软件学报 ISSN 1000-9825, CODEN RUXUEW
*Journal of Software*,2018,29(4):957−972 [doi: 10.13328/j.cnki.jos.005406]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

# 基于同态加密系统的图像鲁棒可逆水印算法*

项世军[1,2], 杨 乐[1,2]

[1](暨南大学 信息科学技术学院 电子工程系,广东 广州 510632)

[2](信息安全国家重点实验室(中国科学院 信息工程研究所),北京 100093)

通讯作者: 项世军, E-mail: Shijun_Xiang@qq.com

**摘 要**: 同态加密技术可用于保护数据隐私并允许对密文数据进行算术操作,在云计算安全上有着很好的应用前景.针对云计算中的隐私保护和数据安全等问题,提出了一种基于同态加密系统的图像鲁棒可逆水印算法,主要思想为:(1) 对原始图像进行分块和利用 Paillier 加密系统进行加密得到密文图像;(2) 在加密域中,通过模乘法逆元 MMI(modular multiple inverse)方法和查询相应的密文映射表得到每个密文分块的统计量,然后利用同态特性对统计量进行直方图平移来嵌入水印信息;(3) 在接收方,可从含水印的密文图像的统计量直方图中完整地提取水印,并可通过对统计量进行与嵌入过程相反的直方图平移操作来恢复原始密文图像;(4) 含水印的密文图像在直接解密后可从其统计量直方图中完整地提取水印信息和恢复原始图像;(5) 解密后的含水印图像在受到一定程度的攻击后(如 JPEG/JPEG 2000 压缩和叠加高斯噪声等),水印仍能正确提取.该算法实现了在不对原始图像进行预处理的情况下可直接在加密后的密文图像中嵌入水印,并可分别在加密域或明文域提取水印和恢复原始密文图像或原始明文图像,而且嵌入的水印对常见的图像处理操作具有一定的鲁棒性.实验仿真结果验证了该算法的有效性.

**关键词**: 同态加密;鲁棒可逆水印;隐私保护;数据安全;云计算

**中图法分类号**: TP309

中文引用格式: 项世军,杨乐.基于同态加密系统的图像鲁棒可逆水印算法.软件学报,2018,29(4):957−972. http://www.jos.org.cn/1000-9825/5406.htm

英文引用格式: Xiang SJ, Yang L. Robust and reversible image watermarking algorithm in homomorphic encrypted domain. Ruan Jian Xue Bao/Journal of Software, 2018,29(4):957−972 (in Chinese). http://www.jos.org.cn/1000-9825/5406.htm

## Robust and Reversible Image Watermarking Algorithm in Homomorphic Encrypted Domain

XIANG Shi-Jun[1,2], YANG Le[1,2]

[1](Department of Electronic Engineering, School of Information Science and Technology, Ji'nan University, Guangzhou 510632, China)

[2](State Key Laboratory of Information Security (Institute of Information Engineering, The Chinese Academy of Sciences), Beijing 100093, China)

**Abstract**: Homomorphic encryption technique can be used for protection of data privacy, and some algebraic operations can be implemented on the ciphertext data. This is very useful in the field of cloud computing security, such as analyzing and processing the encrypted data in cloud without exposing the content of data. Addressing privacy protection and data security problems in cloud computing, this paper proposes a robust and reversible image watermarking algorithm in homomorphic encrypted domain. The algorithm includes five aspects: (1) The original image is divided into a number of non-overlapping blocks and each pixel in a block is encrypted

* 基金项目: 国家自然科学基金(61772234, 61272414); 信息安全国家重点实验室开放课题基金(2016-MS-07)

Foundation item: National Natural Science Foundation of China (61772234, 61272414); Open Research Fund from the State Key Laboratory of Information Security (2016-MS-07)

本文由"多媒体大数据处理与分析"专题特约编辑赵耀教授、李波教授、华先胜研究员、文继荣教授、蒋刚毅教授、常冬霞副教授推荐.

收稿时间: 2017-04-29; 修改时间: 2017-06-26; 采用时间: 2017-10-13; jos 在线出版时间: 2017-12-01

CNKI 网络优先出版: 2017-12-04 06:46:53, http://kns.cnki.net/kcms/detail/11.2560.TP.20171204.0646.007.html

with Paillier cryptosystem to obtain the encrypted image; (2) The statistical values of the encrypted blocks can be retrieved in encrypted domain by employing modular multiplicative inverse (MMI) method and looking for a mapping table. After that, watermark information can be reversibly embedded into encrypted image by shifting the histogram of the statistical values with the homomorphic property of Paillier cryptosystem; (3) On the receiver side, the marked histogram of the watermarked and encrypted image can be obtained for extraction of the watermark from the marked histogram. The encrypted image can be restored by inverse operations of histogram shifting in the embedding phase; (4) The marked histogram can be obtained from the directly decrypted image. This is followed by the watermark extraction and restoration of original image; (5) The watermark can still be extracted correctly under some attacks (such as JPEG/JPEG2000 compression and additive Gaussian noise) to some extent on the watermarked and decrypted image. The proposed method achieves embedding information bits directly into the encrypted image without preprocessing operations on the original image, and can extract the watermark and restore the encrypted image in encrypted domain or the original image in plaintext domain after decryption. Besides, the watermark is robust to those common image processing operations. The experimental results have shown the validity of the proposed scheme.

**Key words**: homomorphic encryption; robust and reversible watermarking; privacy protection; data security; cloud computing

可逆数字水印技术通过利用多媒体信息中存在的冗余,将水印信息(如载体特征信息、版权信息等)嵌入到数字多媒体载体当中,并可在接收方完整地提取水印且无损地恢复原始载体[1].该技术可用于媒体的内容标识、完整性认证和版权保护等功能,已广泛应用于对保密性、安全性以及保真度要求较高的领域,如军事及医学图像、法律文书图片等[2].基于图像的可逆水印算法主要可以分为以下四大类:无损压缩[3]、差值扩展[4]、直方图平移[5]以及预测误差扩展[6].可逆数字水印技术一般不考虑水印的鲁棒性,现有的大多数可逆水印算法在受到噪声干扰或经过处理后无法正确地提取嵌入的水印.实际中,含水印的数字多媒体在网络等信道中传输时,不可避免地会遭到各种处理或干扰,所以,在许多应用场景中希望嵌入的水印具有一定的鲁棒性,在含水印的载体未受到信号处理或恶意攻击时,在接收方能够正确地提取水印和无损地恢复原始载体;而在含水印的载体经过一些信号处理操作时,在接收端仍然能够正确地提取水印.近几年来,鲁棒可逆水印技术逐渐成为了信息隐藏领域的一个重要的研究方向,已取得一些有意义的成果[7−13].

随着互联网技术和云计算技术的快速发展,用户可通过互联网将资料和数据上传到远程服务器或云端进行存储,当需要时再下载使用.云存储节省了购买储存设备的开支并提高了获取资源的便利性.另外,云服务中心的强大计算能力使得用户可以享受到第三方提供的数据处理等服务.然而,云计算技术在方便人们生活的同时,也引发了数据安全和隐私保护的问题[14].数据加密作为保护多媒体内容隐私的一种方式,可以在上传到云端之前先对数据进行加密,有效地降低内容隐私泄露的风险.为了对云端中海量的加密数据进行有效管理和安全保护,云端管理者希望将一些用户资料相关的信息嵌入到密文数据中,并通过提取嵌入的信息来实现密文检索和数据保护.因此,加密域可逆水印技术成为了近年来大数据云计算背景下信息隐藏领域的研究热点.加密域可逆水印技术结合加密技术和可逆水印技术的优点,在不暴露明文内容的情况下直接将水印嵌入到密文载体中,并在解密及信息提取后能无损地恢复原始载体.目前的加密域图像可逆水印算法主要分为两大类:基于对称加密系统的可逆水印算法[15,16]和基于非对称加密系统的可逆水印算法[17−20].目前,加密域中的可逆水印算法绝大多数都不具有鲁棒性,而在一些云计算的应用场景中,信息隐藏者希望嵌入的水印信息不仅具有可逆性,而且具有一定的鲁棒性,以便在解密后图像受到处理或攻击时仍能认证媒体的版权.

与对称加密系统相比,同态加密系统为非对称加密系统,安全性更高,且允许对密文进行算术运算,更适合用于云计算背景下的第三方数据处理.本文结合同态加密系统和鲁棒可逆水印技术,提出了一种基于同态加密系统的图像鲁棒可逆水印算法.该算法首先对原始图像进行 8×8 分块,并利用 Paillier 加密系统[21]进行加密得到密文图像.在加密域中,通过模乘法逆元 MMI(modular multiple inverse)方法和查询密文映射表来得到每个密文分块的统计量,然后利用同态特性对统计量进行直方图平移嵌入水印.在接收方可分别在含水印的密文图像或解密图像中得到统计量直方图并提取水印,同时,通过对统计量进行与嵌入过程相反的直方图平移来恢复密文图像或原始图像.另外,解密后含水印的图像在受到一定程度的图像处理操作(如 JPEG/JPEG 2000 压缩和叠加高斯噪声等)后仍能正确地提取水印.该算法实现了在不对原始图像进行预处理的情况下可直接在加密后的密

文图像中嵌入水印,并可分别从加密域或明文域提取水印和恢复密文图像或原始图像,而且嵌入的水印对常见的图像处理操作具有一定的鲁棒性.本文算法嵌入失真较小,鲁棒性良好,具有足够的嵌入容量来嵌入加密图像相关标签信息、版权信息或图像特征信息等,适用于云计算中加密图像的内容标识、完整性认证及版权保护.

