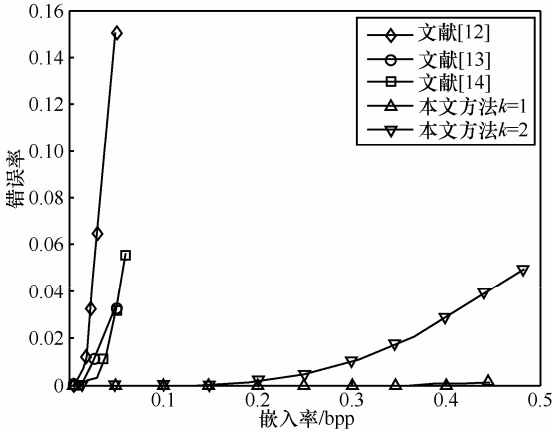


图 5　本文方法与文献[12～14]方法的错误率

2）PSNR

同样使用错误率实验中控制嵌入率变化的方法，计算文献[12~14]方法中直接解密图像和本文方法中含标记的解密图像的 *PSNR*，用以衡量不同方法生成标记图像的质量。本文实验中，分别计算图 3 所示 3 幅图像在不同质量因子($q=65,75,85$)下含标记解密图像的 *PSNR*。实验结果如图 6~图 14 所示。

从图 6~图 14 看出，本文方法在相同嵌入率下有更高的 *PSNR*，并且在可接受的 *PSNR* 范围内，增加了嵌入容量。双层嵌入的 *PSNR* 要高于单层嵌入，因为本文算法的主要失真在于 PA 部分差值的可逆嵌入，相同条件下，双层嵌入的嵌入率是单层嵌入率的 2 倍，但双层嵌入对 *PSNR* 的影响并不是单层嵌入的 2 倍。对同一幅图像，质量因子越高，*PSNR* 越高。因为质量因子越高，JPEG 压缩过程引入的截断误差和取整误差越少，使压缩后的 PA 部分更接近原始状态，并且有更小的 E_{PA}，自嵌入过程中对 PB 部分的修改也越小，造成的失真越小。

