

三维动画与交互技术课程读书报告



题目 基于假色可靠且可逆的图像隐私保护方法

作者姓名 姜涛

作者学号 21851467

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二○一八年十二月

A Reliable and Reversible Image Privacy Protection

Based on False Colors

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Li Qilei

By

Jiang Tao

Zhejiang University, P.R. China

2018

摘要

保护视觉隐私已成为不可或缺的视频监控系统的组成部分由于普遍使用摄像机进行监视。在本文中，我们提出了两种在JPEG体系结构中实现的完全可逆的隐私保护方案。在两种方案中，通过使用假色来实现隐私保护，其中第一方案适用于其他隐私保护过滤器，而第二方案是假色特定的。两种方案都支持无损模式，其中原始未受保护的内容可以被完全提取，或者支持有损模式，这虽然限制了文件大小，但维持了图形在可懂度。我们的方法不是基于兴趣区域（ROI），可以应用于整个帧而不会影响可懂度。这使得用户不必定义ROI并提高安全性，因为动态内容下的跟踪ROI可能会失败，从而暴露敏感信息。我们的实验结果表明我们的方法优于其他常用解决方案来保护视觉隐私。

**关键词**：隐私保护，假色，JPEG

Abstract

Protection of visual privacy has become an indispensable component of video surveillance systems due to pervasive use of video cameras for surveillance purposes. In this paper, we propose two fully reversible privacy protection schemes implemented within the JPEG architecture. In both schemes,privacy protection is accomplished by using false colors with the first scheme being adaptable to other privacy protection filters while the second is false color specific. Both schemes support either a lossless mode in which the original unprotected content can be fully extracted or a lossy mode, which limits file size while still maintaining intelligibility. Our method is not regionof-interest (ROI) based and can be applied on entire frames without compromising intelligibility. This frees the user from having to define ROIs and improves security as tracking ROIs under dynamic content may fail, exposing sensitive information.Our experimental results indicate the favorability of our method over other commonly used solutions to protect visual privacy.

**Keywords：**Privacy protection, false color, JPEG.

**1引言**

**1.1研究背景**

英国安全工业管理局最近的一份报告显示，仅英国就有近600万台闭路电视摄像机，或者每11个人就有一台摄像机。过度使用视觉监控会引起公众对个人隐私的担忧。除了监视之外，摄像机现在经常用于环境辅助生活应用，其中确保视觉隐私也是一个关键问题[1]。

由于在本地服务器中存储和分析大量多媒体数据的困难，将这些任务推迟到云服务器已经普及。然而，这进一步加剧了隐私问题，因为这些数据也可能被未经授权的方获取。这产生了各种数据隐藏方案，其中存储在服务器中的数据以可逆方式加密，或者作为密文或者作为明文。

然而，现有的多媒体内容隐私保护方法似乎都没有得到普及。其背后的原因之一是迄今为止提出的所有视觉隐私保护解决方案都依赖于手动识别敏感区域或者需要计算机视觉模块这样做，导致复杂的操作往往缺乏稳健性和可靠性。基于JPEG的云安全解决方案也不适合监控任务，因为它们要么将内容完全加扰或映射到不同的目标图像，这与原始图像无关。

**1.2 研究问题**

我们需要的解决方案旨在在视觉隐私保护中重要的各种标准之间取得平衡，即隐私，可懂度，可逆性，安全性和稳健性。更具体地说，我们的目标是：（1）安全视频中记录的个人不应该被人类观察者和面部识别算法（隐私）轻易识别;（2）受隐私保护的视频仍然应该允许识别可疑行为和收集非 - 敏感信息，例如给定区域内的人数（可懂度）; （3）在犯罪的情况下，可以撤销受隐私保护的镜头以获得授权用户原始未受保护的镜头（可逆性）; （4）此撤销只能由合法授权的当事人进行，而不能由可能通过某种方式获得受保护内容的任何第三方（担保）进行; （5）隐私保护应该是健壮的，因为它不应该依赖于脆弱的计算机视觉算法或可能无法在某些帧中检测到敏感区域的手动注释（鲁棒性）。

