

适应性 JPEG 图像隐私保护方案

文泓鑫, 胡飞, 俞能海

(中国科学技术大学网络空间安全学院, 安徽 合肥 230026)

摘要: 由于 JPEG 图像格式是一种高压缩率图像格式, JPEG 图像中的隐私保护方案不仅需要保护隐私区域的内容, 同时要控制文件的大小。另外, JPEG 图像中包含隐私信息的区域, 即感兴趣区域 (ROI), 往往只占整个图像的一部分。仅对 ROI 进行保护有助于控制文件的大小。基于此, 提出了一种适应性 JPEG 图像隐私保护方案。利用人脸检测等方法从 JPEG 图像中提取隐私区域, 使用基于块的适应性可逆修改算法对隐私区域进行保护, 用可逆信息隐藏方法将复原所需的信息嵌入图像。上述隐私保护过程采用了可逆的算法, 合法的被授权者可以完全恢复经过所提方案隐私保护后的区域。由于引入了可变的参数, 可以选择合适的参数以适应不同的实际需求, 并达到最优效果。对多个 ROI 重叠的复杂情形进行了讨论并提出了解决方案。在峰值信噪比 (PSNR, peak signal-to-noise ratio) 和文件膨胀率等指标下, 通过实验探索了不同参数对方案性能的影响并找到了推荐的参数组合。实验结果表明, 在合适的参数下, 所提方案在隐私保护、文件膨胀率等方面中的表现优于已有方案。

关键词: JPEG 图像; 隐私保护; 可逆信息隐藏; 人脸识别

中图分类号: TP37

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024005

Adaptable privacy protection scheme in JPEG images

WEN Hongxin, HU Fei, YU Nenghai

School of Cyber Science and Technology, University of Science and Technology of China, Hefei 230026, China

Abstract: The JPEG format is widely used for its high compression rate in image storage. Privacy protection schemes applied to JPEG images not only need to safeguard the content of the privacy region but also maintain the file size. Furthermore, the region of interest (ROI), which contains private information, often occupies only a portion of the entire JPEG image. Protecting only the ROI can help preserve the file size. In light of this, an adaptable privacy protection scheme for JPEG images was proposed. Privacy regions were identified using methods like face detection within the JPEG images. These regions were then modified using the block-based adaptable reversible modification algorithm presented in this study. Reconstruction information was embedded into the image through

收稿日期: 2023-01-06; 修回日期: 2023-08-27

通信作者: 俞能海, ynh@ustc.edu.cn

基金项目: 国家重点研发计划 (2018YFB0804100)

Foundation Item: The National Key R&D Program of China (2018YFB0804100)

引用格式: 文泓鑫, 胡飞, 俞能海. 适应性 JPEG 图像隐私保护方案[J]. 网络与信息安全学报, 2024, 10(1): 91-101.

Citation Format: WEN H X, HU F, YU N H. Adaptable privacy protection scheme in JPEG images[J]. Chinese Journal of Network and Information Security, 2024, 10(1): 91-101.

reversible data hiding. The application of reversible algorithms in the privacy protection process enabled authorized users to fully restore the protected area. By introducing variable parameters, this scheme allowed for the selection of appropriate parameters to meet different practical requirements and achieve optimal encryption results. Additionally, the complex scenario of multiple overlapping ROIs was discussed and resolved. Experimental evaluations were conducted to investigate the influence of parameters on performance, considering both peak signal-to-noise ratio (PSNR) and file size increment. The results demonstrate that the proposed scheme outperforms existing schemes in terms of privacy protection and file size when suitable parameters are employed.

Keywords: JPEG image, privacy protection, reversible data hiding, face detection

0 引言

随着第五代移动通信技术的普及,云计算、云存储等新兴技术迅速发展,各种类型的信息媒体(如图片、视频等)在通信网络中迅速便捷地传输。然而,通信技术的发展在给人们的生活带来便利的同时,给网络空间安全带来了新的挑战。由于网络的纷繁复杂,私人隐私信息、商业机密消息,甚至国家重要情报在网络中传输时不能保证安全和可靠,这些数据在传输中可能会被恶意攻击者窃听、截获和篡改。保证隐私数据的安全,成为网络空间安全的重要课题。出于对网络空间安全的重视,我国陆续出台了《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等相关法律。

图像作为数据的一种形式,包含隐私信息,因此图像加密是信息安全的一个重要分支。与文本相比,图像具有直观性,更加受到人们的青睐。如今,图像已经是互联网上最流行的媒介之一。

由于图像具有更大的信息量,加密所产生的代价(即加密用时、计算量、传输带宽、存储空间等)相应也更大。另外,并不是整个图像是需要保护的敏感信息,图像中只有部分区域包含隐私信息,只有这些隐私区域需要保护,这些区域被称为感兴趣区域(ROI, region of interest)。为了节省不必要的开支,需要对图像进行细粒度的加密,只对选出的图像隐私区域,即 ROI 进行加密保护。例如,用户在社交平台(如微博、微信朋友圈)上发布图片,图片中包含的隐私信息就有可能因此泄露,对包含隐私信息的 ROI 进行加密保护可以有效解

决这个问题。

JPEG 图像格式^[1]是较流行的图像格式之一,互联网上的绝大部分图片是以 JPEG 格式传输与存储的,因此针对 JPEG 图像的 ROI 保护具有重大意义。

为了减少传输和存储成本,隐私保护方法要具有可逆性。被授权用户除了少数必要信息,如验证身份的密钥,不需要额外信息就能从被保护的 JPEG 文件中恢复原始图像,从而不需要保留原始图片与被加密处理后图片的多个版本,降低传输与存储成本。为了实现可逆性,现有的图像隐私保护方案将重构信息用可逆信息隐藏(RDH, reversible data hiding)嵌入处理后的图像中。然而,由于 JPEG 格式的高压缩率,几乎没有冗余信息,难以嵌入大量额外信息,否则会严重失真。

