

图像隐写分析算法

研究及其应用

刘祖根 / 著

TUXIANG YINXIE FENXI SUANFA
YANJIU JI QI YINGYONG



西南交通大学出版社

云南财经大学前沿研究丛书

云南财经大学博士学术基金资助出版

云南省自然科学基金项目“稳健的图像隐写分析算法研究”(2011FZ148)

云南省哲学社会科学规划课题“图书馆数字资源共享和永久保存研究”(YB2012080)

云南财经大学校级课题“数字图书馆学术网络构建算法研究”(YC2012D11)

图像隐写分析算法

研究及其应用

TUXIANG YINXIE FENXI SUANFA
YANJIU JI QI YINGYONG

刘祖根 / 著

西南交通大学出版社

• 成 都 •

图书在版编目 (C I P) 数据

图像隐写分析算法研究及其应用 / 刘祖根著. —成都: 西南交通大学出版社, 2016.1
(云南财经大学前沿研究丛书)
ISBN 978-7-5643-4408-5

I. ①图… II. ①刘… III. ①电子计算机—密码术—研究 IV. ①TP309.7

中国版本图书馆 CIP 数据核字 (2015) 第 284615 号

云南财经大学前沿研究丛书

图像隐写分析算法研究及其应用

刘祖根 著

责任编辑 孟秀芝

封面设计 墨创文化

出版发行 西南交通大学出版社
(四川省成都市二环路北一段 111 号
西南交通大学创新大厦 21 楼)

发行部电话 028-87600564 028-87600533

邮政编码 610031

网 址 <http://www.xnjdcbs.com>

印刷 成都中铁二局永经堂印务有限责任公司

成品尺寸 165 mm × 230 mm

印 张 8.75

字 数 151 千

版 次 2016 年 1 月第 1 版

印 次 2016 年 1 月第 1 次

书 号 ISBN 978-7-5643-4408-5

定 价 52.00 元

图书如有印装质量问题 本社负责退换

版权所有 盗版必究 举报电话: 028-87600562

前 言

本书是著者结合博士学位论文和最近几年相关研究完善而成的。自1996年第一届国际信息隐藏会议召开以来,以图像为载体的隐写分析研究在国内外学者的多年努力下,取得了丰硕的成果,然而传统研究距离实际应用的高检测性能和强鲁棒性等要求,仍然任重而道远。本书是对根据隐写分析算法的这些实际应用中的要求所做研究工作和有益尝试的总结。

隐写术可用于实现隐蔽通信,该技术可能被应用于军事通信,甚至为不法分子所利用。隐写分析是能检测隐藏信息,获知其长度,甚至抽取隐藏信息的技术。研究隐写分析算法对防止机密数据通过隐写术泄漏、确保危害性信息不借助隐写术传递非常重要,因此获得了广泛关注。

本书研究了以图像为载体的隐写分析算法及其应用。在研究隐写术、隐写分析理论和算法的基础上,应用“载体相关”的隐写分析思想,选择具有代表性的有损压缩类图像中 JPEG、原始类图像中 BMP 和有限调色板类图像中 GIF 三种格式图像为载体,研究和提出了三类隐写分析算法。本书主要研究成果和工作为:

(1) 提出了基于“多域特征”的 MDF_C (Multi-Domain Features in Color images) 和 MDF_G (Multi-Domain Features in Gray images) 隐写分析算法。多域特征使用 DCT 域和空域统计量代表原始图像及其校准图像的函数关系。MDF_C 算法在 DCT 域计算彩色 JPEG 图像中交流系数的一维和二维统计量,在空域抽取原始 JPEG 图像的解压图像内各颜色信道中和颜色信道两两之间 DCT 块边界两侧两个像素颜色之差的一维统计量以及相邻像素位置的“颜色差值对”的二维统计量。采用相同算法从校准图像抽取如上统计量,并使用曼哈顿公式从两类图像的相同统计量计算得到 36 维特征向量。MDF_G 算法在 DCT 域抽取原始图像和校准图像中相同 DCT