**1.3研究方法**

为了实现这些目标，我们提出了在JPEG体系结构中实现的两种基于伪彩色的方案。两种方案都是相关的;然而，第一个并不特定于假色，并且如果需要，可以与其他隐私保护算法一起使用。另一方面，第二种是专门用于假色。第二种方案的好处是它通过利用原始图像和伪彩色版本之间的一致性显着减小了受保护内容的文件大小。

我们的实验结果涉及感知上有意义的质量指标和人脸识别算法，表明我们的方法优于其他隐私保护方法。此外，我们使用主观实验验证我们的技术，这证实了我们的方法实现了比比较更好的可懂度 - 隐私平衡。最后，我们表明我们的方法能够抵御旨在从受保护内容中恢复原始信息的攻击。

**2 相关工作**

视觉隐私保护的目标是防止图像（或视频）中存在的敏感信息被泄露给该内容的观看者。已经提出了许多方法来实现该目标，从简单的过滤方法到基于复杂计算机视觉的算法的复杂性变化。在本节中，我们首先概述了视觉隐私保护算法的主要类别。然后，我们审查基于伪彩色的隐私保护。最后，我们讨论使用JPEG元数据作为恢复信息的JPEG和相关方法。

**2.1** **常用视觉隐私保护（vpp）方法**

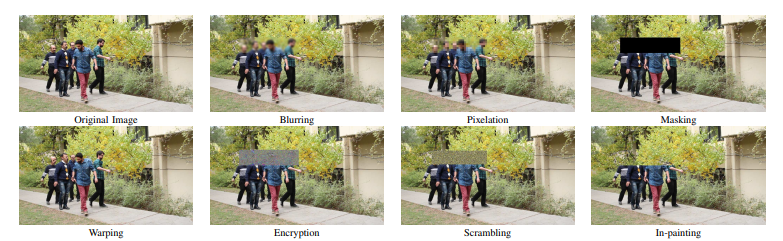


图1

最常用的三种VPP方法是屏蔽（Masking）（图一上4），模糊(Bluring)（图一上2）和像素化(pixelation)（图一上3）。涉及加密的算法是最安全的的算法。这些算法将整个图像或选定的ROI视为比特流。由于这些算法的时间复杂性，已经提出了针对数字视频的更轻量级加密方法。效果如图一下2和图一下3。面部匿名化（即去识别）的另外两种值得注意的方法被称为扭曲（morphing）和变形(warping)(图一下1)。两种方法的缺点在于，取决于插值参数和翘曲强度，原始面可能是不可恢复的。

以上所有的VPP方法的主要缺点是它们需要用户定义或自动提取的ROI来应用隐私保护。 对全帧应用保护严重损害了捕获数据的可懂度。 手动定义ROI是不切实际的，自动提取ROI的鲁棒性取决于对象/人体检测算法的稳健性，可知这些算法在恶劣的捕获条件下会失败。

**2.2假色方法**

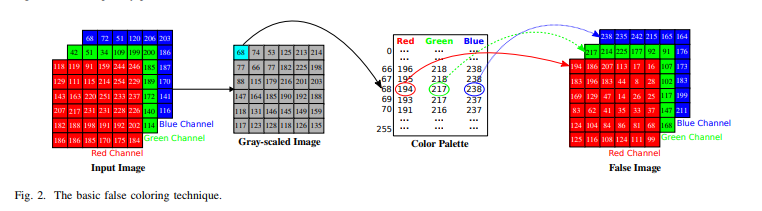


图2

在图像处理中，假色通常用作可视化辅助以表示否则不可见的信息。最近，假色被用于视觉隐私保护的目的。为此，首先将RGB输入图像变换为灰度。然后使用8位灰度值来索引RGB颜色表（即调色板），并使用相应的RGB三元组来替换原始像素值（参见图2）。基于伪彩色的VPP的主要优点是它可以应用于整个图像而不会影响可懂度。换句话说，不需要选择ROI，这使得该方法能够抵抗旨在检测敏感区域的计算机视觉算法的脆弱性。然而，由于两个原因，该方法不是完全可逆的：（1）原始颜色到灰度转换消除颜色信息和（2）调色板通常不是一对一，这意味着两个不同的灰度值可能获取映射到相同的颜色值，使得无法恢复原始灰度值。在本文中，我们建议扩展基于伪彩色的VPP，以便可以完美地恢复原始未受保护的内容，同时如果关注点在文件大小，我们仍支持有损恢复。此外，我们在行业标准JPEG格式中实施此方法。