## 1 Paillier 加密系统

Paillier 加密系统[21]是一种加性同态公钥加密系统,这种加密技术已广泛应用于加密信号处理或第三方数据处理领域.其同态特性表现为:在加密后可直接对密文进行相应的算术运算,其运算结果与明文域中对应的运算结果一致.其概率特性表现为:对于相同的明文,可通过不同的加密过程得到不同的密文,从而保证了密文的语义安全.其加密和解密机制如下.

**密钥生成**:随机选择两个较大的质数 $p$ 和 $q$,计算它们的乘积 $N$ 以及 $p-1$、$q-1$ 的最小公倍数$\lambda$.然后再随机选取一个整数 $g\in Z_{N^2}^*$, 且 $g$ 满足:

$$gcd\left(L\left(g^{\lambda} \bmod N^2\right),N\right)=1 \tag{1}$$

其中,函数 $L(u)=(u-1)/N$,函数 $gcd(\cdot)$用于计算两数的最大公约数, $Z_{N^2}$ 为小于 $N^2$ 的整数的集合,而 $Z_{N^2}^*$ 为 $Z_{N^3}$ 中与 $N^2$ 互质的整数的集合.$(N,g)$和$\lambda$分别为公钥和私钥.

**加密过程**:随机选取一个整数 $r\in Z_N^*$, 对于任意一个明文 $m\in Z_N$, 利用公钥$(N,g)$加密后得到对应密文 $c$ 为

$$c=E[m,r]=g^m\cdot r^N \bmod N^2 \tag{2}$$

根据 Paillier 加密系统的性质,密文 $c\in Z_{N^2}^*$.利用相同的公钥进行加密时,由于 $r$ 的选取是随机的,对于同一个明文 $m$,可得到不同的密文 $c$.但是解密后可以还原出相同的明文 $m$,从而保证了密文的语义安全.

**解密过程**:利用私钥$\lambda$,对密文 $c$ 解密后得到对应的明文 $m$:

$$m=D[c]=\frac{L\left(c^{\lambda} \bmod N^2\right)}{L\left(g^{\lambda} \bmod N^2\right)} \bmod N \tag{3}$$

另外,Paillier 加密系统具有两个重要的性质.

**定理 1**. 当 $g$ 满足公式(1)时,则 $c=E[m,r]$是双射的,即 $\forall(m,r)\,|\,m\in Z_N,r\in Z_N^*$ 都有唯一的 $c=E[m,r]$ 与之一一对应.也就是说,对于两个明文 $m_1,m_2\in Z_N$ 和 $\forall r_1,r_2\in Z_N^*$, 根据公式(2)分别得到对应的密文 $c_1,c_2\in Z_{N^2}^*$, 则当且仅当 $m_1=m_2$ 和 $r_1=r_2$ 时,等式 $c_1=c_2$ 成立.

本文将利用该定理实现查询密文映射表得到密文分块统计量.

**同态乘法性质**:对于两个明文 $m_1,m_2\in Z_N$ 和 $\forall r_1,r_2\in Z_N^*$, 对应密文 $c_1=E[m_1,r_1],c_2=E[m_2,r_2]$ 满足:

$$c_1\cdot c_2=E[m_1,r_1]\cdot E[m_2,r_2]=g^{m_1+m_2}\cdot\left(r_1\cdot r_2\right)^N \bmod N^2 \tag{4}$$

解密后得到:

$$D[c_1\cdot c_2]=D\left[E[m_1,r_1]\cdot E[m_2,r_2] \bmod N^2\right]=m_1+m_2 \bmod N \tag{5}$$

本文将利用该性质对密文分块统计量进行直方图平移,实现在加密域中嵌入水印信息.

## 2 同态加密域图像鲁棒可逆水印算法

本文提出的基于同态加密系统的图像鲁棒可逆水印算法框架如下文图 1 所示.首先,数据所有者对原始图像进行 8×8 分块,并利用公钥$(N,g)$、密钥 $K_s$ 和 Paillier 加密系统进行加密得到密文图像.信息隐藏者在加密域中利用 $K_s$、模乘法逆元 MMI 方法和查询相应密文映射表得到每个密文分块的统计量,然后利用嵌入密钥$(T,G)$对统计量进行直方图平移嵌入水印.在接收方利用与嵌入过程相同的方法,可在加密域中得到含水印的加密图像的统计量直方图,然后利用$(T,G)$提取水印并通过与嵌入过程相反的直方图平移恢复密文图像.在接收方利用私钥$\lambda$可对含水印的加密图像进行解密并得到统计量直方图,然后利用$(T,G)$可从未受到攻击或图像处理的图像

中提取水印和恢复原始图像,或者从受到一定程度攻击的图像中提取水印.

### 2.1 图像分块加密

数据所有者首先把原始图像 $I$ 分为若干个互不重叠的大小为 8×8 的明文分块,记第 $k$ 个明文分块为 $P^{(k)}$.然后按照第 1 节 Paillier 加密系统中描述的加密过程,随机选取一个整数 $r_1(k)\in Z_N^*$,对 $P^{(k)}$中的每个明文像素值 $P^{(k)}(i,j)$利用公钥$(N,g)$和 $r_1(k)$进行加密得到密文 $C^{(k)}(i,j)$:

$$C^{(k)}(i,j)=E[P^{(k)}(i,j),r_1(k)]=g^{P^{(k)}(i,j)}\cdot r_1(k)^N \bmod N^2 \tag{6}$$

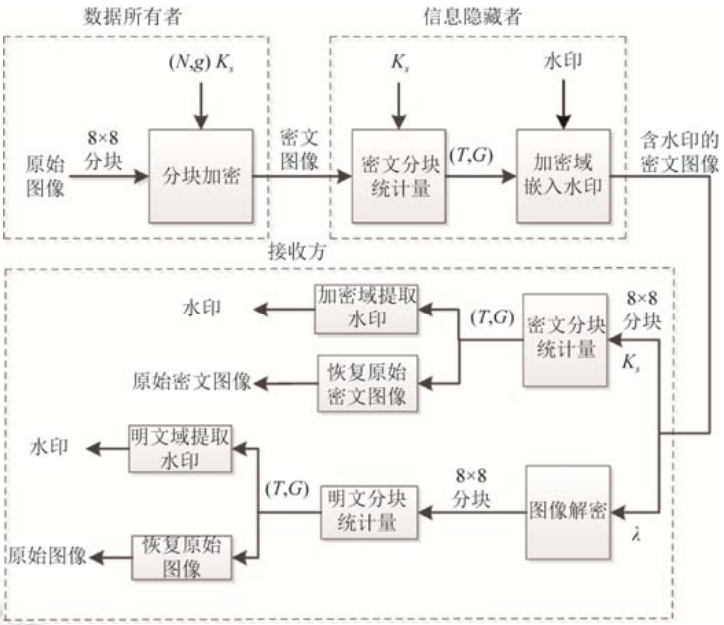



Fig.1 Sketch of the proposed robust and reversible image watermarking algorithm in encrypted image based on homomorphic cryptosystem

图 1 基于同态加密系统的图像鲁棒可逆水印算法框架

其中, $i\in[1,8]$, $j\in[1,8]$, 记 $P^{(k)}$加密后的密文分块为 $C^{(k)}$.为了提高下文提出的密文分块统计量的安全性,数据所有者根据密钥 $K_s$ 随机选取另外 8×8 个整数 $r_{(i,j)}(k)\in Z_N^*$ 对 0 进行同态加密,记加密过程为 $E[0,r_{(i,j)}(k)]$, 并对 $C^{(k)}$中的密文 $C^{(k)}(i,j)$ 进行同态乘法:

$$C'^{(k)}(i,j)=C^{(k)}(i,j)\cdot E[0,r_{(i,j)}(k)]=g^{P^{(k)}(i,j)}\cdot(r_1(k)\cdot r_{(i,j)}(k))^N \bmod N^2 \tag{7}$$

记改变后的密文为 $C'^{(k)}(i,j)$, 密文分块为 $C'^{(k)}$. 记 $D[C'^{(k)}(i,j)]$ 为 $C'^{(k)}(i,j)$ 解密后的结果,根据式(4)和式(5), $D[C'^{(k)}(i,j)]$满足:

$$D[C'^{(k)}(i,j)]=P^{(k)}(i,j) \bmod N \tag{8}$$

因此, $C'^{(k)}$ 中的密文 $C'^{(k)}(i,j)$ 仍然是明文像素值 $P^{(k)}(i,j)$ 的一种加密结果, $C'^{(k)}(i,j)$ 分别由不同的 $r_1(k)\cdot r_{(i,j)}(k)$ 加密而成.在没有密钥 $K_s$ 的情况下,将不能得到下文提出的密文分块统计量.记加密后的图像为 $E[I]$.

