多媒体信息安全研究综述

孔祥维[1] 王 波[1] 李晓龙[2]
[1](大连理工大学电信学部 大连 116023)
[2](北京大学计算机科学技术研究所 北京 100080)
(kongxw@dlut.edu.cn)

Multimedia Information Security: A Review

Kong Xiangwei[1], Wang Bo[1], and Li Xiaolong[2]
[1](*Department of Electronics and Information Engineering, Dalian University of Technology, Dalian* 116023)
[2](*Department of Computer Science, Peking University, Beijing* 100080)

Abstract In the information and network era, multimedia is confronted with severe information security risks, though it has provided great audio and visual enjoyments to people for last few decades. In view of the typical research areas of multimedia information security, including steganography and steganalysis, digital watermarking, multimedia forensics, perceptual hash and multimedia content privacy, this paper firstly expounds the background of specific problems and explains the typical concept. Based on the analysis of the previous works, the potential problems and challenges in the future are summarized and discussed separately. Finally, this paper draws the conclusion and shows the prospect of multimedia information security.

Key words multimedia information security; steganography and steganalysis; digital watermarking; multimedia forensics; perceptual Hash; multimedia privacy

摘 要 信息和网络时代的大环境下,多媒体给人们提供了音视觉享受的同时,也存在着更为严峻的信息安全问题.针对当今主流的多媒体信息安全的典型研究方向,信息隐藏和信息隐藏分析、数字水印、多媒体内容取证、感知哈希和多媒体内容隐私等,首先阐述了各个方向领域的研究背景、总结了其针对性的科学技术问题,解释了典型的概念,然后在介绍和分析研究现状的基础上,给出了当前存在的问题和挑战,最终给出结论和展望.

关键词 多媒体信息安全;信息隐藏和信息隐藏分析;数字水印;多媒体取证;感知哈希;多媒体隐私

中图法分类号 TP391

信息安全是世界性的现实问题,它牵涉到国家的政治安全、经济安全、社会安全、军事安全乃至文化安全,世界主要国家和地区均将信息安全视为国家安全战略的重要基石.

近5年来互联网发达国家密集出台国家网络安全新战略,加速战略核心内容的落地部署.自2013

收稿日期:2015-07-30
基金项目:国家自然科学基金创新团队基金项目(71421001);国家自然科学基金项目(61172109,61502076)

年投入103亿美元的网络安全预算以来，美国用于网络安全的资金投入近年来呈稳步增长趋势，美国总统奥巴马提议在2016财年预算中，拟拨款140亿美元用于加强美国网络安全，以便更好地保护联邦政府和私有企业网络免遭黑客威胁①. “斯诺登事件”震撼了整个世界，让我们清晰地看到核心技术受制于人带来的可怕后果. 斯诺登爆料：美国情报机构正致力于准备网络战争，美国网络战的内容涉及网络战争、远程控制、植入性病毒、黑客攻击与反攻击等. NSA正在秘密开发多种方式，“网络曼哈顿计划”、“爱因斯坦计划”正在实施，以绕开互联网上的多种加密技术，西方多个大国都在不断加大赛博战力量建设. 美国等西方国家正在利用信息霸权、信息威慑谋求主宰世界，对此我国信息安全形势非常严峻. 目前我国对信息安全的重视程度日益提高，已制定和实施《国家安全战略纲要》.《国家中长期科学技术发展规划》也将面向核心应用的信息安全列为重点发展的优先主题②. 为实施国家安全战略，“网络空间安全”已设为一级学科；中央网络安全和信息化领导小组的成立标志着网络空间信息安全上升至国家战略高度.

互联网络的普及和多项业务的飞速发展，已经渗透到经济、社会各个领域. 社交网络、在线媒体、即时通信、移动互联网、物联网等成为主要增值形态[1]，人与人通信的业务模式从单一语音走向丰富多彩的多媒体数据业务，用户可以随心所欲地享受多媒体信息展现的数字新生活. 互联网、社交网、物联网的数量和规模呈爆炸式增长，2012年多媒体中的图像和视频数据已经占到大数据的80%，2013年图像和视频数据在整个大数据的比例已经接近90%.

历来的新事物都是双刃剑，多媒体大数据在提供人们音视觉享受的同时存在着更为严峻的信息安全问题. 多媒体(multimedia)是多种媒体的综合，一般包括文本、声音、图像、视频等多种媒体形式. 国际标准化组织ISO对信息安全定义是：为数据处理系统建立和采取的技术和管理的安全保护. 保护计算机硬件、软件、数据不因偶然的或恶意的原因而受到破坏、更改、泄漏. 多媒体信息安全指的是与多媒体相关的信息安全.

由于多媒体形式的特殊性，沿用传统的密码加密多媒体存在很多现实问题. 早期的研究是将传统的数据加密算法直接用于多媒体数据加密，但由于多媒体数据信息量大、图像和视频相关性高、数据存在冗余等特点，直接加密方法存在计算量大、时间长、有延时、功耗大等局限. 因此多媒体安全这一新的信息安全方向应运而生.

从多媒体的生命周期来看：多媒体从传感器产生后依次以原始格式、多媒体内容、多媒体处理以及多媒体通信和表现等形态存在于多媒体的应用中. 目前在公开文献中出现了多种围绕多媒体各种形态的相关安全研究. 典型的有以多媒体为掩护载体的信息隐藏(information hiding)等；对多媒体内容版权保护和追踪的数字水印(digital watermarking)；对多媒体原始性鉴别的多媒体取证(multimedia forensics)；对多媒体认证的多媒体感知哈希(perceptual Hash)；以及对多媒体敏感内容保护的多媒体内容隐私(multimedia privacy)等与多媒体内容安全相关的研究.