**3** **两种解决方案**

我们提出的方法包括两种方案。 第一种方案更为通用，除了虚假着色外，它还可以与任何隐私保护策略一起使用。 另一方面，第二种方案利用调色板来减小受保护图像的文件大小而不影响其可懂度。 两种方案都支持无损和有损模式。

**3.1** **方案一**

1）保护管道：第一种方案的保护流水线如图3所示。这里，输入图像（I）首先转换为灰度。 接下来，通过使用灰度值作为调色板的索引，获得伪彩色图像（FI）。 此图像将作为主JPEG图像保存在输出文件中。 然后对FI进行JPEG编码和解码以模拟解码器在解码端将看到的内容。 我们将此图像称为FI'。 之后，差异图像（DI）计算如下：



其中（x，y）表示像素索引，c∈{R，G，B}。由于此差异有时为负，因此也会计算伴随的符号图像（SI）：



为了有效存储，我们对每个差异使用一个比特，然后使用zlib压缩算法无损压缩它[2]。 差异图像可以无损压缩或有损压缩。 对于无损压缩，我们使用zlib而对于有损压缩，我们使用JPEG压缩或下采样（两者也可以同时使用）。 然后将压缩和加密的DI和SI作为元数据保存在JPEG应用程序标记中。

2）恢复管道：为了恢复（图4），首先解码JPEG文件以获得伪彩色图像，差异图像和符号图像。 注意，解码的伪彩色图像将等于上面介绍的FI'。 在无损模式中，差异和符号图像将等于DI和SI。 首先使用授权密钥解密这些流，然后解压缩。 恢复的图像R通过以下方式获得：





注意，在差分图像（DI）被无损压缩的情况下，恢复的图像R将等于原始图像I. 否则，R将根据压缩工件的指示偏离I。

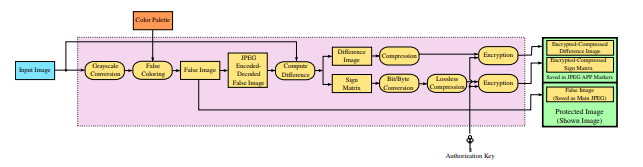


图3 保护管道

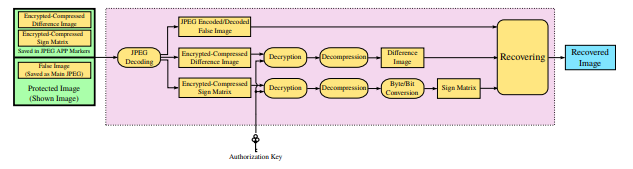


图4 恢复管道

**3.2** **方案二**

1）保护管道：我们的第二个方案的主要工作流程与第一个方案类似（图5）。 然而，在该方案中，我们利用原始图像和反转的伪彩色图像之间的相干性来显着减小受保护图像的尺寸。 在该方案中，虚假彩色图像FI的计算方式略有不同，以避免颜色到灰度的转换（参见图6）。 对于原始图像中的每个颜色值Ic（x，y），相应的伪色值FIc（x，y）被计算为：



其中Pc代表调色板第个通道。

接下来，在对进行编码和解码以获得之后，我们首先应用逆表查找以获得I'，而不是直接从I中减去它:



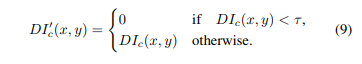
这里，表示调色板的第个通道的伪逆。 我们将其称为伪逆，因为大多数调色板不是一对一的，因此是不可逆的。 在实践中，给定，我们在Pc内部搜索以找到最相似颜色值的索引：



如果有多个这样的值可以最小化这种差异，我们根据原始图像的直方图选择索引。 例如，如果i = 5且i = 125是等式8的两个解，并且hist（Ic）[125]> hist（Ic）[5]，我们选择125作为逆。 这确保了反转值将与最大像素数的原始值类似。

一旦以这种方式计算I'c，就从I中减去它以获得差异和符号图像（在用I'c代替F I'c之后使用等式2和3）。 注意，与方案一不同，在这种情况下，差异图像将具有可以有效压缩的许多零或小组件。