### 2.2 信息嵌入

#### 2.2.1 模乘法逆元

Zheng[22]提出了模乘法逆元 MMI 方法.对于两个互质的整数 $y$ 和 $z$,存在整数 $\theta$ 满足:

$$\theta\cdot y=1 \bmod z \tag{9}$$

称$\theta$为$y$的模乘法逆元,$\theta$可根据扩展欧几里德算法[23]求得.利用MMI方法可以实现模运算中的除法.设另一

个整数 $x$,$v$ 是对 $x$ 和 $y$ 的乘积进行模 $z$ 运算的结果:

$$v = x \cdot y \bmod z \tag{10}$$

当 $y$ 已知时,利用 MMI 方法可从 $v$ 中求出 $x$:

$$v \cdot \theta = x \cdot y \cdot \theta = x \bmod z \tag{11}$$

信息隐藏者接收到加密图像 $E[I]$后,首先把 $E[I]$分为若干个互不重叠的大小为 8×8 的密文分块,则第 $k$ 个密文分块为 $C'^{(k)}$. 然后利用密钥 $K_s$ 和 MMI 方法将 $C'^{(k)}$ 中的密文 $C'^{(k)}(i,j)$ 恢复为 $C^{(k)}(i,j)$. 具体方法如下所述:由第 1 节可知密文 $c \in Z_{N^2}^*$, 即 $E[0,r_{(i,j)}(k)]$,$C^{(k)}(i,j)$ 都与 $N^2$ 互质,则可根据扩展欧几里德算法求得 $E[0,r_{(i,j)}(k)]$ 的模乘法逆元 $\theta_{Er}(k)$,$\theta_{Er}(k)$ 满足:

$$\theta_{Er}(k) \cdot E[0,r_{(i,j)}(k)] = 1 \bmod N^2 \tag{12}$$

由于 $C'^{(k)}(i,j)$ 是 $C^{(k)}(i,j)$ 和 $E[0,r_{(i,j)}(k)]$ 的乘积,可通过 $C'^{(k)}(i,j)$ 和 $\theta_{Er}(k)$ 求得 $C^{(k)}(i,j)$:

$$C'^{(k)}(i,j) \cdot \theta_{Er}(k) = C^{(k)}(i,j) \cdot E[0,r_{(i,j)}(k)] \cdot \theta_{Er}(k) = C^{(k)}(i,j) \bmod N^2 \tag{13}$$

其中, $i \in [1,8], j \in [1,8]$. 密文恢复为 $C^{(k)}(i,j)$ 后得到原密文分块 $C^{(k)}$.

2.2.2 密文分块统计量

Zeng[11]提出了一种图像分块的统计量,首先把图像分为若干个互不重叠的大小为 $m \times n$ 的明文分块,然后定义一个大小为 $m \times n$ 的矩阵 $M$:

$$M(i,j) = \begin{cases} 1, & \bmod(i,2) = \bmod(j,2) \\ -1, & \bmod(i,2) \neq \bmod(j,2) \end{cases} \tag{14}$$

其中, $i \in [1,m], j \in [1,n]$, $\bmod(x,2)$ 是模 2 运算的函数.例如,大小为 2×2 的矩阵 $M$ 如图 2 所示.

记 $d^{(k)}$为第 $k$ 个明文分块 $P^{(k)}$的统计量,$P^{(k)}(i,j)$为 $P^{(k)}$中点$(i,j)$位置上的明文值,则 $d^{(k)}$为

$$d^{(k)} = \sum_{i=1}^{m}\sum_{j=1}^{n}(P^{(k)}(i,j) \times M(i,j)) \tag{15}$$

在加密域中,信息隐藏者在不具有私钥$\lambda$的情况下不能将密文解密得到相应的明文,无法直接进行明文值的运算得到统计量 $d^{(k)}$.因此,本文利用 MMI 方法和查询相应密文映射表得到密文分块统计量,其值与对应的明文分块统计量相同.记密文分块统计量为 $d$,$d^{(k)}$是第 $k$ 个密文分块 $C^{(k)}$的统计量.首先,根据扩展欧几里德算法求出密文分块 $C^{(k)}$中密文 $C^{(k)}(i,j)$的模乘法逆元 $\theta_{C^{(k)}(i,j)}$,$\theta_{C^{(k)}(i,j)}$ 满足:

$$\theta_{C^{(k)}(i,j)} \cdot C^{(k)}(i,j) = 1 \bmod N^2 \tag{16}$$

之后定义大小为 8×8 的矩阵 $M_1$ 和 $M_2$,第 $k$ 个矩阵 $M_1^{(k)}$ 和 $M_2^{(k)}$ 满足:

$$M_1^{(k)}(i,j) = \begin{cases} C^{(k)}(i,j), & \text{if } \bmod(i,2) = \bmod(j,2) \\ \theta_{C^{(k)}(i,j)}, & \text{if } \bmod(i,2) \neq \bmod(j,2) \end{cases} \tag{17}$$

$$M_2^{(k)}(i,j) = \begin{cases} \theta_{C^{(k)}(i,j)}, & \text{if } \bmod(i,2) = \bmod(j,2) \\ C^{(k)}(i,j), & \text{if } \bmod(i,2) \neq \bmod(j,2) \end{cases} \tag{18}$$

其中, $i \in [1,8], j \in [1,8]$. 利用 $M_1^{(k)}$ 和 $M_2^{(k)}$ 可以计算出 $C^{(k)}$ 的密文形式的统计量 $c_{d1}(k)$ 和 $c_{d2}(k)$:

$$\begin{cases} c_{d1}(k) = \prod_{i=1}^{m}\prod_{j=1}^{n} M_1^{(k)}(i,j) \bmod N^2 \\ c_{d2}(k) = \prod_{i=1}^{m}\prod_{j=1}^{n} M_2^{(k)}(i,j) \bmod N^2 \end{cases} \tag{19}$$

然后通过密文形式的统计量 $c_{d1}(k)$和 $c_{d2}(k)$便可在密文映射表中查询出对应的统计量 $d^{(k)}$.下面是相应的公式推导和证明.为了方便理解,以分块大小为 2×2 的密文分块加以举例证明.分块大小为 2×2 的第 $k$ 个密文分块 $C^{(k)}$如图 3 所示.


Fig.2 Block $M$ sized 2×2

图 2 2×2 矩阵 $M$ 示意图

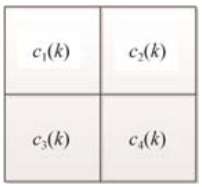



Fig.3 Cipher block $C^{(k)}$ sized 2×2

图 3 2×2 密文分块 $C^{(k)}$示意图

其中, $i\in[1,2]$, $j\in[1,2]$, $c_1(k),c_2(k),c_3(k),c_4(k)$分别是密文分块 $C^{(k)}$在点$(i,j)$位置上的密文 $C^{(k)}(i,j)$.设 $c_1(k),c_2(k),c_3(k),c_4(k)$分别是由明文 $P_1(k),P_2(k),P_3(k),P_4(k)$加密而成,并且对应的模乘法逆元为$\theta_1(k),\theta_2(k),\theta_3(k),\theta_4(k)$,它们满足:

$$\begin{cases}\theta_1(k)\cdot c_1(k)=\theta_1(k)\cdot g^{P_1(k)}\cdot r_1(k)^N=1\bmod N^2\\ \theta_2(k)\cdot c_2(k)=\theta_2(k)\cdot g^{P_2(k)}\cdot r_1(k)^N=1\bmod N^2\\ \theta_3(k)\cdot c_3(k)=\theta_3(k)\cdot g^{P_3(k)}\cdot r_1(k)^N=1\bmod N^2\\ \theta_4(k)\cdot c_4(k)=\theta_4(k)\cdot g^{P_4(k)}\cdot r_1(k)^N=1\bmod N^2\end{cases} \tag{20}$$

则密文分块 $C^{(k)}$的密文形式的统计量 $c_{d1}(k)$和 $c_{d2}(k)$为

$$\begin{cases}c_{d1}(k)=\prod_{i=1}^{2}\prod_{j=1}^{2}M_1(i,j)=c_1(k)\cdot\theta_2(k)\cdot\theta_3(k)\cdot c_4(k)=g^{P_1(k)+P_4(k)}\cdot r_1(k)^{2N}\cdot\theta_2(k)\cdot\theta_3(k)\bmod N^2\\ c_{d2}(k)=\prod_{i=1}^{2}\prod_{j=1}^{2}M_2(i,j)=\theta_1(k)\cdot c_2(k)\cdot c_3(k)\cdot\theta_4(k)=g^{P_2(k)+P_3(k)}\cdot r_1(k)^{2N}\cdot\theta_1(k)\cdot\theta_4(k)\bmod N^2\end{cases} \tag{21}$$

根据 Carmichael 理论[21],对于 $\forall a\in Z_{N^2}^*$ 有:

$$a^{N\lambda}=1\bmod N^2 \tag{22}$$

因此,按照第 1 节 Paillier 加密系统的描述,对于 $g\in Z_{N^2}^*$ 和 $r_1(k)\in Z_N^*$ 满足:

$$\begin{cases}g^{N\lambda}=1\bmod N^2\\ r_1(k)^{N\lambda}=1\bmod N^2\end{cases} \tag{23}$$

并且,

$$g^{N\lambda}\cdot r_1(k)^{N\lambda}=1\bmod N^2 \tag{24}$$

根据式(20)和式(24),可以推导出模乘法逆元的表达式:

$$\begin{cases}\theta_1(k)=g^{N\lambda-P_1(k)}\cdot r_1(k)^{N(\lambda-1)}\bmod N^2\\ \theta_2(k)=g^{N\lambda-P_2(k)}\cdot r_1(k)^{N(\lambda-1)}\bmod N^2\\ \theta_3(k)=g^{N\lambda-P_3(k)}\cdot r_1(k)^{N(\lambda-1)}\bmod N^2\\ \theta_4(k)=g^{N\lambda-P_4(k)}\cdot r_1(k)^{N(\lambda-1)}\bmod N^2\end{cases} \tag{25}$$

根据式(23)和式(25)中模乘法逆元的表达式,并考虑到统计量 $d^{(k)}$可能的取值,公式(21)可以化简为

当 $d^{(k)}\geqslant 0$ 时,

$$\begin{cases}c_{d1}(k)=g^{P_1-P_2+P_4-P_3}\bmod N^2\\ c_{d2}(k)=g^{2N\lambda+P_2-P_1+P_3-P_4}\bmod N^2\end{cases} \tag{26}$$

否则,

$$\begin{cases}c_{d1}(k)=g^{2N\lambda+P_2-P_1+P_3-P_4}\bmod N^2\\ c_{d2}(k)=g^{P_1-P_2+P_4-P_3}\bmod N^2\end{cases} \tag{27}$$

由于灰度图像中两个像素值差值的绝对值的取值范围为 0~255,则对于每个 2×2 分块的统计量 $d^{(k)}$的绝对

值的取值范围为 0~$255\times\frac{2\times 2}{2}$.由于图像像素之间存在相关性,$d^{(k)}$通常是一个较小的值(在本文实验中,对于 8×8 分块 $d^{(k)}$的值均小于 500).另外,在下文中对统计量进行直方图平移嵌入水印时,会使 $d^{(k)}$产生大小为 $T+G$ 的改变,记统计量可能取值的绝对值为 $d_p$,则 $d_p\in\left[0,255\times\frac{2\times 2}{2}+T+G\right]$.利用公钥$(N,g)$,当 $d_p\in Z_N$ 时可以加密得到:

$$c_{d_p}=g^{d_p}\bmod N^2, d_p=0,1,...,255\times\frac{2\times 2}{2}+T+G \tag{28}$$

创建一个与 $d_p$ 一一对应的密文映射表 $c_{dp}$ 如图 4 所示.