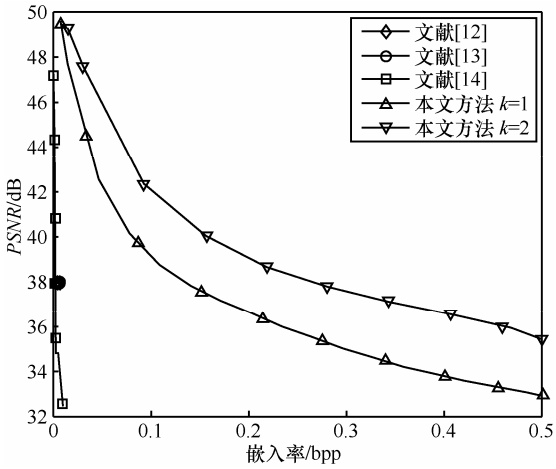


图 6　Baboon.bmp 在 q=65 时的实验结果

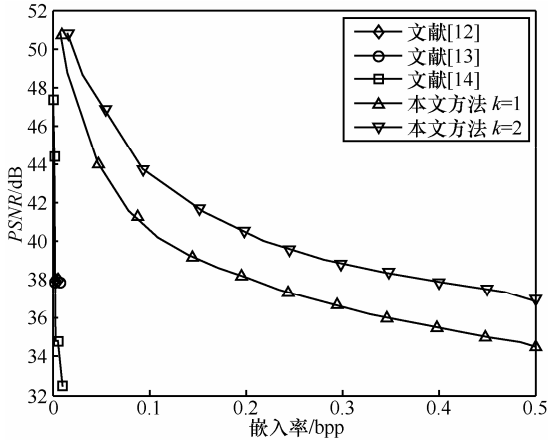


图 7　Baboon.bmp 在 q=75 时的实验结果

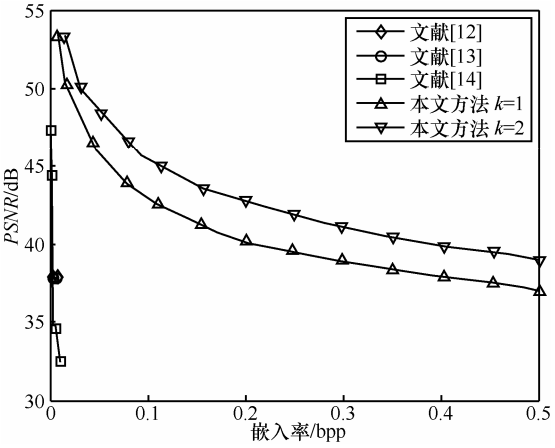


图 8　Baboon.bmp 在 q=85 时的实验结果

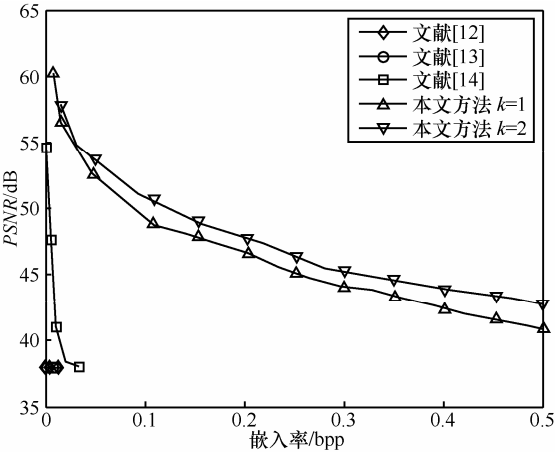


图 9　Lena.bmp 在 q=65 时的实验结果

本文实验中，将 PA 部分和 PB 部分设为大小相同。实际上，PA 部分越小，其能通过重复压缩成功恢复的概率越高，同时需要自嵌入的数据越少，带标记解密图像的 *PSNR* 也会越高，但是会降低最大的嵌入容量。所以，对嵌入容量需求小的时候，可适当将 PA 部分缩小以获取更好的性能。