位置交流系数的二维统计量以及两种图像中两相邻 DCT 块内同一位置交流系数的二维统计量；在空域抽取两种解压图像亮度的二维统计量和一阶偏微分的二维统计量，求统计量的“直方图特征函数”的“块中心”得到 26 维特征向量。

使用支持向量机设计的隐写分析算法的实验结果表明，同传统隐写分析算法比较，MDF_C 算法的误检率降低了 0.54%~48.23%，误检率大致相等时，检测率提高了 0.47%~66%；MDF_G 算法的误检率降低了 0%~24.99%，误检率大致相等时，检测率提高了 1.24%~67.81%；MDF_C 和 MDF_G 算法对多种隐写术的总体检测性能比传统算法更高。

(2) 提出了基于“高阶微分特征”的 HODF_G (High-Order Differential Features in Gray images)、HODF_C (High-Order Differential Features in Color images) 和 PCF (Principal-Component Features) 隐写分析算法。HODF_G 算法首先计算像素亮度、一阶和二阶全微分的直方图；然后计算像素亮度，一阶、二阶和三阶偏微分，以及一阶和二阶全微分等 6 个对象在图像中两相邻空间位置的共生矩阵；计算得到一个 30 维特征向量。HODF_C 算法计算彩色 BMP 图像各颜色信道中上述统计量，并计算两颜色信道之间的统计量，得到一个 136 维的特征向量。PCF 算法使用“主元分析法”对 HODF_C 算法的 136 维特征向量降维得到主元特征。以支持向量机为分类算法实现隐写分析算法。

实验结果表明，同传统隐写分析算法比较，HODF_G 算法的误检率降低了 2.13%~29.63%，误检率大致相等时，检测率提高了 3.56%~14.13%；HODF_C 算法的误检率降低了 5.42%~18.63%，检测率有不同程度的提高。PCF 算法表现出更好的检测性能和鲁棒性。

(3) 提出了基于“图像子块颜色复杂度”的 CCSB (Color Complexity in Sub-Block)、CCSB_DP (Color Complexity in Sub-Block Differentiated with Parity) 和 CCSB_CNC (Color Complexity in Sub-Block Classified with Number of Different Colors) 隐写分析算法。CCSB 算法将“图像子块的颜色复杂度”定义为一个像素索引值与它的邻近像素索引值之差的绝对值之和。将“颜色复杂度”区分成正、负颜色复杂度，并将两者划分为主、次

两类；然后从图像中抽取上述多种“颜色复杂度”的一维统计量和二维统计量，计算得到 28 维特征向量。CCSB_DP 算法根据“最佳奇偶匹配”法将 CCSB 算法中的统计量分为 2 类并计算得到 56 维特征向量。CCSB_CNC 算法根据“图像子块中不同颜色数”将 CCSB 算法的统计量分布到更高维统计空间，计算得到 44 维特征向量。以支持向量机为分类算法设计隐写分析算法。

实验结果表明，同传统隐写分析算法比较，CCSB 算法的误检率大致降低了 0.07% ~ 81.17%，误检率大致相等或更低时，检测率提高了 11.24% ~ 50.78%。CCSB_DP 算法的检测性能略逊于 CCSB 算法，而 CCSB_CNC 算法表现更优。