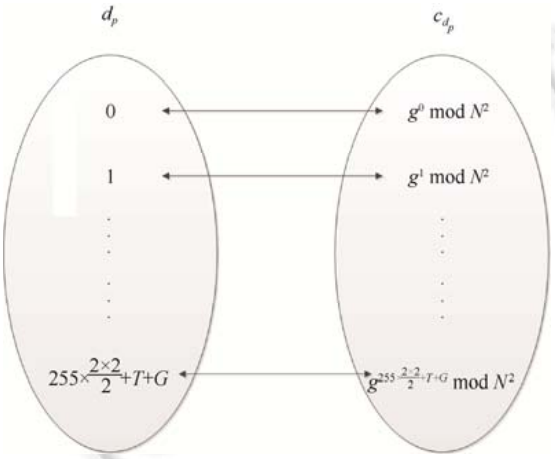



Fig.4 Mapping table $c_{dp}$ of $d_p$

图 4 $d_p$的密文映射表 $c_{dp}$

其中,密文映射表 $c_{dp}$ 就是对所有统计量可能取值的绝对值 $d_p$ 进行加密后得到的所有密文值的集合.当 $c_{d1}(k)$ 与密文映射表 $c_{dp}$ 中的值 $c_{d_p}[x]$ 匹配时,则表明 $d^{(k)}\geqslant 0$, 得到 $d^{(k)}=d_p[x]$. 当 $c_{d2}(k)$ 与密文映射表中的值 $c_{dp}$ 匹配时,则表明 $d^{(k)}<0$, 得到 $d^{(k)}=-d_p[x]$. 其中, $x\in\left[0,255\times\frac{2\times 2}{2}+T+G\right]$, 代表在密文映射表中的第 $x$ 个值.因此,信息隐藏者可以在没有私钥$\lambda$的情况下在加密图像中得到分块统计量 $d^{(k)}$.

利用 MMI 方法和查询密文映射表的策略,对大小为 512×512 的 Lena 灰度图像进行 8×8 分块和加密后得到的密文分块统计量 $d$ 的直方图分布如图 5 所示.

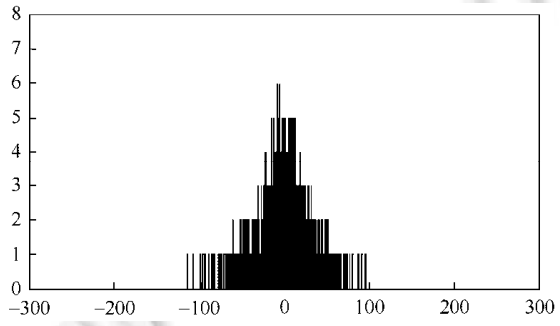



Fig.5 Histogram of statistical values of the encrypted Lena image blocks

图 5 Lena 图像分块在加密后的统计量直方图

2.2.3 统计量直方图平移

信息隐藏者首先分别选取两个正整数 $T$ 和 $G$ 作为嵌入密钥$(T,G)$,其中,$T>d_{\max}$,$d_{\max}$ 是统计量 $d$ 中绝对值最大的数,通常,$T$ 和 $G$ 取为(8×8)/2 的倍数.然后计算得到嵌入系数 $B$:

$$B=\left\lceil \frac{(T+G)\times 2}{8\times 8} \right\rceil \tag{29}$$

其中,函数 $\lceil\cdot\rceil$ 为向上取整.通过对统计量 $d$ 进行直方图平移,可以把水印嵌入到加密图像 $E[I]$中,每个 8×8 密文分块 $C^{(k)}$可以嵌入 1 比特水印,记嵌入水印后的密文分块为矩阵 $C_w^{(k)}$. 当嵌入比特为 0 时,不需要对 $C^{(k)}$进行处理,即 $C_w^{(k)}=C^{(k)}$. 当嵌入比特为 1 时,对 $C^{(k)}$中的密文 $C^{(k)}(i,j)$ 进行处理得到 $C_w^{(k)}(i,j)$:

$$C_w^{(k)}(i,j)=\begin{cases} C^{(k)}(i,j)\cdot g^B=g^{P^{(k)}(i,j)+B}\cdot r_1(k)^N \bmod N^2, & \text{if } d^{(k)}\in[0,T) \text{ and } \bmod(i,2)=\bmod(j,2) \\ C^{(k)}(i,j)\cdot g^B=g^{P^{(k)}(i,j)+B}\cdot r_1(k)^N \bmod N^2, & \text{if } d^{(k)}\in(-T,0) \text{ and } \bmod(i,2)\neq\bmod(j,2) \\ C^{(k)}(i,j), & \text{else} \end{cases} \tag{30}$$

其中,$i\in[1,8]$,$j\in[1,8]$,$C_w^{(k)}(i,j)$ 为嵌入水印后的密文.记 $P_w^{(k)}(i,j)$ 为 $C_w^{(k)}(i,j)$ 解密后的明文值,加密域中的处理相当于在明文域中使明文像素值 $P^{(k)}(i,j)$ 变为 $P_w^{(k)}(i,j)$:

$$P_w^{(k)}(i,j)=\begin{cases} P^{(k)}(i,j)+B, & \text{if } d^{(k)}\in[0,T) \text{ and } \bmod(i,2)=\bmod(j,2) \\ P^{(k)}(i,j)+B, & \text{if } d^{(k)}\in(-T,0) \text{ and } \bmod(i,2)\neq\bmod(j,2) \\ P^{(k)}(i,j), & \text{else} \end{cases} \tag{31}$$

记嵌入水印后的密文分块 $C_w^{(k)}$ 的统计量为 $d_w^{(k)}$, 嵌入比特 0 后, $d_w^{(k)}$ 在范围$(-T,T)$内,因此,称$(-T,T)$为比特 0 区;嵌入比特 1 后,$d^{(k)}$发生大小为平移量 $T+G$ 的改变,使 $d_w^{(k)}$ 在范围 $[-2T-G,-T-G)$ 或 $[T+G,2T+G)$ 内,因此,称 $[-2T-G,-T-G)$ 和 $[T+G,2T+G)$ 为比特 1 区.同样,为了提高统计量的安全性,利用密钥 $K_s$ 对含 $C_w^{(k)}$ 中的密文按照公式(7)与 $E[0,r_{(i,j)}(k)]$ 进行同态乘法得到矩阵 $C_w'^{(k)}$. 最终,得到含有水印的密文图像 $E[I_w]$. 例如,图 5 所示 Lena 图像在加密后进行 8×8 分块的 $d_{\max}$ 为 114,选取 $T$=128,$G$=64,嵌入 4 096 比特水印后统计量 $d_w$ 的直方图分布如图 6 所示.

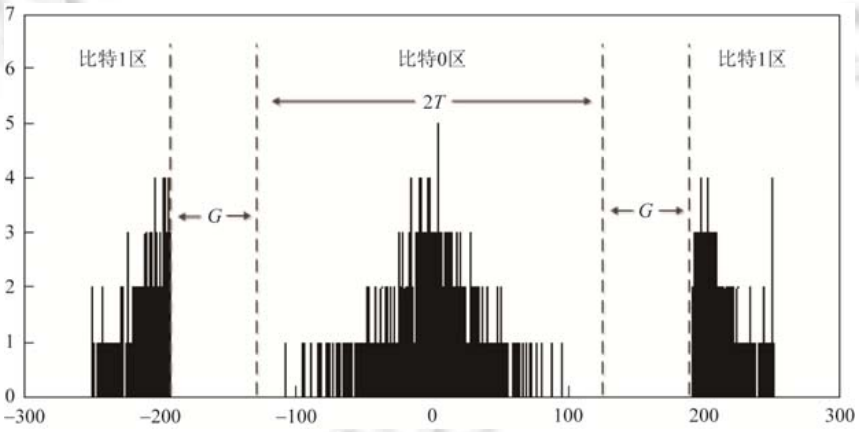



Fig.6 Histogram of statistical values of the encrypted and watermarked Lena image blocks

图 6 Lena 图像在加密和嵌入水印后的分块统计量直方图

### 2.3 信息提取

2.3.1 加密域提取水印和恢复加密图像

与嵌入过程相同,接收者首先对含有水印的密文图像 $E[I_w]$ 进行 8×8 分块,并利用密钥 $K_s$ 和 MMI 方法按照公式(13)把 $C_w'^{(k)}$ 恢复为 $C_w^{(k)}$, 然后再按照第 2.3.2 节描述的模乘法逆元 MMI 方法和查询密文映射表得到 $C_w^{(k)}$ 的

统计量 $d_w^{(k)}$.通过嵌入密钥$(T,G)$可以得到嵌入系数 $B$,则 $C_w^{(k)}$ 中提取的水印 $w^{(k)}$为

$$w^{(k)}=\begin{cases}0, & \text{if } d_w^{(k)}\in(-T,T)\\ 1, & \text{else}\end{cases} \tag{32}$$