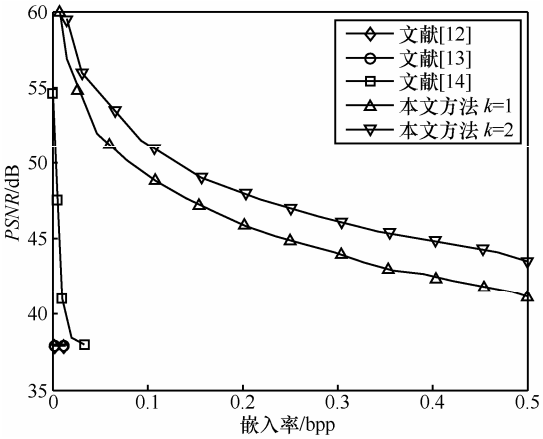


图 10　Lena.bmp 在 q=75 时的实验结果

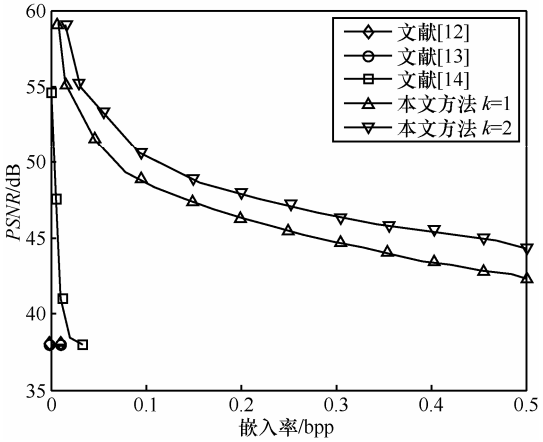


图 11　Lena.bmp 在 q=85 时的实验结果

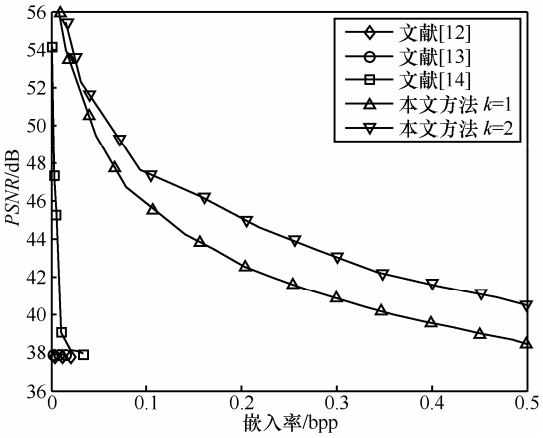


图 12　Peppers.bmp 在 q=65 时的实验结果

由以上分析，相同条件下，双层嵌入比单层嵌入有更大的嵌入率，生成的带标记解密图像的 *PSNR* 更高，但恢复错误率也更高；l 越大，嵌入容量越大，*PSNR* 越低，恢复错误率越高，但最好保证 $l \leqslant l_{\max}$；q 越大，*PSNR* 越高但恢复错误率越低，一般情况下，$q=75$ 较为适宜。实际应用中可根据不同的应用需求选择合适的参数。

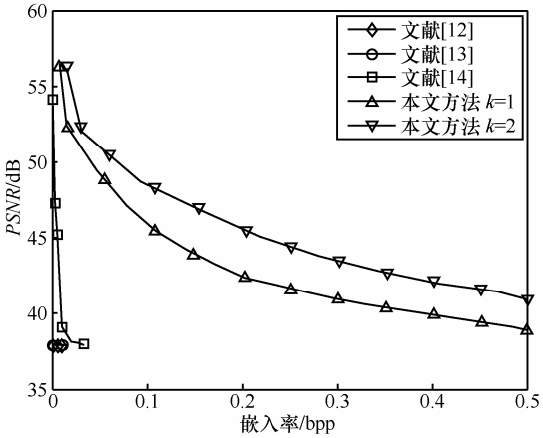


图 13　Peppers.bmp 在 q=75 时的实验结果

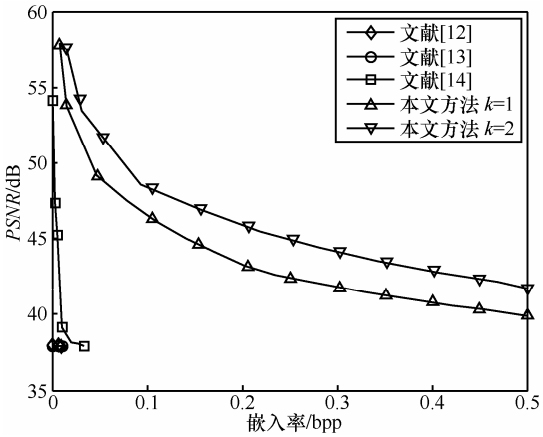


图 14　Peppers.bmp 在 q=85 时的实验结果

6　结束语

由于云数据处理中对隐私保护的需求，加密图像上的可逆数据隐藏受到越来越多的关注。本文提出了一种基于重复压缩的密文图像上的可逆数据隐藏算法。首先给出了算法可行的理论基础，讨论了 JPEG 解压缩载密块经过重复压缩能恢复原始载体的充分条件和理论概率，并利用重复压缩实现密文图像上的可逆数据隐藏算法。本文算法沿用文献[17]提出的模型，在加密之前预留出秘密消息承载空间，实现了数据提取与图像解密的完全分离，并在以下 2 个方面取得良好性能：1)在嵌入率一定的情况下，提高了含标记解密图像的 *PSNR*，同时对于可接受的 *PSNR*，提高了嵌入容量；2)能以高概率准确恢复出原始图像，降低了图像恢复的错误率。本文使用了经典的文献[5]算法来描述自嵌入过程，若使用性能更优的可逆隐藏算法，本文算法性能会更好。

参考文献：

[1] KALKER T, WILLEMS F M J. Capacity bounds and constructions for reversible data hiding[A]. Proceedings of the 14th International Conference on Digital Signal Processing[C]. Santorini, Greece, 2002. 71-76.