本文针对当今主流的多媒体信息安全的典型研究方向信息隐藏和信息隐藏分析、数字水印、多媒体内容取证、感知哈希和多媒体内容隐私等，首先阐述了针对性的背景问题，解释了典型概念，在分析了研究现状的基础上，给出问题和挑战，最终给出结论和展望.

1 信息隐藏及分析的研究与发展

用于隐蔽通信的信息隐藏或隐密(steganography)是具有古老历史并沿用至今的隐蔽通信方式. 其基本思想是把秘密消息隐藏在正常载体中，通过隐藏秘密消息的存在性来构建隐蔽通信. 信息隐藏和隐密术与传统的密码不同，密码术的核心在于让传递的秘密信息“不可读”、“不可懂”，而信息隐藏和隐密术关注的是如何更好地将秘密信息隐藏在载体中进行通信，这使得通信行为本身得以隐藏，让监听者或者分析人员无法得知通信的发生.

与古典隐蔽通信相比，现在利用信息隐藏进行秘密数据的嵌入，其方法更为复杂，伪装所用的载体

① 奥巴马拟拨款140亿美元增强网络安全建设(http://world.chinadaily.com.cn/2015-02/13/content_19573777.htm 2015，2，13)

② 国家中长期科学和技术发展规划纲要(http://www.gov.cn/jrzg/2006-02/09/content_183787.htm 2006，2，9)

更为广泛和多样. 多媒体数据的海量化、多元化、异构化、网络化、云碎片化，为信息隐藏和隐密术的快速发展提供了良好的环境. 不同格式的数字图像、数字音频、视频文件以及其他媒体都可以作为载体进行信息隐藏. 网络上已经出现了上百种利用不同格式的数字媒体进行隐密术的公开软件，其中大部分是免费甚至开源的. 其他诸如不同格式的文本、流媒体[2]、语言翻译系统、文件系统、网络协议、量子信息[3]等也都被尝试用于信息隐藏.

信息隐藏和隐密术的研究者们常用“囚犯模型”来描述典型的信息隐藏通信过程. 在“囚犯模型”中，通信双方 Alice 和 Bob 是要进行隐密通信的主体，其目的是在监听者 Wendy 的严密监视下进行隐密通信；而监听者 Wendy 的角色是攻击者，其任务是要发现 Alice 和 Bob 之间的通信. 这种对抗也就形成了以多媒体为载体的信息隐藏两大对抗式研究方向：信息隐藏(或隐密)以及信息隐藏分析.

1.1 安全信息隐藏或隐密

信息隐藏和隐密术的安全性可被归纳为 4 个方面：格式安全、感知安全、统计安全和系统安全.

1) 格式安全：网络上公开的一些典型信息隐藏和隐密术软件，将秘密信息嵌入到载体文件某些固定的冗余位置而不破坏多媒体的文件定义和结构. 例如放在文件末尾或者中间保留/注释位置. 这样的信息隐藏和隐密术方法可以轻易地通过简单的格式审查、文件信息判断等方式来分析其格式安全性.

2) 感知安全：秘密信息通过隐藏在多媒体载体当中进行传递，由于多媒体本身是具有视听内容的媒体，因此在隐蔽通信过程中，多媒体的自身价值和被感知的内容及质量不能发生改变. 因此，几乎所有的信息隐藏方法都有不可感知性的基本要求[4].

3) 统计安全：当前信息隐藏和隐密术的研究和关注的重点都在统计安全上. 以针对 JPEG 数字图像的隐密术为例，JSteg 隐密算法会由于系数成对翻转，使得原本服从广义高斯分布的量化 DCT 系数直方图出现“对效应”；F5 方法存在收缩效应. OutGuess 算法虽然保留了量化 DCT 系数直方图分布，但会由于频域的信息嵌入而导致空域块连续性遭到破坏[5]. 可以看出，早期的信息隐藏和隐密方法都是通过有针对性地保持已知统计模型来实现其统计安全性. 近年来 Fill 等人提出的最小失真嵌入框架[6]比较有代表性，例如 UNIWARD 算法[7]以及均匀嵌入策略[8]相对于早期的方法更为安全.

4) 系统安全：隐密算法在工程实现过程中的安全性涉及到系统安全. 典型的例子是 ImageHide 和 JPEGX 这 2 个图像隐密软件，为了系统识别含密图像而引入了隐密软件的“特征码”，使得隐密分析人员可以通过简单快速的分析准确检测含密图像，甚至提取秘密信息[9]. 因此，系统安全也是信息隐藏安全实际应用中必须考虑的重要环节.

实际应用中的安全信息隐藏和隐密通信需要综合考虑 4 个方面的安全性能，根据其应用环境、安全等级要求等作出多方面的平衡和优化.

就当前而言，统计安全仍然是隐密算法设计者首要考虑的最重要的问题之一：如何更好地构造失真函数，使其能够尽可能全面地约束多媒体的统计失真[10-11]；同时，设计鲁棒的隐密技术将是一个对实际应用十分重要并且很有挑战性的问题；最后，如何在容量、安全和鲁棒性之间进行平衡，则是隐密系统在实现和应用当中的难点.

1.2 信息隐藏分析对抗

信息隐藏的安全分析——信息隐藏分析暨隐密分析(steganalysis)——关注于检测隐蔽通信的存在. 不失一般性，下文的分析和描述以图像载体为例.

目前的隐密分析可以归结为一个二类分类的匹配分析问题. 隐密分析一般包含 2 个主要步骤：特征提取和分类，其核心大多是通过对载体残差的统计分布进行建模，然后进行分类. 目前面向通用隐密分析最有效的是 Rich Model 模型[12]. 还有文献利用选择信道先验信息，构造了改进的 SRM 特征集[13]. 深度学习也被用于隐密分析当中[14]. 这些方法都采用了对已知标签样本进行统计特征训练构建高维空间模型的方式，因此不可避免地隐含了训练样本和测试样本具有相似甚至相同的统计特征分布这样一个假设条件. 然而在实际的隐密分析情况中这一点难以满足，导致产生隐密分析的失配问题.