通过对统计量 $d_w$进行与嵌入过程相反的直方图平移,可以恢复原加密图像,方法如下所述.首先根据扩展欧几里德算法求出 $g^B$ 的模乘法逆元 $\theta_{g^B}$,$\theta_{g^B}$ 满足:

$$\theta_{g^B}\cdot g^B=1\bmod N^2 \tag{33}$$

对 $C_w^{(k)}$ 中的密文 $C_w^{(k)}(i,j)$ 进行处理得到 $C^{(k)}(i,j)$:

$$C^{(k)}(i,j)=\begin{cases}C_w^{(k)}(i,j)\cdot\theta_{g^B}=C^{(k)}(i,j)\cdot g^B\cdot\theta_{g^B}\bmod N^2, & \text{if } d_w^{(k)}\in[T+G,2T+G) \text{ and} \bmod(i,2)=\bmod(j,2)\\ C_w^{(k)}(i,j)\cdot\theta_{g^B}=C^{(k)}(i,j)\cdot g^B\cdot\theta_{g^B}\bmod N^2, & \text{if } d_w^{(k)}\in[-2T-G,-T-G) \text{ and} \bmod(i,2)\neq\bmod(j,2)\\ C_w^{(k)}(i,j), & \text{if } d_w^{(k)}\in(-T,T)\end{cases} \tag{34}$$

其中,$i\in[1,8]$,$j\in[1,8]$,加密域中的处理相当于在明文域中使明文值 $P_w^{(k)}(i,j)$ 变为 $P^{(k)}(i,j)$:

$$P^{(k)}(i,j)=\begin{cases}P_w^{(k)}(i,j)-B, & \text{if } d^{(k)}\in[T+G,2T+G) \text{ and} \bmod(i,2)=\bmod(j,2)\\ P_w^{(k)}(i,j)-B, & \text{if } d^{(k)}\in[-2T-G,-T-G) \text{ and} \bmod(i,2)\neq\bmod(j,2)\\ P_w^{(k)}(i,j), & \text{if } d^{(k)}\in(-T,T)\end{cases} \tag{35}$$

处理的结果使统计量 $d_w^{(k)}$ 恢复为 $d^{(k)}$.同样,为了提高统计量的安全性,利用密钥 $K_s$ 对 $C^{(k)}$中的密文按照公式(7)与 $E[0,r_{(i,j)}(k)]$ 进行同态乘法得到 $C'^{(k)}$.最终,恢复出原密文图像 $E[I]$.

2.3.2 解密后提取水印和恢复原始图像

接收者在有私钥$\lambda$和嵌入密钥$(T,G)$的情况下,可以对含有水印的密文图像 $E[I_w]$进行解密后提取水印和恢复原始图像.利用$\lambda$对 $E[I_w]$进行解密,解密的过程如第 1 节所述,记解密后的结果为 $I_w$.首先对 $I_w$ 进行 8×8 分块,则第 $k$ 个 8×8 的含水印的明文分块为 $P_w^{(k)}$.通过嵌入密钥$(T,G)$可以得到嵌入系数 $B$,根据公式(15)得到 $P_w^{(k)}$ 的统计量 $d_w^{(k)}$.水印 $w^{(k)}$可根据公式(32)来提取.原始图像的恢复可以通过对 $P_w^{(k)}$ 中的明文像素值按照公式(35)进行处理,处理的结果使统计量 $d_w^{(k)}$ 恢复为 $d^{(k)}$,并且 $P_w^{(k)}$ 中的明文像素值 $P_w^{(k)}(i,j)$ 恢复为 $P^{(k)}(i,j)$ 得到原明文分块 $P^{(k)}$,最终恢复出原始图像 $I$.

2.3.3 在攻击或图像处理后提取水印

解密后含水印的图像在传递过程中可能会受到图像处理或干扰.通过将水印嵌入到分块的统计直方图,水印算法对常见的图像处理操作(JPEG/JPEG 2000 压缩和高斯噪声等)具有一定的鲁棒性.如图 6 所示,比特 0 区和比特 1 区间隔着大小为 $G$ 的鲁棒区间.因此,当解密后含水印的图像受到一定程度的图像处理或操作后,虽然造成了分块统计量 $d$ 小范围的变动但在未进入错误区间时,接收方仍能正确提取水印.为了提高水印图像在遭受干扰或处理后提取水印的正确率,参照文献[11,24]的算法思路,本文采用 3 种提取方案和多数投票系统策略来最终判定提取的水印比特.

**提取方案 1**.

$$w_1^{(k)}=\begin{cases}0, & \text{if } d_w^{(k)}\in(-T,T)\\ 1, & \text{otherwise}\end{cases} \tag{36}$$

**提取方案 2**. 重新定义比特 0 区为 $(-T-G/3,T+G/3)$, 则水印提取为

$$w_2^{(k)}=\begin{cases}0, & \text{if } d_w^{(k)}\in(-T-G/3,T+G/3)\\ 1, & \text{otherwise}\end{cases} \tag{37}$$

**提取方案 3**. 利用 $k$-means 聚类[12]来提取水印比特.图 7 为经过 JPEG 2000 压缩后分块统计量的直方图分布,可以看出,图像处理后分块统计量的分布大致分为 3 类,分别为 *class* 1、*class* 2 和 *class* 3,则水印的提取为

$$w_3^{(k)}=\begin{cases}0, & \text{if } d_w^{(k)}\in class\ 2\\ 1, & \text{if } d_w^{(k)}\in class\ 1 \text{ or } class\ 3\end{cases} \tag{38}$$

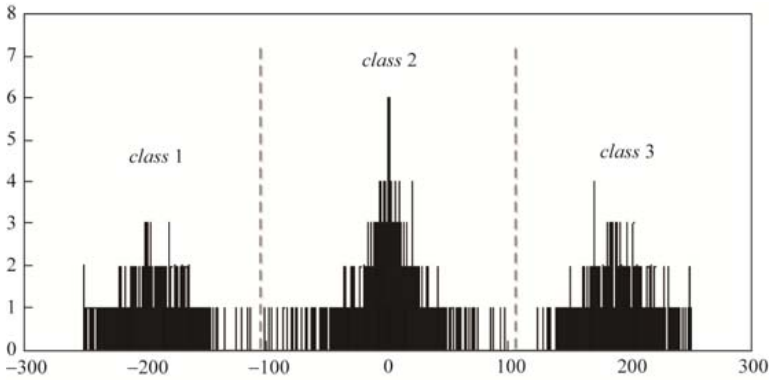



Fig.7 Histogram of statistical values of the watermarked blocks under JPEG 2000 compression

图 7 经过 JPEG 2000 压缩后分块统计量的直方图

最终通过多数投票系统来判决提取的水印比特,多数投票判决为

$$w^{(k)}=\begin{cases}w_1^{(k)}, & \text{if } w_1^{(k)}=w_2^{(k)} \text{ and } w_1^{(k)}=w_3^{(k)}\\ w_1^{(k)}, & \text{if } w_1^{(k)}=w_2^{(k)} \text{ and } w_1^{(k)}\neq w_3^{(k)}\\ w_2^{(k)}, & \text{if } w_1^{(k)}\neq w_2^{(k)} \text{ and } w_2^{(k)}=w_3^{(k)}\\ w_3^{(k)}, & \text{if } w_1^{(k)}\neq w_2^{(k)} \text{ and } w_1^{(k)}=w_3^{(k)}\end{cases} \tag{39}$$

在图像未受到干扰或处理的情况下,公式(32)和 3 种提取投票系统的方案均可无损地提取水印.而在图像受到干扰或处理的情况下,3 种提取投票系统提取水印的准确率更高.因此,本文采用 3 种提取方案和多数投票系统作为水印提取的方案.

### 2.4 溢出处理

在加密域中嵌入水印后,密文值的改变相当于其相应的明文值加上了一个大小为嵌入系数 $B$ 的值,改变后的明文值可能会超出灰度图像中像素值的范围[0,255],则解密后可能会出现溢出问题.$E[I_w]$解密后的结果 $I_w$ 中值的分布有 3 种类型:类型 1,值在范围[0,255]内;类型 2,值在范围[$B$,255+$B$]内;类型 3,值在范围[0,255+$B$]内.类型 1 没有溢出问题,而类型 2 和类型 3 出现了溢出问题,本文参照文献[11]中的算法思路,做出相应处理.

分块统计量具有特殊性,对于任意一个分块中的所有值加上或减去一个值,分块统计量的值不变.因此,对于类型 2,可以将 $I_w$ 中 $d_w^{(k)}$ 在比特 1 区分块中的所有值减去 $B$, $d_w^{(k)}$ 的值不变,$I_w$ 由类型 2 转换为类型 1 得到 $I'_w$. 在接收端,将 $I'_w$ 中 $d_w^{(k)}$ 在比特 1 区分块中的所有值加上 $B$ 可恢复出 $I_w$.

对于类型 3,首先统计 $I_w$ 中的值属于范围[0,$B$]和[255,255+$B$]的数量.如果值属于范围[0,$B$]的数量少于[255,255+$B$]的数量,则把这种类型记为类型 3.1.然后把值属于范围[0,$B$]的位置信息标记下来,并将全部标记位置上的值加上 B,则类型 3.1 转换为类型 2,然后可以再转换为类型 1.最后用一种低失真的可逆水印的方案将这些标记位置信息作为附加信息嵌入图像当中得到 $I'_w$. 在接收方,首先从 $I'_w$ 提取嵌入的附加信息,然后将类型 1 恢复为类型 2,最后将标记位置上的所有值减去 $B$ 则可恢复为类型 3.1,即 $I'_w$ 恢复为 $I_w$;如果值属于范围[255,255+$B$]的数量少于[0,$B$]的数量,则把这种类型记为类型 3.2.然后把值属于范围[255,255+$B$]的位置信息标记下来,并将全部标记位置上的值减去 $B$,则类型 3.2 转换为类型 1.最后用一种低失真的可逆水印的方案将这些标记位置信息作为附加信息嵌入图像当中得到 $I'_w$. 在接收方,首先从 $I'_w$ 提取嵌入的附加信息,然后将标记位置上的所有值加上 $B$ 则可恢复为类型 3.2,即 $I'_w$ 恢复为 $I_w$.