一个典型的 JPEG 图像隐私保护场景如图 1 所示。图像所有者对 JPEG 图像进行处理后发布到互联网上。被授权者拥有获取 ROI 的权限,可以从处理后的 JPEG 文件中恢复出 ROI,进而得到完整的原始图像。而非授权者没有获取 ROI 的权限,不能恢复 ROI,仅能获取处理后的 JPEG 图像。

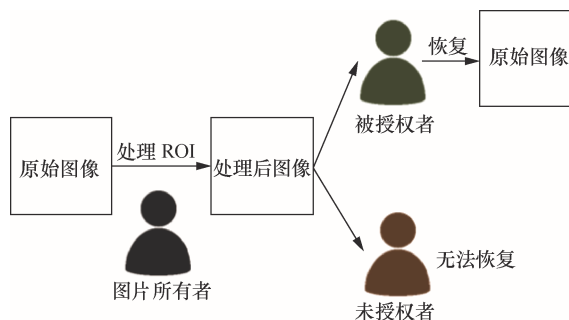


图 1 一个典型的 JPEG 图像隐私保护场景
Figure 1 A typical scene of privacy protection in JPEG image

理想的 JPEG 图像隐私保护方案应该满足以下指标: 第一, 保护隐私能力, 即被保护的区域难以辨认, 尽量不暴露任何信息; 第二, 可逆性, 即原始图像能被完全恢复; 第三, 安全性, 即只有被授权用户具有恢复原始图像的能力, 恶意用户使用任何攻击方式都无法得到隐私信息; 第四, 低文件膨胀率, 即被保护后的 JPEG 文件的大小不会显著变大, 以免占用过多传输带宽和存储空间。其中, 第 4 个指标是与传统隐私保护领域相比 JPEG 图像隐私保护问题独有的特征。

1 相关工作

基于 ROI 的图像保护方案的设计思路一般分为两部分: 一是对 ROI 的提取与加密; 二是将重构信息嵌入加密保护后的图像中。基于这一思路, 学术界提出了很多相关方法。

Niimi 等^[2]改造了一个基于 bmp (bitmap, 位图) 图像的可逆信息隐藏方法, 使之可以运用在 JPEG 图像格式中: 将 JPEG 图像文件的可交换图像文件格式 (Exif, exchangeable image file format) 段中的隐私信息作为秘密消息, 嵌入 ROI 对应的量化离散余弦变换 (DCT, discrete cosine transform) 系数中; 选取合适的参数使得失真尽可能大, 以此来保护 ROI。但是该方法产生的失真不足以保护隐私, 从视觉上仍能辨认出隐私信息的大致轮廓。

Yuan 等^[3]先用任意方法加密保护 ROI, 然后计算加密保护后图像与原图像的差值得到一个子图像, 最后将子图像和元数据 (即子图像数据的大小、加密方法、子图像位置等信息) 加入 JPEG 图像文件的 APPn 段中。但是由于获取子图像时只记录了差值超过阈值的部分, 差值低于阈值的部分不被记录, 这个方法不能完全重构原始图像。Li 等^[4]通过引入稀疏表示改进了上述方法。在获得子图像后, 计算子图像的亮度通道的稀疏表示, 通过稀疏表示计算重构子图像, 再计算重构子图像与原始子图像的残差, 将稀疏表示、残差、色度通道以及元数据等嵌入图像。引入了稀疏表示方法, 大大减少了嵌入数据, 从而降低文件膨胀率, 但是由于使用了不可逆的数据隐藏方法, 这

个方法不完全可逆。

Lin 等^[5]提出对隐私区域的量化 DCT 系数进行修改来保护 JPEG 图片隐私的方法。根据密钥控制的随机数流改变 DCT 系数的符号, 然后将元数据加入 JPEG 图像文件的 APPn 段中。

Cao 等^[6]对 JPEG 图像的亮度通道进行修改, 将 ROI 的量化 DCT 系数的直流分量修改成 ROI 的直流分量的平均值, 并将修改量编码为 9 bit 的数据, 隐藏到 zigzag 顺序的第 3 到第 11 个交流分量中, 再利用可逆信息隐藏方法将元数据嵌入保护后的图像中。Huang 等^[7]对上述方法进行了改进, 修改了包括亮度和色度在内的全部 3 个通道, 并且修改亮度通道的前 6 个量化 DCT 系数, 实现更强的安全性。但是该方法修改量较大, 有可能做了多余的修改, 提高了性能的空间。

以往工作并未考虑多 ROI 的情形, 或者简单地假设多个 ROI 不会重叠。在实际应用中, ROI 重叠的情形并不少见, 因此如何处理多 ROI 重叠是个值得探讨的问题。

本文的主要贡献如下。

1) 引入了可变的参数, 在不同的实际应用中可以根据具体的需求选择合适的参数, 以达到更好的性能。

2) 对多 ROI 的情形进行了讨论并提出了解决方案。

3) 通过实验探索了不同参数对算法性能的影响, 并找到了推荐的参数配置。

2 本文方案

适应性 JPEG 图像隐私保护方案分为两个模块: 图像加密保护和图像恢复。图像所有者根据需要对图像中的 ROI 执行图像加密保护模块; 被授权者对接收到的处理后图像执行图像恢复模块, 得到原始图像。

2.1 感兴趣区域

一般来说, 图像中包含隐私信息、需要被保护的区域均可认为是 ROI。ROI 的大小、形状根据实际情况并不唯一。简单起见, 假设 ROI 为任意大小的矩形, 隐私信息为图像中的人脸, 如无特殊说明, 图像中的 ROI 仅有一个。