隐密分析失配问题已经引起学术界的普遍关注，2013 年，Ker 等 8 位[15]本领域最有名望的学者联名撰文，呼吁学术界加强实用隐密技术和隐密分析技术的研究，力求使这 2 项技术从实验室推向实际应用. Ker 等人[16]针对现实世界中利用非公开图像源或算法的隐密行为，提出了基于聚类分析的隐密分析框架[16]. 黄炜等人[17]在此基础上提出了利用

核 Fisher 鉴别指标计算样本间差异度量的聚类方法. 针对训练集和测试集中载体图像的失配问题，Lubenko 等人[18]分析了相同统计特征情况下支持向量机与集成分类器的性能差异. Kodovsky 等人[19]提出了依据 JPEG 图像量化表进行预分类的隐密分析解决方案. Li 等人[20]将迁移学习引入到失配隐密分析的研究当中. 平西建等人[21]则尝试使用四叉树分割来解决训练与测试图像统计特征分布差异问题. Ker 等人[22]将不同载体来源所造成的隐密分析失配问题进行了分类，并给出了降低该因素影响的具体策略. Zeng 等人[23]则尝试利用域对齐的方法来解决隐密分析失配.

对于信息隐藏和隐密分析来说，在互联网上实用仍然是目前研究面临的最大难题：高准确率意味着更为客观和准确的建模，准确的建模意味着数量更为庞大，分布更为相似的训练样本. 训练样本和测试样本数目的平衡实际中难以满足；多种参数在实用环境中由于多种失配会导致目前的分析方法性能失效，如何提高失配分析的准确率是现实的难题；数据不均衡导致安全分析难度增加，训练数据多、实际数据量少导致的数据分布不均时，安全分析的难度会大大增加. 而作为信息隐藏和隐密术的安全评估手段，隐密分析的图像库建立也同样是一项意义大、难度大的工作[24].

2 数字水印的研究与发展

鲁棒数字水印的目的是用于保护多媒体作品的版权，主要技术是内嵌入不影响感知的水印标识在多媒体作品历经流通后仍能检测出作为版权拥有者标记，其核心是寻求嵌入量、安全性、鲁棒性三者之间的平衡. 从数字水印的设计思想上看，当前主要有 2 类比较流行的鲁棒水印.

1) 扩频水印：Cox 等人[25]最早建立了扩频水印的设计理论. 扩频水印对压缩和噪声攻击具有很强的鲁棒性，但是其检测是非盲的. 之后关于提升扩频水印性能的研究可归纳为 4 个方面：①选择不同的水印嵌入域[26]；②分析水印检测器的性能[27]；③水印感知性和鲁棒性的折中设计[28]；④抗攻击的策略[29]. 扩频水印将载体信号视为噪声，即便检测器利用了载体信号的分布信息，仍然不能完全去除载体信号对检测器的负面影响. 然而根据脏纸编码理论[30]，对于高斯分布的载体信号和随机高斯噪声信道，载体信号的干扰是可以完全剔除的，因此，大量的研究工作致力于逼近脏纸编码的理论结果.

2) 量化水印：量化索引调制水印(QIM)是这方面研究的一个典型代表. QIM 构造一组量化器，然后根据隐藏信息选择一个量化器量化载体信号得到含水印的信号，水印检测使用最小距离检测. 根据 QIM 的思想，人们提出了多种不同的实现方案. 扩展变换抖动调制水印(STDM)利用抖动量化器量化载体信号在某个扩展向量上的投影[31]. Eggers 等人[32]给出了二元 SCS(scalar costa scheme)的多元设计方案，具有更优的性能. 原始的 QIM 方案都使用了固定的量化步长，对载体信号各个分量的修改量是随机的，不符合人的感知系统特征，使用自适应的量化步长则可有效改善水印的性能. Li 等人[33]提出了图像 DCT 域内的自适应 DM，每个 DCT 系数的量化步长使用 HVM 模型来设计. Bao 等人[34]提出了基于小波域奇异值分解的自适应 DM 技术，其中量化步长由各个图像块的统计特性确定.

虽然已有很多鲁棒水印算法提出，但现有算法的鲁棒性仍然不能满足现有需求，同步问题一直没有很好地解决. 鲁棒水印仍是多媒体安全研究的一个难点和重点.

为了兼顾信息隐藏和原始载体的无失真恢复，可逆水印作为一种特殊的水印技术被提出. 这类隐藏技术使得接收者不仅能正确提取嵌入信息，还能将原始载体无失真复原.

现有空间域可逆水印根据嵌入原理，主要可分为基于无损压缩的可逆水印、基于整数变换的可逆水印、基于直方图修改的可逆水印三大类算法. 对于基于无损压缩的可逆水印，为了提高嵌入率，需要压缩更多的位平面同时也产生了明显的嵌入失真[35]. 基于整数变换的可逆水印想法新颖、设计巧妙[36]. 这类方法通常使用一个像素块的平均值来预测块中的每个像素点，不能很好利用图像冗余的同时，也不能有效控制嵌入失真. 相对于这 2 类可逆水印，基于直方图修改特别是基于预测误差直方图修改的方法具有更好的嵌入性能[37-38]. 相对一维直方图而言，基于二维直方图的可逆水印能够达到更好的嵌入性能[39-40]. 但现有的基于二维直方图的方法只是先验地设计了一种直方图修改模式，这种单一修改模式并没有考虑到所生成直方图的特性，无法达到最优

嵌入性能. 除此之外，加密域的可逆水印也是近年来研究的热点之一[41-42].

现有的针对可逆水印的研究尚存在下述主要不足：基于一维直方图的方法当嵌入量增大时性能下降较快，不能实现含密图像的高保真，而基于二维直方图的方法仅能提供较低嵌入容量；基于二维直方图的方法中，单一的直方图修改模式无法达到最优的嵌入性能.