对于自然图像,像素的灰度值集中于[25,230]的范围内,所以在一般情况下,只要嵌入系数 $B$ 的取值适当,图像在嵌入水印后很少会出现溢出问题,并且在绝大多数情况下不会出现类型 3.根据本文的溢出处理方案,有效

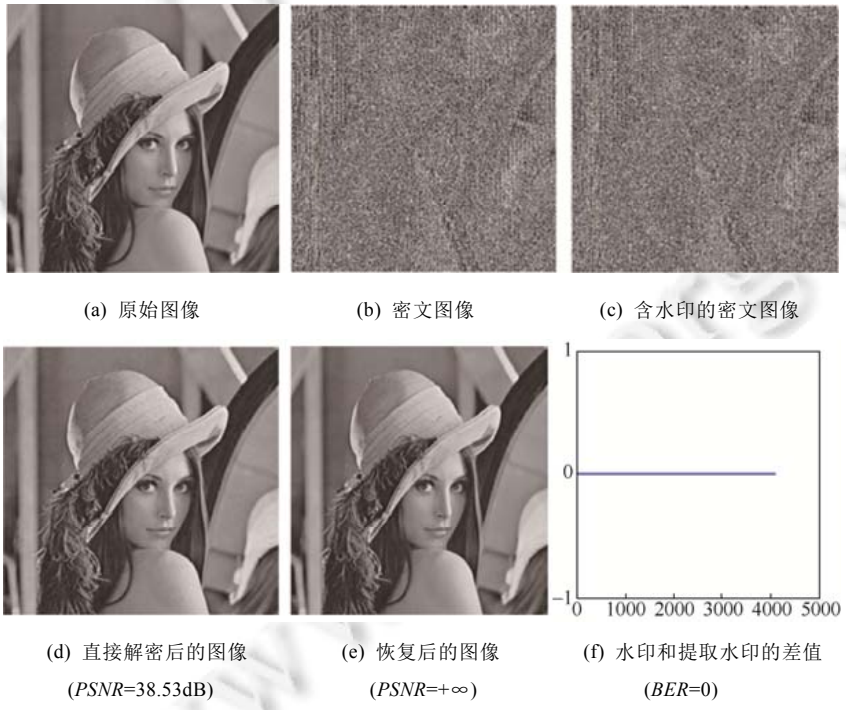
地解决了可能发生的溢出情况.

## 3 实验结果及分析

实验中首先给出以 Lena 图像作为载体的实验结果来验证本文算法的可行性.本文采用峰值信噪比 PSNR(peak signal to noise ratio)来衡量解密后含水印的图像质量,其值越高,则表示嵌入水印的不可感知性越强,图像的质量越好.假设 $I$ 代表原始图像,$I'$ 代表解密后含水印的图像,$(i,j)$表示图像像素坐标,则 PSNR 的计算公式为

$$PSNR = 10 \times \lg \frac{h \times w \times 255^2}{\sum_{i=1}^{h}\sum_{j=1}^{w}[I(i,j) - I'(i,j)]^2} \tag{40}$$

其中,$h$ 和 $w$ 代表图像的尺寸,$i \in [1,h]$, $j \in [1,w]$. 另外,本文采用比特误差率 BER(bit error rate)来衡量提取水印的正确性,其值越低,表明提取水印的正确性越高.实验中选取大小为 512×512 的 8 比特的 Lena 灰度图像(如图 8(a)所示)作为测试图像,嵌入的水印是一段 4 096 比特的伪随机序列,采用 Paillier 加密系统的参数设置为 $p$=61, $q$=67,其可加密的明文的上限值 $N=p\cdot q$=4087.具体的实验结果如图 8 所示.首先对图 8(a)所示的原始图像进行 8×8 分块并利用公钥$(N,g)$和密钥 $K_s$ 进行加密得到密文图像(如图 8(b)所示),然后利用嵌入密钥($T$=128,$G$=64)嵌入水印得到含水印的密文图像(如图 8(c)所示).其中,图 8(b)和图 8(c)皆为归一化后的密文图像,目的是显示图像加密后的效果.通过私钥$\lambda$对图 8(c)进行解密得到直接解密后的图像(如图 8(d)所示),该图像的 PSNR 为 38.53dB.最后通过嵌入密钥($T$=128,$G$=64)提取嵌入的水印和恢复图像(如图 8(e)所示),该图像的 PSNR 为+∞,表明恢复出来的图像与原始图像完全相同,图 8(f)表明对所有水印比特完成了正确提取.实验结果说明,本文算法实现了加密域中水印的可逆嵌入和提取以及原始图像的恢复.



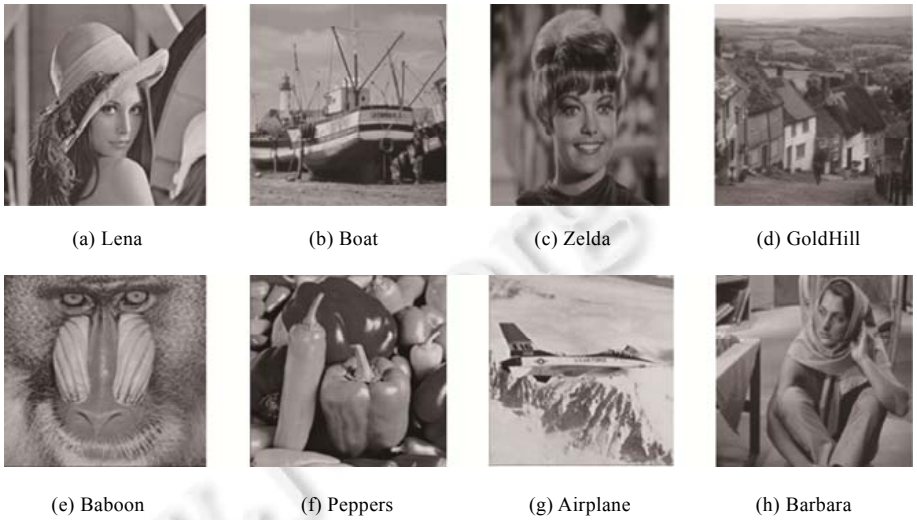

(a) 原始图像 (b) 密文图像 (c) 含水印的密文图像

(d) 直接解密后的图像 (*PSNR*=38.53dB) (e) 恢复后的图像 (*PSNR*=+∞) (f) 水印和提取水印的差值 (*BER*=0)

Fig.8 Watermark embedding and extraction testing results with Lena

图 8 以 Lena 图像为载体的水印嵌入和提取测试

为了更进一步地评估本文算法的性能,实验选取了如图 9 所示的 8 幅大小为 512×512 的 8 比特灰度图像作为载体进行测试.实验中对 8 幅载体图像进行 8×8 分块,嵌入的水印信息为 4 096 比特的伪随机系列.加密域嵌入水印后,密文值的改变相当于其相应的明文值加上了一个大小为嵌入系数 $B$ 的值,因此,$B$ 直接影响到解密后图像的质量.随着 $B$ 的增大,相应的失真就越大,导致 PSNR 降低.经过测试,当 8 幅图像以相同的 $B$ 嵌入水印时,解密后图像的 PSNR 值基本相同,嵌入系数 $B$ 和 PSNR 值的关系如图 10 所示.由图 10 可知,当 $B$ 为 16 时,图像的 PSNR 略大于 30dB,基于不可觉察度的考虑,为了获得较好的图像质量,本文设定最大嵌入系数 $B_{max}$=16,即 $B$ 的值不能超过 16.



(a) Lena (b) Boat (c) Zelda (d) GoldHill

(e) Baboon (f) Peppers (g) Airplane (h) Barbara

Fig.9 Eight standard example images

图 9 8 幅标准测试图像

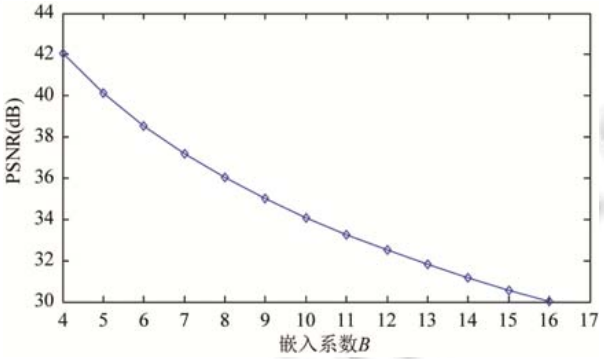



Fig.10 Relationship between embedding strength $B$ and PSNR value

图 10 嵌入强度 $B$ 和 PSNR 值的关系

由图 6 可知,比特 0 区和比特 1 区间隔着大小为 $G$ 的鲁棒区间,$G$ 越大,则鲁棒性越强.但是,$G$ 越大会使嵌入系数 $B$ 随之增大,导致 PSNR 值有所降低.因此,实际中可根据需要调节阈值 $G$,若需要更强的鲁棒性,可以选择较大的阈值 $G$;若需要水印图像质量更好,则可选择较小的阈值 $G$.在 JPEG 压缩鲁棒性实验中,我们采用了 ACDsee 14.0 软件对解密后含水印的图像进行 JPEG 压缩.如图 11 所示,随着压缩质量因子数值的降低(表示压缩强度逐渐增大),提取水印的比特误差率 BER 逐渐升高.在同样的压缩强度下,BER 随着阈值 $G$ 的增大而降低,说明随着阈值 $G$ 的增大,嵌入强度增加,水印对 JPEG 压缩的鲁棒性增强.