2.2 图像加密保护

加密保护的流程如图 2 所示, 分为 ROI 的获取与表示、ROI 的加密保护和元数据的嵌入 3 部分。

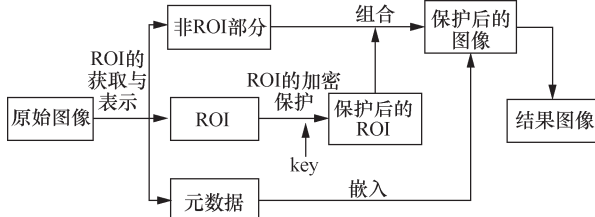


图 2 加密保护的流程
Figure 2 Procedure of encryption

2.2.1 ROI 的获取与表示

本文中的 ROI 为图像中的人脸。ROI 的获取可以由图像所有者手动选取, 也可以借助人脸检测等辅助软件实现。假设 JPEG 图像中包含 N 个 8×8 的块, 分别编号为 1 到 N 。获得 ROI 后, 记录左上和右下两个块的编号, 分别记为 B_1 和 B_2 , 然后将 B_1 和 B_2 分别转换为二进制串 b_1 和 b_2 。

$$b_1 = \text{dec2bin}(B_1, L) \quad (1)$$

$$b_2 = \text{dec2bin}(B_2, L) \quad (2)$$

其中, $\text{dec2bin}(\cdot)$ 表示将数字转换为二进制串。 b_1 和 b_2 的位数 L 满足

$$L = \text{ceil}(\text{lb}N) \quad (3)$$

其中, $\text{ceil}(\cdot)$ 表示向上取整。

2.2.2 ROI 的加密保护

本文方案对 JPEG 图像的包括亮度和色度在内的全部 3 个通道进行加密保护。对每个通道, 先对 ROI 中的每个块进行块操作, 再对 ROI 中的块进行置乱。块操作的流程如图 3 所示。块操作使用 3 个参数: 修改数 α 、修改量位数 β 和嵌入位置 γ 。块操作的具体细节如下。

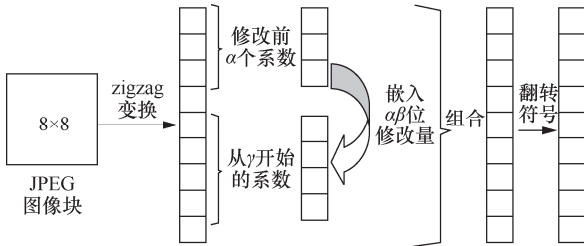


图 3 块操作的流程
Figure 3 Procedure of block-operation

假设 ROI 中有 n 个块, 对每个块 B_i 进行 zigzag 变换后得到的重新排列的量化 DCT 系数为 $\{d_1^i, d_2^i, \dots, d_{64}^i\}$ 。计算前 α 个系数的平均值。

$$\text{avg}_j = \frac{1}{n} \sum_{B_i \in \text{ROI}} d_j^i, j=1, 2, \dots, \alpha \quad (4)$$

将前 α 个系数修改为平均值, 计算修改量 delta 。

$$\text{delta} = d_j - \text{avg}_j \quad (5)$$

由于量化 DCT 系数的取值是 $[-1023, 1023]$, delta 的取值是 $[-2046, 2046]$ 。为了能用 β bit 表示修改量, 通过以下方式对 delta 进行修整。

$$\Delta d_j = \begin{cases} 2 \cdot u \cdot \text{round}\left(\frac{|\text{delta}|}{2u}\right), & \text{delta} \geq 0 \\ 2 \cdot u \cdot \text{floor}\left(\frac{|\text{delta}|}{2u}\right) + u, & \text{delta} < 0 \end{cases} \quad (6)$$

其中, $\text{round}(\cdot)$ 表示四舍五入, $\text{floor}(\cdot)$ 表示向下取整。 u 的值由式(7)决定。

$$u = 2^{11-\beta} \quad (7)$$

令 $D_j = \frac{|\Delta d_j|}{u}$, 于是 D_j 可以表示为 β bit 的比特串

$$\text{str}_j = \text{dec2bin}(D_j, \beta) \quad (8)$$

第 j 个系数修改为

$$d_j' = d_j - \Delta d_j, j=1, 2, \dots, \alpha \quad (9)$$

此时前 α 个系数被修改, 并得到 $\alpha\beta$ bit 的修改量信息 message 。 $\alpha\beta$ bit 的修改量信息 message 接下来被嵌入第 $\gamma+1$ 到第 $\gamma+\alpha\beta$ 个系数中。

$$d_{\gamma+j}' = 2d_j + \text{message}_j, j = \gamma+1, \gamma+2, \dots, \gamma+\alpha\beta \quad (10)$$

其中, message_j 为 message 的第 j 个比特, 即被嵌入的比特。

出于安全性考虑, 根据密钥 key 生成一个二元随机序列, 再根据该随机序列翻转交流系数, 即第 2 到第 64 个系数的符号。

当 ROI 中所有块的块操作完成后, 再在 key 的控制下将块置乱, 进一步增强安全性。

2.2.3 元数据的嵌入

元数据包括 ROI 的位置信息等恢复图像所必需的信息。使用基于直方图平移的可逆信息隐藏方法将元数据 *metadata* 嵌入处理后的图像。

$$a' = \begin{cases} 0, & a = 0 \\ a + \text{metadata}_i, & a = 1 \\ a - \text{metadata}_i, & a = -1 \\ a + 1, & a > 1 \\ a - 1, & a < -1 \end{cases} \quad (11)$$

其中, a 表示未修改的交流系数, metadata_i 表示元数据 *metadata* 待嵌入的比特, a' 表示修改后的交流系数。

2.3 图像恢复

图像恢复的流程如图 4 所示, 分为元数据的提取与 ROI 的定位和 ROI 的恢复。除了被加密保护的 JPEG 图像, 被授权者还需要参数 α 、 β 、 γ 以及密钥 *key* 来恢复原始图像。

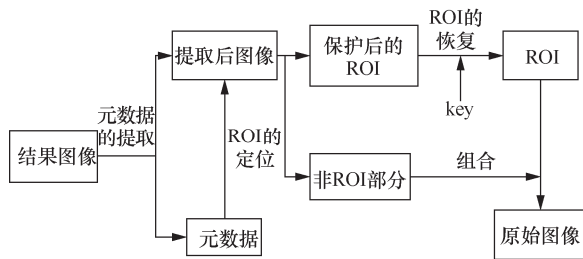


图 4 图像恢复的流程

Figure 4 Procedure of image restoring

2.3.1 元数据的提取与 ROI 的定位

根据对应的可逆信息隐藏方法可以提取出嵌入图像的元数据 *metadata*, 具体过程如下。

$$\text{metadata}_i = \begin{cases} 0, & a' = 1 \text{ 或 } a' = -1 \\ 1, & a' = 2 \text{ 或 } a' = -2 \end{cases} \quad (12)$$

$$a = \begin{cases} 0, & a' = 0 \\ 1, & a' = 1 \text{ 或 } a' = 2 \\ -1, & a' = -1 \text{ 或 } a' = -2 \\ a' - 1, & a' > 2 \\ a' + 1, & a' < -2 \end{cases} \quad (13)$$