3 多媒体内容取证的研究与发展

日益普及的数码设备、众多的多媒体编辑工具、互联网的快速传播和几乎人人持有的移动终端，分别从多媒体获取、处理和传输 3 个方面大大降低了内容篡改伪造的门槛. 在司法、新闻、军事、社会和科学等多个领域频发的多媒体内容造假事件，如周老虎、藏羚羊、广场鸽等，引发对于数字多媒体真实性和来源性的信任危机，促使数字多媒体被动无损取证技术成为了近年研究热点.

对多媒体内容的无损取证溯源技术主要可以分为两大类：一是对多媒体内容进行篡改伪造取证；二是对多媒体的获取设备进行溯源分析.

3.1 多媒体内容的无损取证

对多媒体内容的篡改伪造取证分析，由于增加或删除了部分内容进行篡改伪造，因此重要的一个内容就是对拼接操作的检测. 数字多媒体在统计特性上会呈现共同性和稳定性，这是在获取过程中所引入的固有特征，进行拼接操作后会破坏这种原有的特性，如能构建多媒体的固有属性描述模型，就能够对拼接篡改伪造进行检测. 同样以数字图像为例，相位一致性[43]、小波域统计特征[44]、马尔可夫转移概率模型[45]、图像质量特征[46]、图像光照模型[47]等都曾被用于构建设备获取的原始图像的模型，进而用于拼接篡改伪造图像的检测. 另外还有文献通过检测邻域相关性来检测尺度变换[48]；通过模型化量化 DCT 系数的分布检测 JPEG 压缩历史[49]；通过图像的二元相似性测度分析来检测亮度调整[50]；通过不同色彩空间的相邻像素一致性分析来检测模糊操作[51]；通过图像边缘的振铃效应来检测锐化操作[52]等. 对这些局部性操作的检测结果可以进一步对篡改伪造区域进行初步的定位.

多媒体篡改伪造取证检测的核心问题仍然集中在检测和分析的准确率上. 多媒体篡改伪造的方法和手段具有多样性和复杂性，即便篡改伪造的目的和结果几乎完全一样，其过程和手段也可能完全不同. 这是多媒体篡改的特点决定的，导致多媒体取证方法难以通用. 由于现有的被动取证算法大多基于统计意义上的数据分析，许多模型构建还比较简单，因此已有的取证方法大多针对性较强，而距离实际场景中复杂多操作篡改过程的准确检测仍然还有一定的距离.

3.2 多媒体内容的溯源分析

对多媒体内容的溯源分析，大多是从多媒体数据的来源鉴别方面进行考虑. 以数字图像为例，典型的是以宏观统计特征和成像参数估计进行鉴别的方法. 宏观统计特征指的是图像在成像过程中，由于软硬件等系统函数综合作用于图像，从而形成的固有特征模式. 例如，利用不同来源图像在颜色、图像质量等差异对图像来源进行鉴别[53]，图像邻域的模式也被用于来源的分析[54]. 成像参数估计是在成像过程对特有参数进行估计判断图像的来源. 典型的成像参数估计方法有 Choi 等人利用镜头的非线性失真[55]和非零量化 DCT 系数分布规律[56]进行来源鉴别；CFA 也被广泛用于图像的来源分析[57]；相机光学传感器引入的模式噪声也被认为是图像获取设备个体鉴别的独特“指纹”信息[58].

值得注意的是多媒体内容的无损取证溯源同样具有典型的对抗性. 在多媒体内容无损取证研究蓬勃发展的同时，另一门旨在研究如何“逃避”取证方法检测的多媒体内容反取证也在不断地发展和进步[59].

4 感知哈希的研究与发展

多媒体认证侧重于保证信息内容的真实性和完整性，并具有检测和验证图像真实性和完整性的机制. 感知哈希(perceptual Hashing)是多媒体数据集到感知摘要集的一类单向映射，即将具有相同感知内容的多媒体数字表示唯一地映射为数字摘要，并满足感知鲁棒性和安全性，在数字媒体认证、盗版检测和信息源追踪中起着越来越重要的作用.

Kalker 等人[60]指出感知哈希具有 2 个特点：一是由含有大量数据的多媒体对象映射为较短的比特

序列;二是由感知相似媒体对象映射为数学意义上相近的哈希值,保证了感知相似的可传递性. 牛夏牧等人在人类感知模型基础上,明确了感知哈希的定义、性质和一般描述,并对感知哈希的一些典型算法、应用模式以及评测基准等进行了综述[61].

张维克等人利用图像DCT低频系数的感知不变性生成了安全的哈希序列索引,对经过图像处理操作的图像内容进行认证,可以抵抗内容保持的修改操作,具有较强的鲁棒性和安全性[62]. Bertino等人[63]使用生物特征识别和感知哈希来生成生物密钥的身份认证问题. 图像感知哈希函数使用了奇异值分解方法和支持向量机的分类方法进行设计,保证了生物识别特征识别器的独特性和可重复性.

丁凯孟等人[64]提出了一种基于边缘特征的遥感影像感知哈希算法,实现了遥感影像的完整性认证. 对边缘特征矩阵进行奇异值分解,选择较大的奇异值作为格网单元的特征,并采用感知哈希函数进行归一化处理,最后串联所有格网单元的归一化特征,得到感知哈希码. 计算待认证影像的感知哈希码,并与收到的感知哈希码进行匹配,从而实现遥感影像的完整性认证. 针对遥感影像,他们还提出了一种基于自适应格网划分的遥感影像感知哈希算法[65],实现遥感影像的完整性认证,解决了遥感影像数据量大和信息分布不均匀的问题.