此外，在这种特殊情况下，存在不同类型的有损压缩的机会。 根据质量阈值τ，DI中小于τ的所有值都可以设置为零：



在这种情况下，相应的值也应该设置为零，以提高符号图像的压缩效率。 在此过程之后，差异和符号图像在被写入JPEG应用程序标记之前被压缩和加密。 我们还压缩、加密和写入每个通道的直方图和JPEG应用程序标记内的调色板，以便在恢复过程中使用。

2)恢复管道

在第二种方案的恢复流水线中（图7）,首先解密所有加密的元数据，然后根据需要进行解压缩。 然后使用解码的伪彩色图像FI'，输入图像直方图hist(Ic)和调色板P（原始图像的近似值），使用等式7计算I'。然后将该图像与差异和符号组合。 通过替换，如等式4中的图像对于F我来说。 注意，类似于方案一，如果差异和符号图像没有以无损方式被阈值化和压缩，则恢复的图像R将与原始图像I相同。

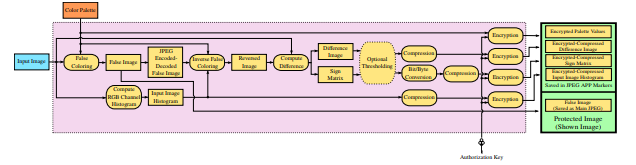


图5 保护管道

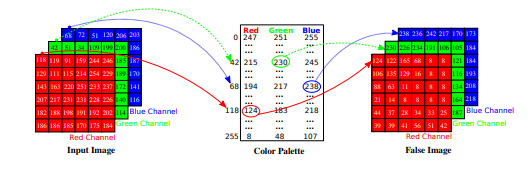


图6 方案二的虚假着色方法

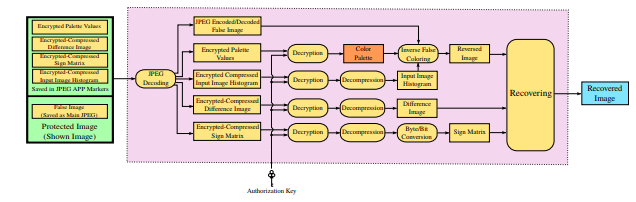


图7 方案二的恢复管道

4 **实验结果分析**

为了证明方法的有效性，进行了一系列实验，实验主要针对以下几个方面：

(1)在无损压缩模式下，两种方案的视觉效果。

(2)在有损压缩对文件大小和恢复图像质量的影响。

(3)使用面部识别基准和其他众所周知的隐私保护方法的比较。

(4)安全性检验。

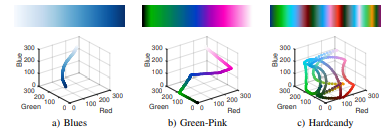


图8 本实验所使用的三种调色板特征

**4.1 无损压缩**

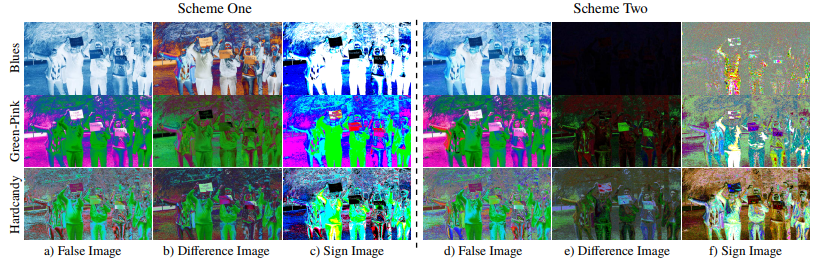
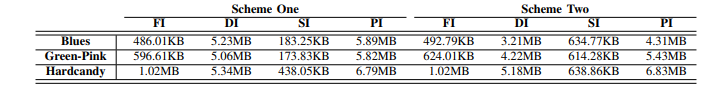


图9

我们首先说明我们对图9所示的抗议场景的结果。每行代表不同的调色板。第一列显示隐私保护的假彩色图像，第二列显示差异图像，第三列显示标志图像（最后三列中的方案二相同）。对于所有三种调色板和两种方案，可以观察到虽然受保护的图像仍然是可理解的，但人们的身份大多是隐藏的。