由于图像的大小没有改变, 块数 N 不变, 被授权者可以根据式(3)计算出 L , 进而从元数据 *metadata* 中提取 ROI 的位置。

2.3.2 ROI 的恢复

被授权者利用密钥 *key* 恢复被置乱的 ROI 中的块, 然后用加密过程中的方法生成相同的二进制随机序列恢复交流系数的符号。

对 ROI 中的每个块, 首先进行 zigzag 变换, 然后从第 $\gamma+1$ 到第 $\gamma+\alpha\beta$ 个系数中提取修改量的信息 *message*, 并恢复这些系数的值。

$$\text{message} = d'_{\gamma+j} \bmod 2 \quad (14)$$

$$d_{r+j} = \text{floor}(d'_{\gamma+j} / 2) \quad (15)$$

其中, $j = 1, 2, \dots, \alpha\beta$ 。

从修改量信息 *message* 中提取前 α 个系数对应的 β 位二进制串 *str_j*, 再转换为 D_j 。

$$D_j = \text{bin2dec}(\text{str}_j) \quad (16)$$

其中, $\text{bin2dec}(\cdot)$ 表示将二进制串转换为数字。

修改量的值由下式得到。

$$\Delta d_j = \begin{cases} u \cdot D_j, & D_j \bmod 2 = 0 \\ -u \cdot D_j, & D_j \bmod 2 = 1 \end{cases} \quad (17)$$

然后恢复前 α 个系数。

$$d_j = d'_j + \Delta d_j, j = 1, 2, \dots, \alpha \quad (18)$$

将恢复的 ROI 替换图像中对应的位置, 即完成了图像的恢复。

2.4 多 ROI 的情形

当图像中的 ROI 不止一个时, 就会产生新的、更复杂的问题。首先是被授权者只有一个的较简单情形。当 ROI 重叠时, 被授权者恢复图像时如果简单地按照加密的顺序恢复 ROI, 是无法正确恢复的。正确的做法是按照与加密顺序相反的顺序恢复 ROI。以上讨论同样适用于只有一类被授权者, 即被授权者的权限均相同的情形。

更复杂的情形是被授权者不止一个, 多 ROI 情形下的多被授权者问题如图 5(a)所示, 每个被授权者只有解密部分 ROI 的权限, 而非所有的 ROI。同时, ROI 之间会有重叠部分。如果两个 ROI 分别属于两个被授权者, 并且每个被授权者都没有解密另一个 ROI 的权限, 当两个 ROI 重叠时, 两个被授权者都应有权限解密重叠部分, 因为重叠部分属于他们有权限解密的范围。但如果按顺序加密两个 ROI, 两个被授权者不

仅不能解密重叠部分, 有权限解密的非重叠部分也可能无法完全恢复。

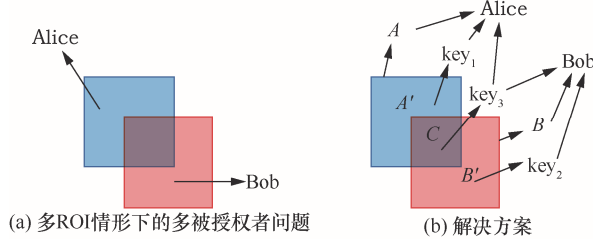


图 5 多 ROI 的情形
Figure 5 Multi-ROI situation

在这种情形下, 应将 ROI 分为互不重叠的部分, 每个部分用不同的密钥加密, 再把密钥分发给对应的被授权者。如图 5(b)所示, Alice 和 Bob 分别有 ROI A 和 ROI B 的权限, 而 A 与 B 存在一处重叠区域 C 。将 ROI 划分为 $A' = A - C$, $B' = B - C$ 以及 C 共 3 部分, 分别用密钥 key_1 、 key_2 和 key_3 加密, 将 key_1 、 key_3 分发给 Alice, 将 key_2 、 key_3 分发给 Bob。这样 Alice 就能恢复 A' 和 C , 从而恢复 A ; 同理 Bob 能恢复 B 。多 ROI 情形的密钥分发如表 1 所示, 列出了本例中各个被授权者对应的权限以及拥有的密钥。

表 1 多 ROI 情形的密钥分发
Table 1 Key distribution in multi-ROI situation

被授权者	Alice	Bob
权限	$A = A' + C$	$B = B' + C$
密钥	key_1, key_3	key_2, key_3

3 实验与分析

3.1 实验准备及对参数的讨论

本文使用的实验数据集是来自 Image of Groups 数据集^[8]的 200 张图片, ROI 的获取步骤使用了来自文献[9]的人脸检测工具。本文在 Matlab 中完成了所提方案的代码实现。

实验采用两个数值指标分别衡量对 ROI 的保护效果和文件大小的增加程度: 峰值信噪比 (PSNR, peak signal-to-noise ratio) 和文件膨胀率 (FSI, file size increment)。PSNR 和 FSI 的定义如下。

$$PSNR = 10 \lg \frac{255^2}{MSE} \quad (19)$$

$$MSE = \frac{1}{|I|} \sum_{(i,j)} [I(i,j) - I'(i,j)]^2 \quad (20)$$

$$FSI = \frac{Size(F') - Size(F)}{Size(F)} \times 100\% \quad (21)$$

其中, I 表示原始图像, I' 表示处理后图像, F 表示原始文件, F' 表示处理后文件。由于 ROI 在完整图像中所占的比例并不固定, 为了消除干扰, 仅计算处理前后的 ROI 的 PSNR。