Hu等人[66]为感知图像哈希提出和设计了一种安全和隐私协议,提出了可交换和加密的感知哈希的概念、距离保持特征加密的感知哈希的概念. Breitinger等人[67]用感知哈希技术来自动完成文件识别. 检测文件的哈希值与数据库中的哈希值进行比较,比较哈希值就可以得到这些文件的结果到底是可疑文件还是可信文件. Wang等人[68]提出一种基于可视化模型的感知哈希方法用于内容认证,他们结合统计分析方法和视觉感知理论,为内容认证开发了一种感知图像哈希方法. 为了实现真正的感知鲁棒性和感知敏感性,该方法使用Watson视觉模型提取视觉敏感特征,通过联合图像块特征和关键点特征,生成鲁棒的感知哈希码,从而实现了图像内容认证.

目前感知哈希的研究需要进一步解决以下问题:1)感知哈希函数构造问题需要理论支撑. 不同应用,感知特征提取方法会有所不同,哈希函数需要准则进行构造,准则的通用性是一个挑战;2)用于多媒体认证的感知哈希函数的安全性分析理论模型. 目前的方法缺乏系统的安全性分析,使得应用存在局限. 感知哈希的鲁棒性与抗碰撞性、篡改检测能力与随机性矛盾,感知哈希码长度和唯一性的平衡和哈希码的重复性和单一性之间的平衡.

5 多媒体内容隐私的研究与发展

多媒体大数据时代的到来是一把双刃剑,它在为人们提供信息价值的同时也为用户隐私埋下了安全隐患[69]. 近年来多起社会名流私密照片和视频被曝光的事件引发轩然大波. 因此,对多媒体大数据的妥善处理至关重要.

另一方面,视频监控的普及引发了大众对于个人隐私泄露的担忧. 政府设立的公共安全图像信息系统,监控范围遍布各种公共场所,其目的是提高预防和处置突发公共事件的能力,保障公共安全和公民的合法权益①. 因此,面对监控内容包含的群众隐私信息,合理的访问与保密机制是十分必要的.

社交网络是图像分享的主要平台,主流的社交网站所提供的隐私保护仅仅是面向大众的基础性保护. 大多数实用解决方案是允许用户自行设定图像的可见用户范围以及有效期,属于访问权限控制. Mazurek等人[70]则通过研究人脑模型,将语义信息与标签信息相结合建立逻辑访问控制. Ra等人[71]提出将照片中小部分重要信息和其余部分格式分别编码存储,以实现隐私保护. Yuan等人[72]提出了一种在线图像隐私保护的分享架构. 该结构以iOS手机应用ProShare为原型,利用置乱作为隐私保护的工具,同时实现对受保护图像的安全访问及在Facebook上的安全分享.

视频中对隐私目标区域的保护方法与图像类似,主要有对数据源的保护及与编码过程结合的保护[73]. 对数据源的保护包括数据替换、多重拷贝、数据分割,与编码结合的保护有加密编码、数据隐藏等. 另外还有设置访问权限,进行分级保护的多级访问控制技术. IBM的研究人员在2003年开发了PrivacyCam系统[74],首次提出了视频隐私保护. 其目

① 北京市公共安全图像信息系统管理办法(http://www.bjgaj.gov.cn/web/gspdAction.do?method=getFlfgInfo&lawid=1842 2006,12,15)

的是重点保护隐私区域，同时不影响视频的正常使用. 熊金波等人[75]对视频的安全等级进行划分，同时通过身份辨别对用户实行动态授权，只有访问权限与安全等级匹配才能访问相应的数据，利用语义交叉层次索引实现多级访问控制.

视频中隐私保护的目标主要是人脸或身体等能够泄露用身份的部分，由于视频的主体通常是运动物体，因此也有部分研究将运动物体作为保护的目标[76]. 对人脸的检测是机器学习中的热门研究，取得了诸多成果. 而在隐私保护中，除了直接对人脸进行识别的技术[77-79]以外，还有利用辅助信息自适应地根据需求选取隐私区域[80-82]. Korshunov 等人[79]通过人脸检测和变换技术，在保证变换后的人脸仍然保持脸部视觉特征的基础上，实现不可识别的结果，从而保护隐私.

针对云存储数据的隐私安全，朱旭东等人[83]提出了一种加密图像检索技术，根据视觉词汇表建立索引向量，并利用安全相似度运算确保索引向量运算和存储过程中的隐私安全. 许杰等人[84]通过对数据源进行几何变换，使得数据以含噪的形式存储于大数据系统中，须经授权后才能逆变换还原，若未经授权访问，得到的数据并非原始数据，实现了隐私保护. 鉴于加密存储不仅增加了额外开销，且被破解后会造成图像全部信息泄露，吕骁博等人[85]提出了图像分割存储的方式，将分割后的图像分别存储于不同的云服务器中，从而降低了泄密的危险，提高了隐私保护的可靠性.

多媒体内容隐私保护技术从最初简单直接的数据移除替换发展到能兼顾视频场景理解，实现二者的平衡，虽然取得了一系列的成果，但仍存在待解决的问题，具体包括以下 3 个方面：

1) 隐私区域的提取困难. 首先，隐私的概念比较模糊，并且其所涵盖的语义信息并不固定. 如何正确提取出真正的隐私区域是需要解决的难题. 即使是单纯提取人脸等显著区域，在实际的应用环境中也是有难度的.

2) 实时性的要求高. 在实际的应用场景中，如监控系统下，对隐私保护系统的速度要求很高. 在保证隐私安全和场景质量的前提下寻找低复杂度的技术，使系统满足实时性的要求仍需研究.

3) 大数据给隐私保护带来更多难题. 大数据所包含的场景之复杂、范围之广泛、用户数量之多都对隐私保护技术的应用实现造成了新的问题. 以访问控制为例，大量的数据信息使用户的身份识别与权限划分变得更加难以预估，需求类型的多样化使得统一描述成为更具挑战性的难题.