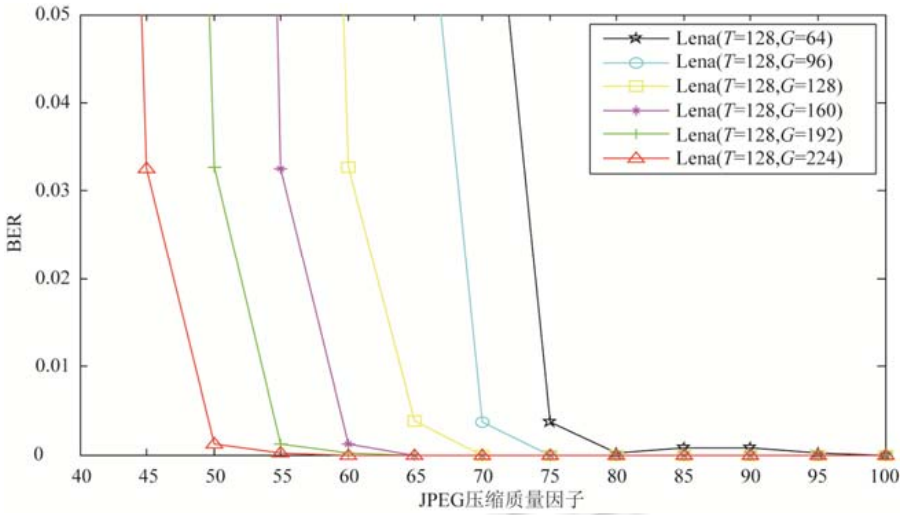



Fig.11 Effection of the value $G$ on the robustness to JPEG compression

图 11 阈值 $G$ 对水印抗 JPEG 压缩的影响

我们采用 ACDsee 14.0 软件对解密后的水印图像进行 JPEG 2000 压缩,采用存活率(surviving bit rate)来衡量水印算法对 JPEG 2000 压缩的鲁棒性.存活率与最大压缩率的关系为:存活率=8/最大压缩率,即存活率越小,其压缩倍数越大,鲁棒性越强.图 12 所示为在不同阈值 $G$ 下能够正确提取水印的最小存活率,即在该阈值下,若存活率高于对应的存活率,即当压缩率更低时,能够完全正确地提取水印.由图 12 可知,随着阈值 $G$ 的增大,最小存活率减小,即水印对抗 JPEG 2000 压缩的鲁棒性增强.

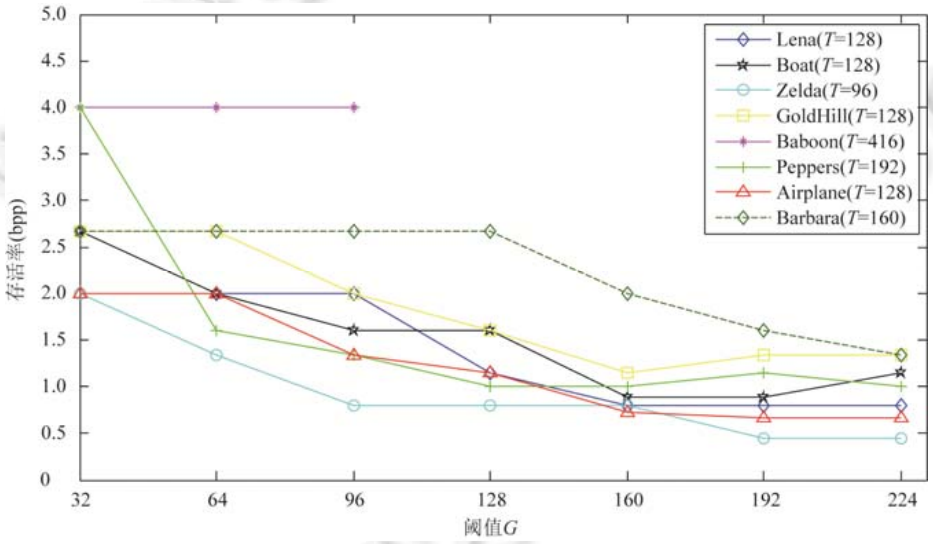



Fig.12 Effection of the value $G$ on the robustness to JPEG 2000 compression

图 12 阈值 $G$ 对水印抗 JPEG 2000 压缩的影响

为了测试本文算法对其他图像处理操作的鲁棒性,本文采用 MATLAB 软件来对水印图像添加高斯噪声和椒盐噪声.表 1 给出了 8 幅测试图像在设定的参数下,嵌入 4 096 比特水印后的解密图像在受到不同图像处理操作后提取水印的比特误差率(BER)(%).嵌入系数 $B$ 设定为 16.从表 1 可以看出,对于质量因子为 25 的 JPEG 压缩,除了 Baboon 图像外,BER 均小于 1%;对于存活率为 0.67bpp 的 JPEG 2000 压缩,除了 Baboon 和 Barbara 图像

外,BER 均小于 1%;对于方差为 0.005 的高斯噪声,除了 Baboon 图像外,BER 均约为 5%;对于方差为 0.01 的椒盐噪声,除了 Baboon 图像外,BER 均小于 5%.与其他测试图像相比,水印以 Baboon 图像为载体性能较差的主要原因是 Baboon 图像的纹理比较复杂,8×8 分块后的 $d_{max}$=388,需要设定 $T$=416,即使设定 $G$=0 都会使嵌入系数 $B$=13,因此图像的失真较大.在本文设定的最大嵌入系数 $B_{max}$=16 条件下,$G$ 最大只能取为 96,因此鲁棒性较差.由此可知,本文算法对于比较平滑的图像有较好的性能.

**Table 1** Robustness of the proposed method against several image processing operations

**表 1** 本文水印算法在几种常见图像处理操作下的鲁棒性

| 测试图像 | $T$ | $G$ | PSNR(dB) | JPEG 压缩 (质量因子 25) | JPEG 2000 压缩 (存活率 0.67bbp) | 高斯噪声 (方差 0.005) | 椒盐噪声 (方差 0.01) |
|---|---|---|---|---|---|---|---|
| Lena | 128 | 384 | 30.01 | 0.10 | 0 | 4.93 | 2.61 |
| Boat | 128 | 384 | 30.01 | 0.02 | 0.02 | 4.96 | 2.83 |
| Zelda | 96 | 416 | 30.01 | 0.02 | 0 | 5.79 | 3.32 |
| Goldhill | 128 | 384 | 30.01 | 0.10 | 0.05 | 4.40 | 2.73 |
| Baboon | 416 | 96 | 30.01 | 2.61 | 48.19 | 9.25 | 6.52 |
| Peppers | 192 | 320 | 30.01 | 0.29 | 0.05 | 4.81 | 2.71 |
| Airplane | 128 | 384 | 30.01 | 0.024 | 0 | 5.27 | 2.83 |
| Barbara | 160 | 352 | 30.01 | 0.34 | 8.45 | 4.98 | 2.32 |

由于每个密文分块都可以嵌入 1 比特水印,所以图像进行分块的尺寸越小,则嵌入容量越大.对于大小为 $h\times w$ 的图像并且分块大小为 $m\times n$ 的最大嵌入容量为$[h/w]\times[w/n]$.其中,函数$\lfloor\cdot\rfloor$为向下取整.但是,通常进行小尺寸分块的嵌入系数 $B$ 会比较大,导致 PSNR 降低.以 Lena 图像为例,当分块尺寸为 4×4 时,它的 $d_{max}$ 为 120,这就意味着:

$$B=\left\lceil\frac{(T+G)\times 2}{m\times n}\right\rceil\geqslant\left\lceil\frac{(128+G)\times 2}{4\times 4}\right\rceil\geqslant 16 \tag{41}$$

设 $T$ 为 128,只有当 $G$ 取 0 时,$B$=16 才能满足本文设定的 $B_{max}$=16.为了衡量嵌入容量和图像失真以及鲁棒性的关系,给出 Lena 图像以不同分块尺寸嵌入最大嵌入容量的性能,见表 2.在不超出最大嵌入系数 $B_{max}$=16 的情况下,除分块尺寸为 4×4 在 JPEG 压缩因子为 100 时提取水印 $BER$=4.52%、存活率为 4 时提取水印 $BER$=3.89%,其他分块尺寸下的 JPEG 压缩因子和存活率都是在给定参数下能够正确地提取水印的最小 JPEG 压缩因子和存活率,并且 $G$ 是能正确提取水印的最小阈值.由表 2 可知,若分块尺寸越大,则最大嵌入容量越小,并且图像的失真越小,鲁棒性越强.经过实验测试,其结果表明,8×8 分块在嵌入容量、图像失真和鲁棒性之间有较好的平衡性.

**Table 2** The performance of Lena image in different block sizes

**表 2** Lena 图像不同分块尺寸的性能

| 分块尺寸 | 嵌入容量(bit) | $T$ | $G$ | $B$ | PSNR(dB) | JPEG 压缩因子 | 存活率(bpp) |
|---|---|---|---|---|---|---|---|
| 4×4 | 16 384 | 128 | 0 | 16 | 30.07 | 100 | 4 |
| 4×8 | 8 192 | 128 | 16 | 9 | 35.03 | 100 | 4 |
| 8×4 | 8 192 | 128 | 16 | 9 | 35.03 | 100 | 4 |
| 8×8 | 4 096 | 128 | 32 | 5 | 40.12 | 98 | 2.66 |
| 8×16 | 2 048 | 256 | 64 | 5 | 40.15 | 93 | 2 |
| 16×8 | 2 048 | 256 | 64 | 5 | 40.15 | 93 | 2.66 |
| 16×16 | 1 024 | 256 | 128 | 3 | 44.62 | 93 | 2 |

目前,在文献中尚未有有效的加密域图像鲁棒可逆水印算法的报道,因此无法将本文提出的同态加密域图像鲁棒可逆水印算法与前人的研究结果进行公平对比.为了进一步说明本文算法的鲁棒性,我们与前期具有代表性的一种明文域图像鲁棒可逆水印算法[10]进行了性能比较.测试中,使用相同的载体进行 8×8 分块嵌入相同的水印容量,本文算法与文献[10]中 Ni 算法的鲁棒性比较见表 3.由表 3 可知,除了 Baboon 图像以外,本文算法的图像质量和鲁棒性都优于文献[10].值得一提的是,本文算法是在加密域中嵌入水印,可以更好地在云端保护用户的数据隐私,相比文献[10]以及其他在明文域嵌入鲁棒可逆水印的方案,本文算法更加适用于当下大数据背景下的云计算安全领域.