(4) 以多类异源高分辨率图像为载体，对多种隐写分析算法进行了增强其鲁棒性的实验研究。

作者

2015 年 1 月

目 录

1 绪 论	1
1.1 研究背景	1
1.2 信息隐藏	1
1.3 隐写术和隐写分析	4
1.4 主要工作	7
1.5 本书结构	10
2 图像隐写分析基础理论研究	11
2.1 隐写术安全模型	11
2.2 隐写分析模型与理论	18
2.3 图像盲隐写分析研究的基础	24
2.4 小 结	31
3 有损压缩图像隐写分析算法研究	32
3.1 有损压缩图像	32
3.2 以 JPEG 图像为载体的隐写术	34
3.3 以 JPEG 图像为载体的传统隐写分析算法	37
3.4 基于多域易变特征的 JPEG 图像隐写分析算法	39
3.5 多域特征向量在灰度图像中的应用	50
3.6 小 结	56
4 原始类图像隐写分析算法研究	58
4.1 原始类图像格式	58
4.2 以 BMP 图像为载体的隐写术	58
4.3 以 BMP 图像为载体的传统隐写分析算法	60
4.4 灰度图像中基于高阶微分的隐写分析算法	62
4.5 彩色图像中基于高阶微分的隐写分析算法	71
4.6 小 结	86

5 有限调色板图像隐写分析算法研究	87
5.1 有限调色板图像格式和存储表示	87
5.2 以 GIF 图像为载体的隐写术	87
5.3 以 GIF 图像为载体的传统隐写分析算法	91
5.4 基于图像子块颜色复杂度的隐写分析算法	91
5.5 基于图像子块颜色复杂度的隐写分析算法的改进	99
5.6 小 结	106
6 鲁棒隐写分析算法研究	108
6.1 图像库间操作	109
6.2 图像库内操作	110
6.3 图像库间与库内操作的综合应用	111
7 总结与展望	112
7.1 隐写分析研究总结	112
7.2 隐写分析研究展望	114
附录 缩略语	117
参考文献	119

1 绪 论

1.1 研究背景

计算机科学技术和网络通信技术的迅猛发展,使人类活动依赖信息技术进步的程度日益加深。“科学技术”这把“双刃剑”在给人类带来便利的同时,也带来了隐患:不法分子可能利用计算机与通信技术从事危害国家和他人的活动。信息安全已成为现代人类社会不得不面对的重大问题。

信息安全指信息网络的硬件、软件及其系统中的数据受到保护,不因偶然的或者恶意的原因而遭到破坏、更改和泄露;使系统连续、可靠和正常地运行,而不中断信息服务。从广义上讲,凡是涉及网络中信息的保密性、完整性、可用性、真实性和可控性的相关技术和理论都属于信息安全的研究领域。

信息隐藏 (information hiding) 是信息安全的子学科。作为信息隐藏的重要分支,隐写术 (steganography) 在载体 (cover) 的冗余信息中隐藏信息,或者改变载体中不足以引起人类感官怀疑的信息实现信息隐藏^{[1]-[5]};而隐写分析 (steganalysis) 技术用于检测隐藏信息的存在、获取隐藏信息时采用的隐写术算法相关信息,甚至从隐密媒体 (stego-media) 中抽取隐藏的的信息。由于隐写术可能被用于实现机密信息的隐蔽传递,造成泄露机密数据或者隐蔽发布不法信息的严重后果。因此,从保护机密信息不外泄和防止隐写术用于犯罪活动的角度而言,研究隐写分析技术尤为重要。

1.2 信息隐藏

1.2.1 信息隐藏研究的历史与现状

信息隐藏是当前受到较多关注和发展较快的一个研究领域。国际信息

隐藏会议自第一、二届分别于 1996 年在英国剑桥大学和 1998 年在美国俄勒冈州召开以后, 每年召开一次。IEEE 期刊关于通信方面的特刊和学报每年刊登相关文章。SPIE 和 IEEE 的一些重要国际会议也开辟了相关专题。从公开发表的文献来看, 由于有美国军方的支持和大公司的介入, 信息隐藏研究发展得非常迅速, 对信息隐藏在数字产品的著作权保护方面(或称为数字水印)的应用研究较多^{[6]~[12]}。以麻省理工学院媒体实验室为代表的一批研究机构和企业已在美国申请了数字水印方面的专利。美国 1998 年政府报告中出现了第一份有关图像数据隐藏的报告。目前已有政府部门、大学和知名企业等机构支持或开展数字水印研究。自“9·11”事件中据称有恐怖组织利用隐写术发布秘密指令以后, 众多研究者在军方支持下开展了隐写术和隐写分析方面的研究^{[13]~[17]}。我国于 1999 年 12 月召开了第一届信息隐藏学术会议。2000 年 1 月, 由国家“863”智能机专家组和中科院自动化所模式识别国家重点实验室组织召开了数字水印学术会议。目前, 有很多研究者从事数字水印、隐写术和隐写分析方面的研究^{[18]~[31]}。