6 结论和展望

本文只是对当前典型的多媒体安全研究进行了分析，可以看出新兴的多媒体信息安全研究正在蓬勃发展. 当前大数据、物联网、互联网＋的兴起带来了更为严峻的安全挑战，相比传统信息安全，多媒体信息安全是尚未成熟的新生代，还需要应用需求的引导、合理的评价机制、成熟的理论模型、公开的测试数据和指标体系建设等等，同时多媒体信息安全从科学研究、技术走向现实仍然具有非常大的挑战.

参 考 文 献

[1] 方滨兴，贾焰，韩毅. 社交网络分析核心科学问题、研究现状及未来展望[J]. 中国科学院院刊，2015，30(2)：187-199

[2] Jian Y，Yongfeng H. A method to build subliminal channel in streaming media with multiple steganography methods [C] //IEEE China Summit & Int Conf on Signal and Information Processing (ChinaSIP)，Piscataway，NJ：IEEE，2013：447-451

[3] Xu Shujiang，Chen Xiubo，Wang Lianhai，et al. A novel quantum information hiding protocol based on entanglement swapping of high-level Bell states [J]. Chinese Physics B，2015，24(5)：050306

[4] 孔祥维，郭艳卿，王波. 多媒体信息安全[M]. 北京：科学出版社，2014

[5] Holub V，Fridrich J. Challenging the doctrines of JPEG steganography [C] //Proc of the SPIE 9028，Media Watermarking，Security，and Forensics 2014. San Francisco，CA：SPIE，2014：1-8

[6] Filler T，Judas J，Fridrich J. Minimizing additive distortion in steganography using syndrome-trellis codes [J]. IEEE Trans on Information Forensics and Security，2011，6(3)：920-935

[7] Holub V，Fridrich J，Denemark T. Universal distortion function for steganography in an arbitrary domain [J]. EURASIP Journal on Information Security，2014，2014(1)：1-13

[8] Guo L，Ni J，Shi Y Q. Uniform embedding for efficient JPEG steganography [J]. IEEE Trans on Information Forensics and Security，2014，9(5)：814-825

[9] 胡昊然，钱萌. 基于待征码的Imagehide与JPEGX图像隐藏信息检测以及提取[J]. 科技广场，2007，(1)：125-127

[10] Fridrich J. Modern trends in steganography and steganalysis [C] //Digital Forensics and Watermarking. Berlin：Springer，2012：1-1

[11] Zielinska E，Mazurczyk W，Szczypiorski K. Development trends in steganography [J]. Communications of the ACM，2014，57(3)：86-95

[12] Fridrich J，Kodovsky J. Rich models for steganalysis of digital images [J]. IEEE Trans on Information Forensics and Security，2012，7(3)：868-882

[13] Denemark T，Sedighi V，Holub V，et al. Selection-channel-aware rich model for steganalysis of digital images [C] //Proc of the IEEE Workshop on Information Forensic and Security. Piscataway，NJ：IEEE，2014：48-53

[14] Qian Y，Dong J，Wang W，et al. Deep learning for steganalysis via convolutional neural networks [C] //Proc of the SPIE 9409，Media Watermarking，Security，and Forensics 2015. San Francisco，CA：SPIE，2015：1-10

[15] Ker A D，Bas P，Böhme R，et al. Moving steganography and steganalysis from the laboratory into the real world [C] //Proc of the 1st ACM Workshop on Information Hiding and Multimedia Security. New York：ACM，2013：45-58

[16] Ker A D，Pevny T. A new paradigm for steganalysis via clustering [C] //Proc of the SPIE 7880，Media Watermarking，Security，and Forensics III. San Francisco，CA：SPIE，2011：1-13

[17] 黄炜，赵险峰，盛任农. 基于KFD指标聚类的高隐蔽性JPEG隐密分析[J]. 计算机学报，2012，35(9)：1951-1958

[18] Lubenko I，Ker A D. Steganalysis with mismatched covers：Do simple classifiers help? [C] //Proc of the ACM on Multimedia and security. New York：ACM，2012：11-18

[19] Kodovsky J，Sedighi V，Fridrich J. Study of cover source mismatch in steganalysis and ways to mitigate its impact [C] //Proc of the SPIE 9028，Media Watermarking，Security，and Forensics 2014. San Francisco，CA：SPIE，2014：1-12

[20] Li X，Kong X，Wang B，et al. Generalized transfer component analysis for mismatched JPEG steganalysis [C] //Proc of the IEEE Int Conf on Image Processing (ICIP). Piscataway，NJ：IEEE，2013：4432-4436

[21] 汪然，平西建，许漫坤，等. 基于四叉树分割的JPEG隐写分析[J]. 电子与信息学报，2014，36(3)：631-638

[22] Ker A D，Pevny T. A mishmash of methods for mitigating the model mismatch mess [C] //Proc of the SPIE 9028，Media Watermarking，Security，and Forensics 2014. San Francisco，CA：SPIE，2014：1-15

[23] Zeng Likai，Kong Xiangwei，Li Ming，et al. JPEG quantization table mismatched steganalysis via robust discriminative feature transformation [C] //Proc of the SPIE 9409，Media Watermarking，Security，and Forensics 2015. San Francisco，CA：SPIE，2015：1-9

[24] 钱思进，张恒，何德全. 基于图像视觉复杂度计算的分类信息隐藏图像库[J]. 解放军理工大学学报：自然科学版，2010，11(1)：26-30

[25] Cox I J，Kilian J，Leighton T，et al. Secure spread spectrum watermarking for multimedia [J]. IEEE Trans on Image Processing，1997，6(12)：1673-1687

[26] Barni M，Bartolini F，Piva A. Multichannel watermarking of color images [J]. IEEE Trans on Circuits System for Video Technology，2002，12(3)：142-156