由于二维 DCT 的对称性, α 的取值为 1、3、6、10... 又由于修改量嵌入容量有限, 过大的 α 应舍去, 因此 α 的取值仅有 1、3、6。由于人眼对色度的敏感性相对较低, 色度通道的重要性较低, 可以选择修改量较少的 $\alpha = 1$ 。

由于 $\beta = 1, 2$ 时, u 的值较大, 容易导致修改量过大, 进而造成溢出, 因此有

$$\beta \geq 3 \quad (22)$$

并且需满足

$$\beta \leq 11 \quad (23)$$

而 γ 需要满足的条件有

$$\gamma \geq \alpha \quad (24)$$

$$\gamma + \alpha\beta \leq 64 \quad (25)$$

3.2 不同参数组合对算法性能的影响

为了探索不同参数组合对算法性能的影响, 本节对所有的参数组合进行了实验。根据 3.1 节中对参数取值的讨论, 可能的参数组合数为

$$\sum_{\alpha \in \{1, 3, 6\}} \sum_{\beta=3}^{11} \max(64 - \alpha\beta - \alpha + 1, 0) = 1043 \quad (26)$$

不同参数下 PSNR 的变化曲线如图 6 所示, 不同参数下 FSI 的变化曲线如图 7 所示。

从图 6 可以看到, PSNR 的最大值仅有 12 dB 左右, 说明图像的质量很差, 足以有效保护 ROI。不同 α 值的 PSNR 曲面基本重合, 说明不同的 α 对 PSNR 的影响不大。

从图 7 中可以看到, α 对 FSI 有显著影响。 $\alpha = 6$ 时, FSI 最大, $\alpha = 3$ 时, 次之, $\alpha = 1$ 时, FSI 最小。其原因是 α 越大, 嵌入的数据量 $\alpha\beta$ 越大, 从而 FSI 越大。

为了进一步探索各参数对算法性能的影响, 在二维坐标系中重新绘制了实验结果。从上述讨

论可知, α 对 PSNR 几乎没有影响, 并且 α 越小 FSI 越小。为了找到使 FSI 最小的参数组合, 只对 $\alpha=1$ 的情形进行讨论。

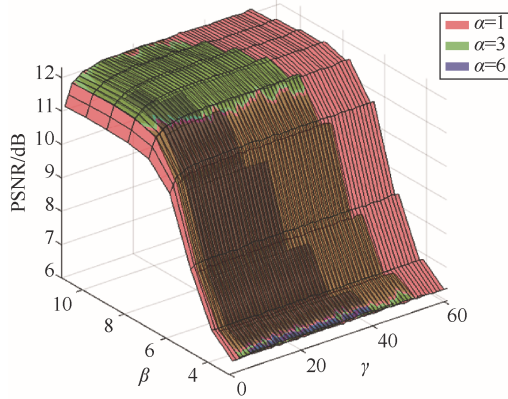


图6 不同参数下 PSNR 的变化曲线
Figure 6 PSNR curves in different parameters

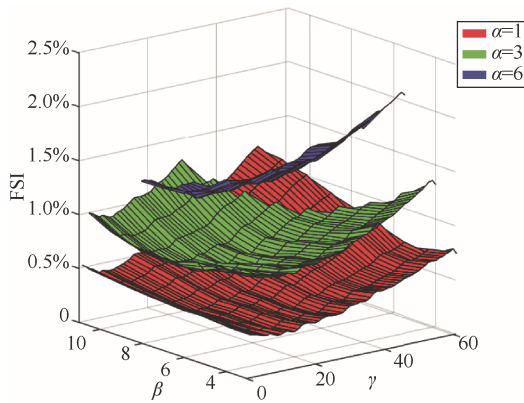


图7 不同参数下 FSI 的变化曲线
Figure 7 FSI curves in different parameters

图8展示了 $\alpha=1$ 时, PSNR 的变化曲线。可以看到, γ 对 PSNR 的影响不大。当 $\beta=3,4,5$ 时, γ -PSNR 曲线几乎为与横轴平行的直线; 而当 $\beta>5$ 时, γ -PSNR 曲线在 $\gamma\leq 3$ 时快速上升, 并在 $\gamma=6$ 之后趋于稳定。造成这种现象可能是由于二维 DCT 的对称性, 第2到第3个系数、第4到第6个系数等具有相似的频率和重要性。其中第2和第3个系数具有仅次于第1个系数(即直流系数)的重要性, 因此 $\gamma\leq 3$ 时, 对这两个系数的修改会造成较严重的失真。而第6个之后的系数的重要性差距相对不大, 因此 $\gamma>6$ 时, PSNR 变化不明显。

而 β 对 PSNR 影响的规律为: β 越大, PSNR 越大, 当 $\beta>7$ 时, PSNR 趋于稳定。造成该现象

的原因: β 越小, u 越大, 对修改量 delta 的修整就越粗糙, 容易产生剧烈的修改, 严重影响图像质量。而当 β 足够大时, u 足够小, 修整变得精细, 修整后的修改量 Δd_j 趋于稳定, 不再随 β 产生可观的变化, 此时 PSNR 趋于稳定。