**Table 3** Performance comparison of the proposed method against the Ni's method in Ref.[10]
**表 3** 与文献[10]中 Ni 算法的鲁棒性能比较

| 测试图像 | 文献[10]中 Ni 算法 | | | 本文算法 | | | | |
|---|---|---|---|---|---|---|---|---|
| | PSNR(dB) | 嵌入容量(bit) | 存活率(bpp) | *T* | *G* | PSNR(dB) | 嵌入容量(bit) | 存活率(bpp) |
| Lena | 40.20 | 792 | 0.80 | 128 | 128 | 43.27 | 792 | 0.66 |
| Boat | 40.50 | 560 | 1.00 | 128 | 160 | 43.54 | 560 | 0.61 |
| Baboon | 38.70 | 585 | 1.60 | 416 | 96 | 38.34 | 585 | 2.00 |

## 4 结 论

加密域鲁棒可逆水印技术通过加密手段来保护数据在云端的隐私,通过可逆水印来实现对敏感载体的完整性认证并通过鲁棒水印来进行保护数据在解密后的版权.通过结合 Paillier 加密系统、构造统计量和直方图平移技术,本文提出了一种新的同态加密域图像鲁棒可逆水印算法.为了保护数据在云端的隐私并允许对密文数据进行算术操作,数据在上传云端之前进行同态加密.为了能在云端的加密数据中嵌入水印,对图像采用了分块加密和在加密域中构造分块统计量,并结合 Paillier 加密系统的同态特性来实现基于统计量直方图平移的水印嵌入.算法的关键在于密文分块统计量的构造、模乘法逆元 MMI 方法的运用、利用同态特性构建密文映射表和嵌入水印,使得在加密域中可以获得统计量直方图进行水印嵌入.

在本文中,我们对水印算法在加密域和明文域的提取和图像的恢复进行了详细分析,并对可能出现的溢出情况给出了处理方案.最后在实验部分,我们选择一些标准的例子图像来测试算法的可逆性和鲁棒性.实验结果表明:(1) 水印算法的嵌入失真较小,具有良好的保真性;(2) 水印算法是可逆的,在未受到攻击的情况下水印能分别在加密域和明文域提取并且原始密文图像或原始明文图像能够无损恢复;(3) 水印具有良好的鲁棒性,解密后的含水印图像在经受一定的图像处理操作下仍能正确地提取水印.在云计算大数据背景下,由于同态加密域鲁棒可逆水印技术在隐私保护和数据安全上的潜在应用前景,本文算法具有很好的理论研究意义和实用价值.

**References**:

[1] Honsinger CW, Jones P, Rabbani M, Stoffel JC. Lossless recovery of an original image containing embedded data. Int'l Cl: G06K 9/00 US 6278791 B1, 2001-08-21.

[2] Feng JB, Lin IC, Tsai CS, Chu YP. Reversible watermarking: current status and key issues. Int'l Journal of Network Security, 2006,2(3):161–171.

[3] Celik MU, Sharma G, Tekalp AM, Saber E. Lossless generalized-LSB data embedding. IEEE Trans. on Image Processing, 2005, 14(2):253–266. [doi: 10.1109/TIP.2004.840686]

[4] Tian J. Reversible data embedding using a difference expansion. IEEE Trans. on Circuits and Systems for Video Technology, 2003, 13(8):890–896. [doi: 10.1109/TCSVT.2003.815962]

[5] Ni ZC, Shi YQ, Ansari N, Su W. Reversible data hiding. IEEE Trans. on Circuits and Systems for Video Technology, 2006,16(3): 354–362. [doi: 10.1109/TCSVT.2006.869964]

[6] Li XL, Yang B, Zeng TY. Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. IEEE Trans. on Image Processing, 2011,20(12):3524–3533. [doi: 10.1109/TIP.2011.2150233]

[7] Vleeschouwer CD, Delaigle JE, Macq B. Circular interpretation of histogram for reversible watermarking. In: Proc. of the IEEE Workshop of Multimedia Signal Process. 2001. 345–350. [doi: 10.1109/MMSP.2001.962758]

[8] Vleeschouwer CD, Delaigle JE, Macq B. Circular interpretation of bijective transformations in lossless watermarking for media asset management. IEEE Trans. on Multimedia, 2003,5(1):97–105. [doi: 10.1109/TMM.2003.809729]

[9] Zou D, Shi YQ, Ni Z, Su W. A semi-fragile lossless digital watermarking scheme based on integer wavelet transform. IEEE Trans. on Circuits and Systems for Video Technology, 2006,16(10):1294–1300. [doi: 10.1109/TCSVT.2006.881857]

[10] Ni Z, Shi YQ, Ansari N, Su W, Sun Q, Lin X. Robust lossless image data hiding designed for semi-fragile image authentication. IEEE Trans. on Circuits and Systems for Video Technology, 2008,18(4):890–896. [doi: 10.1109/TCSVT.2008.918761]

[11] Zeng XT, Ping LD, Pan XZ. A lossless robust data hiding scheme. Pattern Recognition, 2010,43(4):1656–1667. [doi: 10.1016/j.patcog.2009.09.016]

[12] An L, Gao X, Li X, Tao D, Deng C, Li J. Robust reversible watermarking via clustering and enhanced pixel-wise masking. IEEE Trans. on Image Processing, 2012,21(8):3598–3611. [doi: 10.1109/TIP.2012.2191564]

[13] Thabit R, Khoo BE. Capacity improved robust lossless image watermarking. IET Image Processing, 2014,8(11):662–670. [doi: 10.1049/iet-ipr.2013.0862]

[14] Feng DG, Zhang M, Zhang Y, Xu Z. Study on cloud computing security. Ruan Jian Xue Bao/Journal of Software, 2011,22(1):71–83 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/ SP.J.1001.2011.03958]

[15] Zhang XP. Reversible data hiding in encrypted image. IEEE Signal Processing Letters, 2011,18(4):255–258. [doi: 10.1109/LSP.2011.2114651]

[16] Zhang XP. Separable reversible data hiding in encrypted image. IEEE Trans. on Information Forensics and Security, 2012,7(2):826–832. [doi: 10.1109/TIFS.2011.2176120]

[17] Chen YC, Shiu CW, Horng G. Encrypted signal-based reversible data hiding with public key cryptosystem. Journal of Visual Communication and Image Representation, 2014,25:1164–1170. [doi: 10.1016/j.jvcir.2014.04.003]

[18] Zhang XP, Long J, Wang Z, Cheng H. Lossless and reversible data hiding in encrypted images with public key cryptography. IEEE Trans. on Circuits and Systems for Video Technology, 2016,26(9):1622–1631. [doi: 10.1109/TCSVT.2015.2433194]

[19] Xiang SJ, Luo XR, Shi SX. A novel reversible image watermarking algorithm in homomorphic encrypted domain. Chinese Journal of Computers, 2016,39(3):571–581 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2016.00571]

[20] Xiang SJ, Luo XR. Reversible data hiding in encrypted image based on homomorphic public key cryptosystem. Ruan Jian Xue Bao/Journal of Software, 2016,27(6):1592–1601 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5007.htm [doi: 10.13328/j.cnki.jos.005007]

[21] Paillier P. Public-Key cryptosystems based on composite degree residuosity classes. In: Proc. of the Int'l Conf. on the Theory and Application of Cryptographic Techniques Prague. 1999. 233–238. [doi: 10.1007/3-540-48910-X_16]

[22] Zheng PJ, Huang JW. Discrete wavelet transform and data expansion reduction in homomorphic encrypted domain. IEEE Trans. on Image Processing, 2013,22(6):2455–2468. [doi: 10.1109/TIP.2013.2253474]

[23] Donald K. The Art of Computer Programming, Volume 2. 3rd ed., Addison-Wesley, 1997. 325–515.

[24] Xiang SJ, Yang L, Wang Y. Robust and reversible audio watermarking by modifying statistical features in time domain. Advances in Multimedia, 2017,2017(3):1–10. [doi: 10.1155/2017/8492672]

**附中文参考文献**:

[14] 冯登国,张敏,张研,徐震.云计算安全研究.软件学报,2011,22(1):71−83. http://www.jos.org.cn/1000-9825/3958.htm [doi: 10.3724/SP.J.1001.2011.03958]

[19] 项世军,罗欣荣,石书协.一种同态加密域图像可逆水印算法.计算机学报,2016,39(3):571−581. [doi: 10.11897/SP.J.1016.2016.00571]

[20] 项世军,罗欣荣.基于同态公钥加密系统的图像可逆信息隐藏算法.软件学报,2016,27(6):1592−1601. http://www.jos.org.cn/1000-9825/5007.htm [doi: 10.13328/j.cnki.jos.005007]

**项世军**(1974－),男,贵州普定人,博士,教授, CCF 高级会员,主要研究领域为信息隐藏,加密域信号处理.

**杨乐**(1993－),男,硕士,主要研究领域为多媒体信息安全,加密域信号处理.