[27] Ng T M，Garg H K. Maximum-likelihood detection in DWT domain image watermarking using laplacian modeling [J]. IEEE Signal Processing Letters，2005，12(4)：285-288

[28] Barni M，Bartolini F，Piva A. Improved wavelet-based watermarking through pixel-wise masking [J]. IEEE Trans on Image Processing，2001，10(5)：783-791

[29] Dong P，Brankov J G，Galatsanos N P，et al. Digital Watermarking robust to geometric distortions [J]. IEEE Trans on Image Processing，2005，14(12)：2140-2150

[30] Costa M. Writing on dirty paper [J]. IEEE Trans on Information Theory，1983，IT-29(3)：439-441

[31] Brian Chen，Gregory W. Wornell. Quantization index modulation：A class of provably good methods for digital watermarking and information embedding [J]. IEEE Trans on Information Theory，2001，47(4)：1423-1443

[32] Eggers J J，Bäuml R，Tzschoppe R，et al. Scalar costa scheme for information embedding [J]. IEEE Trans on Signal Processing，2003，51(4)：1003-1019

[33] Li Q，Cox I J. Using perceptual models to improve fidelity and provide resistance to valumetric scaling for quantization index modulation watermarking [J]. IEEE Trans on Information Forensics and Security，2007，2(2)：127-139

[34] Bao P，Ma X H. Image adaptive watermarking using wavelet domain singular value decomposition [J]. IEEE Trans on Circuits and Systems for Video Technology，2005，15(1)：96-102

[35] Celik M，Sharma G，Tekalp A. Lossless watermarking for image authentication：A new framework and an implementation [J]. IEEE Trans on Image Processing，2006，15(4)：1042-1049

[36] Peng F，Li X，Yang B. Adaptive reversible data hiding scheme based on integer transform [J]. Signal Processing，2012，92(1)：54-62

[37] Hu Y，Lee H，Li J. DE-based reversible data hiding with improved overflow location map [J]. IEEE Trans on Circuits System for Video Technology，2009，19(2)：250-260

[38] Wu H T, Huang J. Reversible image watermarking on prediction errors by efficient histogram modification [J]. Signal Processing, 2012, 92(12): 3000-3009

[39] Li X, Zhang W, Gui X, et al. A novel reversible data hiding scheme based on two-dimensional difference histogram modification [J]. IEEE Trans on Information Forensics and Security, 2013, 8(7): 1091-1100

[40] Ou B, Li X, Zhao Y, et al. Pairwise prediction-error expansion for efficient reversible data hiding [J]. IEEE Trans on Image Processing, 2013, 22(12): 5010-5021

[41] Cao X, Du L, Wei X, et al. High capacity reversible data hiding in encrypted images by patch-level sparse representation [OL]. [2015-07-30]. http://ieeexplore.ieee.org/xpls/abs all.jsp? arnumber=7098386

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

[43] Chen W, Shi Y Q, Su W. Image splicing detection using 2-D phase congruency and statistical moments of characteristic function [C] //Proc of the 9th Conf on Security, Steganography, and Watermarking of Multimedia Contents. Berlin: Springer, 2007: 65050R

[44] Farid H, Lyu S. Higher-order wavelet statistics and their application to digital forensics [C] //Proc of the Conf on Computer Vision and Pattern Recognition Workshop. Piscataway, NJ: IEEE, 2008 (8): 94-101

[45] Dong J, Wang W, Tan T, et al. Run-length and edge statistics based approach for image splicing detection [C] // Proc of the 7th Int Workshop on Digital Watermarking. Piscataway, NJ: IEEE , 2009: 76-87

[46] Li Ying, Wang Bo, Kong Xiangwei, et al. Image tampering detection using no-reference image quality metrics [J]. Journal of Harbin Institute of Technology (New Series), 2014, 21(6): 51-56

[47] Johnson M K, Farid H. Exposing digital forgeries in complex lighting environments [J]. IEEE Trans on Information Forensics and Security, 2007, 2(3): 450-461

[48] Popescu A C, Farid H. Exposing digital forgeries by detecting traces of re-sampling [J]. IEEE Trans on Signal Processing, 2005, 53(2): 758-767

[49] LUKÁŠ J, Fridrich J. Estimation of primary quantization matrix in double compressed JPEG images [C] //Proc of the Digital Forensic Research Workshop. Piscataway, NJ: IEEE, 2003: 5-8

[50] Avcıbaş I, Kharrazi M, Memon N, et al. Image steganalysis with binary similarity measures [C] //Proc of the IEEE Int Conf on Image Processing. Piscataway, NJ: IEEE, 2002 (3): 645-648

[51] Wang Bo, Kong Xiangwei, Elisa Bertino, et al. Exposing copy-paste-blur forgeries based on color coherence [J]. Chinese Journal of Electronics, 2009, 18(3): 487-490

[52] 尚世泽，王波，孔祥维，等. 针对 Photoshop 软件中 USM 锐化的取证检测[C] //第九届全国信息隐藏暨多媒体信息安全学术大会会议论文集. 成都：中国电子学会，2010：350-356

[53] Kot A C, Cao H. Digital Image Forensics [M]. Berlin: Springer, 2013: 157-178

[54] Xu G, Shi Y Q. Camera model identification using local binary patterns [C] //Proc of the IEEE Int Conf on Multimedia and Expo (ICME). Piscataway, NJ: IEEE, 2012: 392-397

[55] Choi K S, Lam E Y, Wong K K Y. Automatic source camera identification using the intrinsic lens radial distortion [J]. Optics Express, 2006, 14(24): 11551-11565

[56] San Choi K, Lam E Y, Wong K K Y. Source camera identification by JPEG compression statistics for image forensics [C] //Proc of the IEEE Region Conf on TENCON, Piscataway, NJ: IEEE, 2006: 1-4