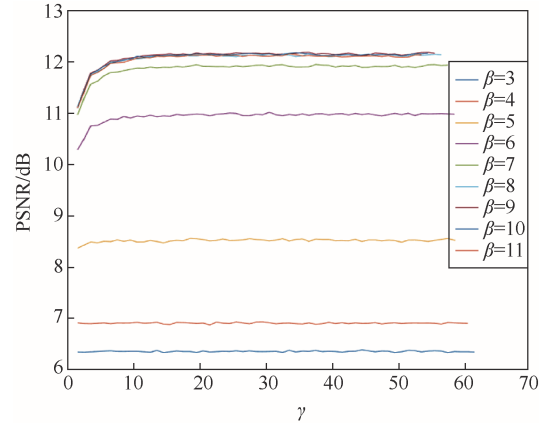


图8 $\alpha=1$ 时, PSNR 的变化曲线
Figure 8 PSNR curves when $\alpha=1$

图9展示了 $\alpha=1$ 时, FSI 的变化曲线。可以看到, FSI 并不随 γ 单调变化, 而随着 γ 的增加先降低后升高。产生这种现象的原因可能是: 当在较前的系数嵌入时, 较前的系数有更大的概率具有较大的值, 嵌入过程中系数值会翻倍变得更大, 从而熵编码时会使用更长的编码, 导致文件变大; 当在较后的系数嵌入时, 较后的系数通常为0, 若嵌入的比特为1则会变为1, 即非0值, 导致连续的0被打断, 在行程长度编码时增加了编码的数量, 进而增加了文件的大小。因此, γ 的取值应该适中, 既不宜过大也不宜过小。另外, 在 γ 的最大取值处, FSI 会急剧下降可能是由于对第64个系数(即最后一个系数)进行了修改, 当嵌入的比特为1时, 该系数变为非0值, 在行程长度编码时减少了EOB的编码, 因此 FSI 显著降低。

β 对 FSI 影响的大致规律为: β 越大, 则 FSI 越大。显然 β 越大时, 嵌入的数据量 $\alpha\beta$ 比特越多, 从而 FSI 越大。而某些地方违反了这一规律, 如 $\beta=3$ 的曲线并没有低于其他曲线。其解释为 β 越小, u 越大, 对修改量 delta 的修整越粗糙, 容易产生剧烈的修改, 使得修改后的系数值显著变大, 增加编码长度, 导致 FSI 增大。

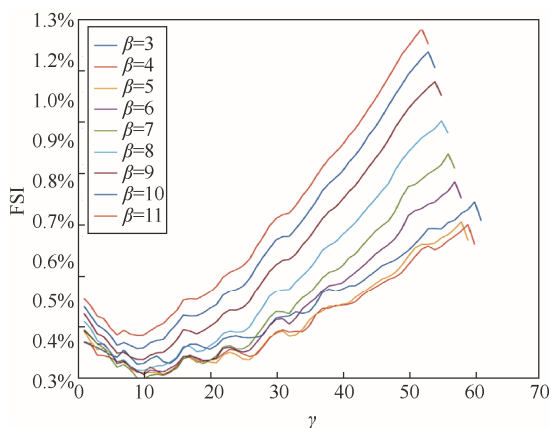


图 9 $\alpha=1$ 时, FSI 的变化曲线
Figure 9 FSI curves when $\alpha=1$

这些曲线的最低点在 $\beta=7$, $\gamma=10$ 时取得, 此时其 FSI 为 0.30%。

3.3 在不同 QF 下的表现

JPEG 是一种有损压缩编码标准, 通过不同的质量因子 (QF, quality factor) 来调整图像质量和压缩率。上述实验中使用的测试数据图像为 QF = 75 的 JPEG 图像。为了测试本文方案在不同 QF 下的表现, 分别对 QF = 60, 70, 80, 90, 100 的测试图像进行了实验。根据 3.2 节, 如无特殊说明, 之后的实验参数选择均为 $(\alpha, \beta, \gamma) = (1, 7, 10)$ 。

不同 QF 下的一组实验结果如图 10 和图 11 所示。

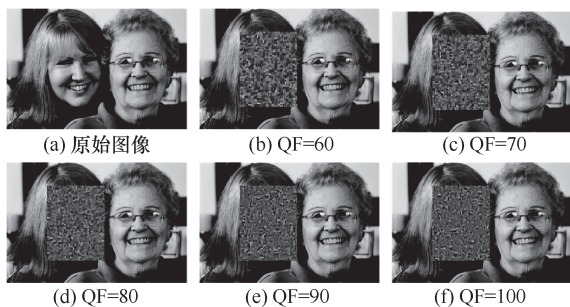


图 10 不同 QF 下的一组实验结果 (黑白)
Figure 10 Experiment results in different QF (monochrome images)

图 10 展示了本文方案对于黑白图像在不同 QF 下的表现。可以看到, 本文方案在不同 QF 下都不泄露 ROI 信息, 能较好地实现隐私保护的功能, 并且 QF 越大, 被保护的区域越平滑, 突变越少。其原因是: QF 越小, 量化步长越大, 像素值对量化 DCT 系数的变化就越敏感, 最终导致相近的量化 DCT 系数对应的图像有较大的

差距。

图 11 展示了本文方案对于彩色图像在不同 QF 下的表现。可以看到, 当 QF 较小时, 被保护的区域和黑白图像的情形一样突变较明显, 并且颜色偏品红色; 而当 QF 较大时, 颜色恢复为正常颜色的平均值。出现颜色方面的这种现象可能是由于当 QF 较小时, 色彩通道的量化步长较大, 色彩通道的色彩空间较小, 色彩表达能力有限, 加上本文算法为了减少修改量位数对修改量进行修整, 减小了修改量的取值空间, 容易产生颜色的偏差。如果想要修正这种偏差, 可以使用更大的 QF, 或者在色彩通道使用更大的参数 β 对修改量进行更精细的修整。

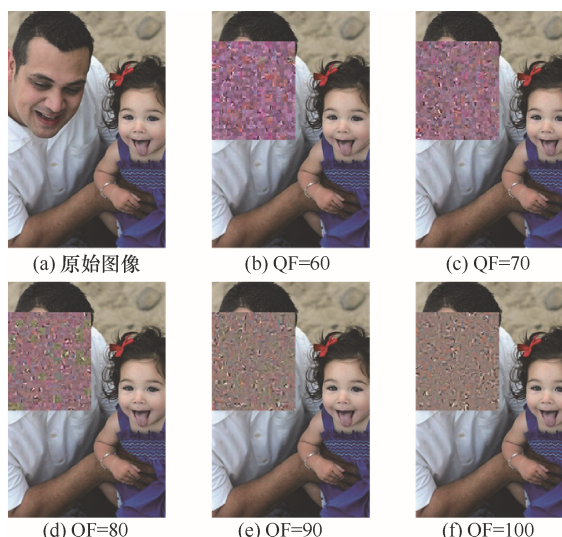


图 11 不同 QF 下的一组实验结果 (彩色)
Figure 11 Experiment results in different QF (color images)