通过检查这些图，可以观察到两种方案之间的差异以及调色板的效果。首先，可以观察到两种方案的假彩色图像非常相似。对于差异图像，第二种方案产生的值比方案一中的值小。这是由于在第二种方案中应用的恢复步骤（方程式7和8）。但是，调色板也会影响恢复质量。对于比其他两个调色板更单调的蓝调调色板（图8），恢复的质量非常好，因此差异图像包含非常小的值。然而，对于其他两个调色板，由于它们在色标上的规则变化较小，因此恢复的效果逐渐降低。

至于符号图像，它们看起来更嘈杂，因为差异图像具有更小的像素值。这是因为小的差异可能是正的或负的，并且这可能在像素之间快速变化（甚至在像素的颜色通道之间）。两种方案中的这些组件的字节大小也支持这些观察结果（表I）。这里，FI表示JPEG质量设置85的伪彩色图像的JPEG压缩大小; DI和SI表示差异和符号图像的无损zlib压缩大小; PI表示受保护图像的总比特流大小。如该表I所示，方案2中的DI小于方案1中的DI，而SI更大。此外，方案二中的DI值随着调色板的方差而增加。



表I

**4.2 有损压缩**

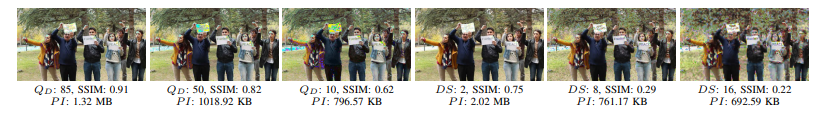
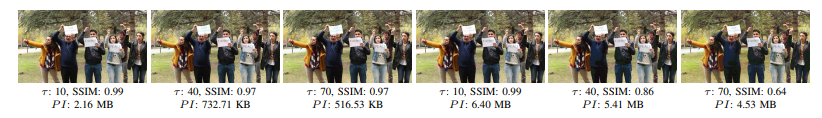
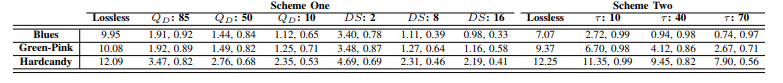


图10



、图11

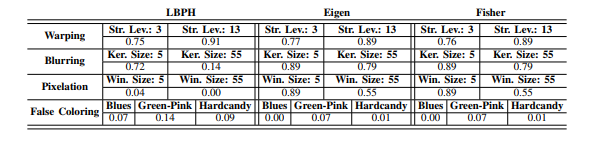


表II

在许多应用程序中，受保护图像的文件大小可能是关键的，并且所需元数据引入的额外开销可能太大。为此，我们还提出了一种有损模式，并在图像质量与文件大小方面展示了我们的实验结果。

为了概括我们的结果，我们捕获了几个模拟各种监控场景的视频。从每个视频中，我们选择了代表性帧，总共产生12个测试图像。然后我们执行了前面描述的有损压缩方法。我们的结果报告在表II中。对于无损情况，第二种方案使用蓝调调色板（7.07）提供最佳压缩比。对于第一种方案的JPEG编码与下采样，前者不仅产生更好的压缩，而且产生更高的平均SSIM分数。通过增加第二种方案中的τ阈值，可以获得甚至小于原始文件的受保护文件。然而，对于高度不规则的调色板，例如Hardcandy，即使高阈值也不会产生非常小的文件，因为差异图像包含许多高于该阈值的像素。

**4.3 面部识别**



表III

汇总结果如表III所示。在此表中，值越低，保护性能越好。对于所有三种FRA和所有三种调色板，基于假色的隐私性能非常好，LBPH-Green-Pink组合的最高识别率为0.14。 Blues和Hardcandy调色板表现非常出色，其识别率不高于Eigenfaces和Fisherfaces算法的1％的机会水平。在比较算法中，我们的方法仅通过使用LBPH算法的像素化来表现。然而，如图12所示，像素化窗口大小为5在保护隐私方面无效。另一方面，55的窗口大小完全阻止了可懂度和可逆性。此外，在其他FRA下，像素化的表现并不好。基于这些结果，可以认为基于伪彩色的隐私保护优于该评估任务的比较方法。