[57] Swaminathan A, Wu M, Liu K J R. Nonintrusive component forensics of visual sensors using output images [J]. IEEE Trans on Information Forensics and Security, 2007, 2(1): 91-106

[58] Fridrich J. Sensor Defects in Digital Image Forensic [M]. Digital Image Forensics. Berlin: Springer, 2013: 179-218

[59] Stamm M C, Lin W S, Liu K J. Forensics vs anti-forensics: A decision and game theoretic framework [C] //Proc of the IEEE Int Conf on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2012: 1749-1752

[60] Kalker T, Haitsma J, Oostveen J C. Issues with digital watermarking and perceptual hashing [C] //Proc of the ITCom 2001: Int Symp on the Convergence of IT and Communications. International Society for Optics and Photonics. Denver: SPIE, 2001: 189-197

[61] 牛夏牧，焦玉华. 感知哈希综述[J]. 电子学报，2008，36(7)：1405-1411

[62] 张维克，孔祥维，尤新刚. 安全鲁棒的图像感知哈希技术[J]. 东南大学学报：自然科学版，2008，37(A01)：188-192

[63] Bhargav-Spantzel A, Squicciarini A, Bertino E, et al. Biometrics-based identifiers for digital identity management [C] //Proc of the 9th Symp on Identity and Trust on the Internet. New York: ACM, 2010: 84-96

[64] 丁凯孟，朱长青. 一种用于遥感影像完整性认证的感知哈希算法[J]. 东南大学学报：自然科学版，2014，44(4)：723-727

[65] 丁凯孟，朱长青，卢付强. 基于自适应格网划分的遥感影像感知哈希认证算法[J]. 武汉大学学报：信息科学版，2015，40(6)：716-720

[66] Hu D, Su B, Zheng S, et al. Security and privacy protocols for perceptual image hashing [J]. International Journal of Sensor Networks, 2015, 17(3): 146-162

[67] Breitinger F, Liu H, Winter C, et al. Towards a process model for hash functions in digital forensics [C] //Digital Forensics and Cyber Crime. Berlin: Springer, 2014: 170-186

[68] Wang X, Pang K, Zhou X, et al. A visual model based perceptual image hash for content authentication [J]. IEEE Trans on Information Forensics and Security, 2015, 10(7): 1336-1349

[69] 冯登国，张敏，李昊. 大数据安全与隐私保护[J]. 计算机学报，2014，37(1)：246-258

[70] Mazurek M L, Liang Y, Melicher W, et al. Toward strong, usable access control for shared distributed data [C] //Proc of the 12th USENIX Conf on File and Storage Technologies. USENIX Association. Piscataway, NJ: IEEE, 2014: 89-103

[71] Ra M R, Govindan R, Ortega A. P3: Toward privacy-preserving photo sharing [C] //Proc of NSDI. New York: ACM, 2013: 515-528

[72] Yuan L, Korshunov P, Ebrahimi T. Privacy-preserving photo sharing based on a secure JPEG [C] //Proc of the 3rd Int Workshop on Security and Privacy in Big Data (BigSecurity 2015). Piscataway, NJ: IEEE, 2015: 1-6

[73] 佟玲玲，李扬曦，黄文廷. 视频隐私保护技术综述[J]. 通信学报，2013，34(8)：154-160

[74] Senior A, Pankanti S, et al. Blinkering surveillance: Enabling video privacy through computer vision, RC22886 [R]. New York: IBM, 2003

[75] 熊金波，姚志强，马建峰，等. 视频数据库多级访问控制[J]. 通信学报，2012，33(8)：147-154

[76] Martin K, Plataniotis K N. Privacy protected surveillance using secure visual object coding [J]. IEEE Trans on Circuits and Systems for Video Technology, 2008, 18(8): 1152-1162

[77] Newton E, Sweeney L, Main B. Preserving privacy by de-identifying face images [J]. IEEE Trans on Knowledge and Data Engineering, 2005, 17(2): 232-243

[78] Chen D, Chang Y, Yan R, et al. Tools for protecting the privacy of specific individuals in video [J]. EURASIP Journal on Advances in Signal Processing, 2007(1): 1-9

[79] Korshunov P, Ebrahimi T. Using face morphing to protect privacy [C] //Proc of the IEEE Int Conf on Advanced Video and Signal Based Surveillance (AVSS). Piscataway, NJ: IEEE, 2013: 208-213

[80] Schiff J, et al. Respectful cameras: Detecting visual markers in real-time to address privacy concerns [C] //Proc of the Protecting Privacy in Video Surveillance. Berlin: Springer, 2009: 65-89

[81] Zhao J, Cheung S C. Multi-camera surveillance with visual tagging and generic camera placement [C] //Proc of the ACM/IEEE Int Conf on Distributed Smart Camera. Piscataway, NJ: IEEE, 2007: 259-266

[82] Luo Y, et al. Anonymous subject identification in privacy-aware video surveillance [C] //Proc of the IEEE Int Conf on Multimedia and Expo. Piscataway, NJ: IEEE, 2010: 83-88

[83] 朱旭东，李晖，郭祯. 云计算环境下加密图像检索[J]. 西安电子科技大学学报，2014，41(2)：151-158

[84] 许杰，聂大成，李明桂，等. 基于几何变形的大数据安全隐私保护方法[J]. 通信技术，2015，48(5)：602-606

[85] 吕骁博，郭耀，陈向群. 基于分割的数字图像云存储机制[J]. 计算机研究与发展，2014，51(5)：1129-1135

孔祥维

博士，教授，博士生导师，主要研究方向为多媒体信息安全、数字图像处理和识别、大数据下的多媒体语义理解、多媒体知识管理和商务智能、多源信息感知和信息融合等.

王　波

博士，讲师，主要研究方向为数字图像取证、信息隐藏与信息隐藏分析.

李晓龙

博士，讲师，主要研究方向为信息隐藏、信息隐藏分析以及可逆信息隐藏等.