[2] ZHANG W, CHEN B, YU N. Capacity-approaching codes for reversible data hiding[A]. Proc of the Information Hiding[C]. Springer Berlin Heidelberg, 2011. 255-269.

[3] ZHANG W, CHEN B, YU N. Improving various reversible data hiding schemes via optimal codes for binary covers[J]. IEEE Transactions on Image Processing, 2012, 21(6): 2991-3003.

[4] FRIDRICH J, GOLJAN M, DU R. Lossless data embedding for all image formats[A]. Proc of the Electronic Imaging, International Society for Optics and Photonics[C]. 2002. 572-583.

[5] TIAN J. Reversible data embedding using a difference expansion[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(8): 890-896.

[6] NI Z, SHI Y Q, ANSARI N, *et al*. Reversible data hiding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3): 354-362.

[7] LI X, YANG B, ZENG T. Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection[J]. IEEE Transactions on Image Processing, 2011, 20(12): 3524-3533.

[8] TSAI P, HU Y C, YEH H L. Reversible image hiding scheme using predictive coding and histogram shifting[J]. Signal Processing, 2009, 89(6): 1129-1143.

[9] LUO L, CHEN Z, CHEN M, *et al*. Reversible image watermarking using interpolation technique[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(1): 187-193.

[10] MENEZES A J, VAN OORSCHOT P C, VANSTONE S A. Handbook of Applied Cryptography[M]. Boca Raton, FL, USA: CRC, 1996.

[11] HWANG K, LI D. Trusted cloud computing with secure resources and data coloring[J]. Internet Computing, IEEE, 2010, 14(5): 14-22.

[12] ZHANG X. Reversible data hiding in encrypted image[J]. Signal Processing Letters, IEEE, 2011, 18(4): 255-258.

[13] HONG W, CHEN T S, WU H Y. An improved reversible data hiding in encrypted images using side match[J]. Signal Processing Letters, IEEE, 2012, 19(4): 199-202.

[14] ZHANG X. Separable reversible data hiding in encrypted image[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2): 826-832.

[15] JOHNSON M, ISHWAR P, PRABHAKARAN V, *et al.* On compressing encrypted data[J]. IEEE Transactions on Signal Processing, 2004, 52(10): 2992-3006.

[16] LIU W, ZENG W, DONG L, *et al.* Efficient compression of encrypted grayscale images[J]. IEEE Transactions on Image Processing, 2010, 19(4): 1097-1102.

[17] MA K, ZHANG W, ZHAO X. Reversible data hiding in encrypted images by reserving room before encryption[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(3): 553-562.

[18] NRCS Photo Gallery[EB/OL]. http://photogallery.nrcs. usda.gov/, 2011.

[19] LUO W, HUANG J, QIU G. JPEG error analysis and its applications to digital image forensics[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(3): 480-491.

[20] LAM E Y, GOODMAN J W. A mathematical analysis of the DCT coefficient distributions for images[J]. IEEE Transactions on Image Processing, 2000, 9(10): 1661-1666.

[21] FRIDRICH J, GOLJAN M, SOUKAL D. Perturbed quantization steganography[J]. Multimedia Systems, 2005, 11(2): 98-107.

[22] Miscelaneous gray level images [EB/OL]. http://decsai.ugr.es/cvg/dbimagenes/g512.php.

作者简介：

刘九芬（1963-），女，河南温县人，博士，信息工程大学教授，主要研究方向为密码学和信息隐藏。

韩涛（1986-），男，四川彭州人，信息工程大学博士生，主要研究方向为密码学和信息隐藏。

田雨果（1986-），男，四川雅安人，信息工程大学硕士生，主要研究方向为密码学和信息隐藏。

刘文彬（1989-），男，河南许昌人，信息工程大学硕士生，主要研究方向为信息隐藏。