图 12 展示了不同 QF 下的 PSNR。可以看到, PSNR 随 QF 增大而增大, 快速上升后趋于平缓。因为小的 QF 会导致粗糙的修改, 造成更严重的失真; 而当 QF 足够大时, 更精细的修改使系数更接近修改的目标 (即平均值), 此时 PSNR 变得稳定。

图 13 展示了不同 QF 下的 FSI。可以看到, FSI 随 QF 单调下降。这是因为 QF 越大, 量化 DCT 系数越大, 编码越长, 原文件越大, 相应地, 文件膨胀越不明显。

3.4 不同 ROI 获取方法下的表现

本文假设 ROI 为图像中包含隐私人脸的区域, 实验中使用人脸检测程序获取图像中的 ROI。

事实上, ROI 可以根据实际需求采用其他的定义, 获取 ROI 的方法多种多样。

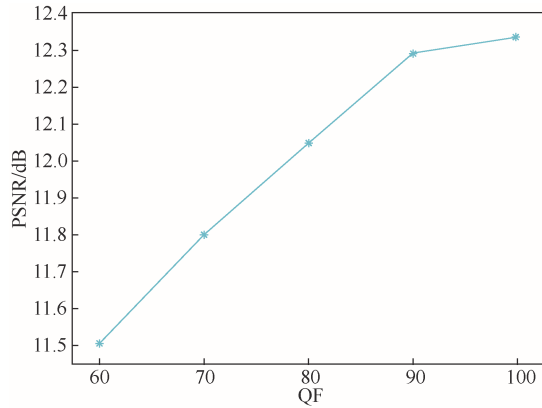


图 12 不同 QF 下的 PSNR
Figure 12 PSNR in different QF

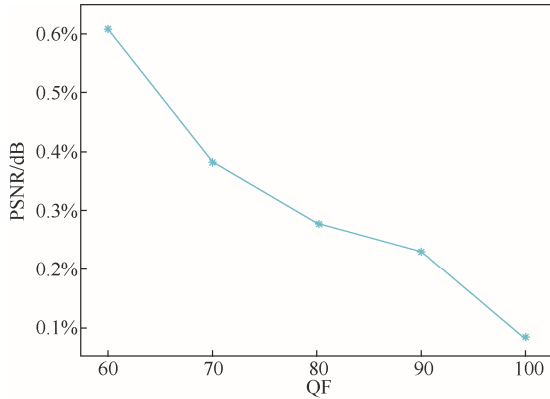


图 13 不同 QF 下的 FSI
Figure 13 FSI in different QF

图 14 展示了不同 ROI 获取方法的差异。图 14(a)为需要保护的原始图像, 图 14(b)、图 14(c)分别为使用人脸检测程序和人工标记获取 ROI 的保护后图像, 图 14(d)、图 14(e)分别为图 14(b)、图 14(c)的复原结果。可以看出, 不同的 ROI 获取方法仅仅对 ROI 的判定区域存在差异, 对保护效果并无显著影响, 都能很好地实现隐私保护和图像的无损恢复。



图 14 不同 ROI 获取方法的差异
Figure 14 Differences with different ROI obtaining methods

图 15 展示了在其他类型隐私的情形使用人工标记方法获取 ROI 的例子。可以看到, 对于非人脸的隐私, 本文方案能有效地保护。

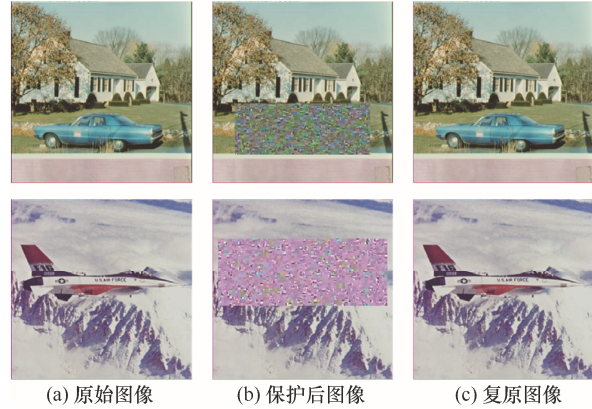


图 15 使用人工标记方法获取 ROI 的例子
Figure 15 Examples of obtaining ROI using manual marking method

3.5 与已有方案的对比

图 16 展示了与已有方案的一组实验结果对比。为了证明本文方案的稳定性, 展示了 3 组不同的参数对同一图片进行保护的效果。3 组参数 (α, β, γ) 的取值分别为 $(1, 7, 10)$ 、 $(1, 9, 10)$ 和 $(6, 9, 10)$ 。可以看到, 当 $(\alpha, \beta, \gamma) = (1, 7, 10)$ 时, 被保护区域颜色为偏品红; 当将 β 增加到 9 后, 被保护区域的颜色变为肉色, 即合理的平均色; 对于修改最精细的参数组合 $(\alpha, \beta, \gamma) = (6, 9, 10)$, 被保护区域更加平滑细腻。由本文方案处理后的 ROI 均难以辨认, 足以被有效保护。而 Cao^[6]的方案由于没有处理色度通道, 人脸的颜色轮廓较为明显。在 Huang^[7]的方案结果中, 具有复杂纹理的区域, 如人脸的眼口鼻部分, 能被辨认出相对位置, 具有泄露信息的风险。上述结果表明, 本文方案比已有方案更能有效保护 ROI。

本文方案实验得到的 FSI 与已有方案的 FSI 对比如表 2 所示。可以看到, 本文方案在参数设置 $(\alpha, \beta, \gamma) = (1, 7, 10)$ 下的 FSI 仅为 0.30%, 低于已有方案。而当参数 $(\alpha, \beta, \gamma) = (1, 9, 10)$ 时, FSI 为 0.35%, 依然优于已有方案。当参数 $(\alpha, \beta, \gamma) = (6, 9, 10)$ 时, FSI 为 1.33%。综上所述, 在合适的参数选择下, 本文方案具有比已有方案更优的性能。