总而言之，人脸识别评估表明所提出的伪着色算法比比较的方法更好地保护隐私。此外，即使在通过有损压缩算法压缩时，也可以反转受保护的图像以获得可由FRA识别的恢复图像。



图 12

**4.4 安全性评估**

除了隐私和可懂度之外，对未授权用户的安全性是有效VPP算法的另一个理想特性。 在这里，我们评估假冒着色对于试图将受保护的面部图像还原为其未受保护的版本的攻击者是否安全。 为此，我们从代表不同种族人群的FERET数据集中选择了12张面部图像（图15）。 从每个面部，我们选择对应于皮肤，唇部和头发区域的像素。 对于每个区域，平均5×5邻域内的像素值以获得代表性颜色。

接下来，选择由我们的第二方案转换的目标面部。 计算相应面部区域的假色RGB值，产生以下映射：



这里，Gr表示来自FERET图像的区域r∈{皮肤，头发，唇部}的灰度值，RGBr表示来自目标面部的对应的假色值。然后通过增加Gr值对该映射进行分类，并且通过线性插值计算调色板的缺失值。请注意，当我们使用12个输入图像时，这导致每个目标面重建12个调色板。

接下来，产生对每个重建调色板的逆查找以找到每个伪色像素的对应灰度值。一些调色板产生比其他调色板更好的结果，我们在视觉上确定最佳调色板作为目标面的最准确反转。对于在该研究中评估的每种假色调，结果显示在图16中。从图中可以看出，虽然Blues和GreenPink调色板容易受到这种攻击，但Hardcandy调色板仍然可以抵抗每个目标面。

**5 结论与展望**

我们提出了一种在JPEG体系结构中实现的基于伪彩色的隐私保护算法，并通过进行大量实验证明了其性能。特别是，我们已经证明了我们的方法不仅可以有效地对抗人脸识别算法，而且可以通过客观和主观的评估来对抗人类观察者。

我们的方法的主要优点是它可以应用于整个图像，从而无需定义隐私敏感的ROI。这很重要，因为虽然手动定义ROI是麻烦的并且其实用性仅限于静态场景，但是通过检测算法自动选择ROI受到这些算法的鲁棒性的影响。

然而，就安全性而言，正如我们的安全评估实验所证明的那样，通过映射伪彩色像素值和真实物体颜色之间的关系，常规调色板更有可能被未经授权的个人辨认。发现诸如Hardcandy之类的不规则调色板可提供更高的安全性。然而，在安全性和可懂度之间存在平衡：完全随机的调色板将非常安全但不易理解，因为所有结构细节都将丢失。此外，如上一节所示，使用更多不规则的调色板会导致更大的受保护文件大小。也许，最理想的方法是定义专门为隐私保护目的而设计的自定义调色板 - 这是我们将来留给未来工作的问题。

最后，尽管针对JPEG图像示出了所提出的解决方案，但是可以以简单的方式将其应用于支持包含元数据的任何其他图像和视频格式。例如，在AVC / H.264和HEVC / H.265格式中，该机制可以通过补充增强信息（SEI）标记来实现。这种扩展可能有助于通过现实世界的监视系统更广泛地采用基于伪彩色的隐私保护。

参考文献

[1] X. Yuan, X. Wang, C. Wang, J. Weng, and K. Ren, “Enabling secure and fast indexing for privacy-assured healthcare monitoring via compressive sensing,” IEEE Trans. on Multimedia, vol. 18, no. 10, pp. 2002–2014,Oct 2016.

[2] L. Sweeney, “k-anonymity: A model for protecting privacy,” Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 05,pp. 557–570, 2002.

[3] P. Deutsch and J.-L. Gailly, “Zlib compressed data format specification version 3.3,” Tech. Rep., 1996.

[4] P. Korshunov, A. Melle, J.-L. Dugelay, and T. Ebrahimi, “Frameworkfor objective evaluation of privacy filters,” in SPIE Optical Engr.Applications. Intl. Society for Optics and Photonics, 2013, pp. 88 560T–88 560T.