3.6 安全性分析

根据文献[6]中的实验, 大部分 ROI 包含 25

到 400 个块。另外,文献[10]中的实验显示,平均每个块中有 26 个非 0 的交流系数。因此,如果一个攻击者试图使用穷举攻击恢复亮度通道,搜索空间大约为 $2^{26 \times 25} \times 25! \approx 2^{734}$,如果攻击者试图进一步恢复色度通道,则搜索空间进一步增大,因此,本文方案足以抵抗穷举攻击。

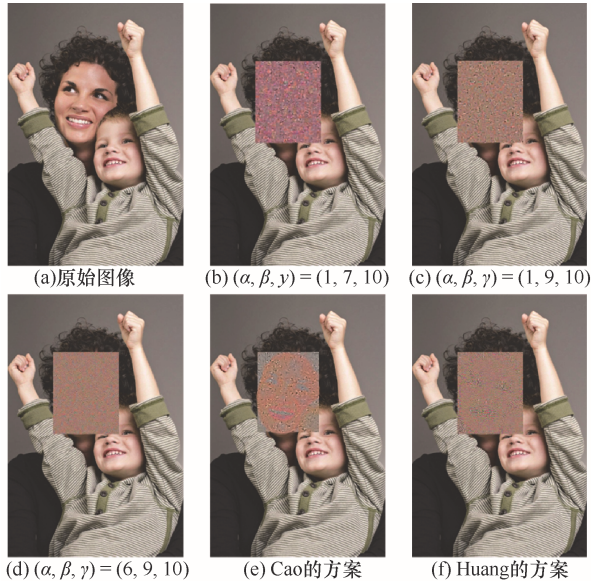


图 16 与已有方案的一组实验结果对比
Figure 16 Comparison with existing schemes

表 2 本文方案与已有方案的 FSI 对比
Table 2 Comparison of FSI between the proposed scheme and existing schemes

方案	FSI
本文方案, 参数 $(\alpha, \beta, \gamma) = (1, 7, 10)$	0.30%
本文方案, 参数 $(\alpha, \beta, \gamma) = (1, 9, 10)$	0.35%
本文方案, 参数 $(\alpha, \beta, \gamma) = (6, 9, 10)$	1.33%
Yuan ^[3] 的方案	10.80%
Li ^[4] 的方案	0.51%
Cao ^[6] 的方案	0.39%
Huang ^[7] 的方案	1.15%

图像修复方法^[11]通过神经网络推理预测可以对图像缺失部分填补修复。基于图像修复的攻击将被保护的 ROI 视为图像的缺失部分,试图通过图像修复方法预测修复 ROI,进而得到其中包含的隐私信息。本文方案可以阻止图像修复算法从 ROI 中获取与隐私相关的有效信息,图像修复基于训练数据的预测人脸看起来合理,但与原本的隐私人脸毫无关联。因此本

文方案可以抵抗基于图像修复的攻击。

4 结束语

本文针对 JPEG 图像中 ROI 保护的问题,提出了适应性 JPEG 图像隐私保护方案,通过引入修改数、修改量位数、嵌入位置等参数,可以灵活控制算法的 ROI 保护效果、文件膨胀率等性能。对于多 ROI 的复杂情况,本文进行了相应的讨论,提出了对应的解决方案,通过实验探索了参数对性能影响的规律,并且实验表明,在合适的参数下,本文方案可以实现比已有方案更优的性能。

由于本文方案中使用的可逆信息隐藏方法较为朴素,在未来的工作中存在改进的空间。例如,在元数据的嵌入中使用的可逆信息隐藏方法可以改用更先进的方法,实现更低的非 ROI 的图像失真以及文件膨胀率。

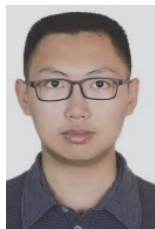
参考文献:

- [1] WALLACE G K. The JPEG still picture compression standard[J]. IEEE Transactions on Consumer Electronics, 1992, 38(1).
- [2] NIIMI M, MASUTANI F, NODA H. Protection of privacy in JPEG files using reversible information hiding[C]//Proceedings of 2012 International Symposium on Intelligent Signal Processing and Communications Systems. 2012: 441-446.
- [3] YUAN L, EBRAHIMI T. Image trans morphing with JPEG[C]//Proceedings of 2015 IEEE International Conference on Image Processing(ICIP). 2015: 3956-3960.
- [4] LI W, NI R, ZHAO Y. JPEG photo privacy-preserving algorithm based on sparse representation and data hiding[C]//Proceedings of International Conference on Image and Graphics. 2017: 575-586.
- [5] LIN Y, KORSHUNOV P, EBRAHIMI T. Secure JPEG scrambling enabling privacy in photo sharing[C]//Proceedings of 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, 2015, (4): 1-6.
- [6] CAO X, HUANG Y, WU H, et al. Content and privacy protection in JPEG images by reversible visual transformation[J]. Applied Sciences, 2020, 10(19).
- [7] HUANG Y, CAO X, WU H T, et al. Reversible data hiding in jpeg images for privacy protection[C]//Proceedings of 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2021: 2715-2719.
- [8] GALLAGHER A, CHEN T. Understanding groups of images of people[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2009.
- [9] YU S. Lib face detection[EB].
- [10] RUCHAUD N, DUGELAY J L. JPEG-based scalable privacy pro-

tection and image data utility preservation[J]. IET Signal Process. 2018(12): 881-887.

- [11] LI J, WANG N, ZHANG L, et al. Recurrent feature reasoning for image inpainting[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). 2020: 7760-7768.

[作者简介]



文泓鑫（1998- ），男，广东汕头人，中国科学技术大学硕士生，主要研究方向为图像隐私保护。



胡飞（1993- ），男，安徽六安人，中国科学技术大学博士生，主要研究方向为图像隐私保护。



俞能海（1964- ），男，安徽无为市人，中国科学技术大学教授、博士生导师，主要研究方向为视觉计算与安全、信息隐藏、数据安全、人工智能安全。