1.2.2 隐写术和信息隐藏

1.2.2.1 隐写术在信息隐藏学科的位置

信息隐藏是指将信息隐藏在宿主对象中实现安全数据传输的应用。隐藏信息对宿主对象造成的影响不应让人类感官察觉到明显差异。信息隐藏领域随着 Internet 和其他数字通信网络的出现取得了很大的进步; 多种数字媒体如文本、图像、音频和视频的发展进一步为信息隐藏研究注入了新的活力。军事和情报部门重视的“通信保密”和“秘密通信”方面也引起了一些人的关注。信息隐藏应用包括信息认证、广播监测、所有权鉴定、所有权证明、版权控制、非法拷贝的泄漏源追踪、标题嵌入、产品的无缝升级和隐蔽通信等^[1]。

信息隐藏主要分为 4 个分支^[1]: 隐蔽信道、隐写术、匿名通信和版权标志(如图 1-1 所示)。

① 多级安全的系统中既非专门设计也不打算用来传输消息的通信路径称为隐蔽信道。

② 隐写术这个名词源于 15 世纪的希腊文, 是“被掩盖的笔迹”的意思。语言上的隐写术类似于中国古代的“藏头诗”。技术上的隐写术通过改变数字媒体中的冗余信息或者改变其中不足以引起人类感官怀疑的信

息实现信息隐藏。

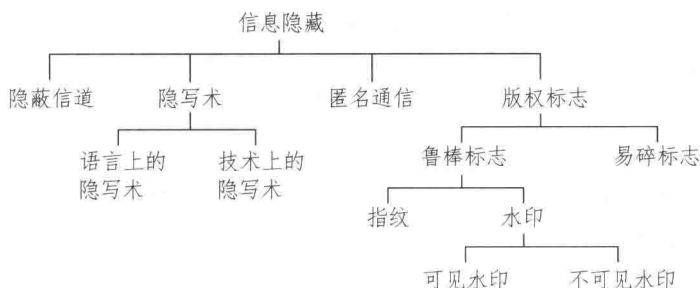


图 1-1 隐写术在信息隐藏学科中的位置

③ 匿名通信，即寻找各种途径来隐藏通信的发送者和接收者。

④ 版权标志的应用主要分为数字水印和指纹识别两类。数字水印在数字媒体中包含所有者信息和数字时间戳，通常用于版权保护。数字水印嵌入的信息本身不存在任何秘密，但应用者希望它在载体中永久存在。同数字水印紧密相关的是指纹识别，指纹^{[32][33]}（或序列号）唯一地标识某个数据集的特定用户，它使所有者有可能从数据集的一次非授权使用跟踪到该用户。

1.2.2.2 隐写术同其他技术的区别和联系

（1）隐写术同密码学不一样

密码学对通信内容进行加密处理，通信双方需要密钥获取秘密信息，没有密钥的第三方无法获知通信内容，却知道双方正在通信这样的事实；隐写术则要隐藏通信这个事实的存在。密码学隐藏信息的含义；而隐写术隐藏信息本身。因此，隐写术同密码学不是竞争的关系，两者可以相互补充^{[34][35]}，进一步保证隐蔽通信的安全性^[36]。

（2）隐写术同数字水印的异同点

隐写术在通信双方之间建立隐蔽通道，第三方无法察觉隐蔽通道的存在；数字水印则将版权或者其他认证信息嵌入载体，信息本身可能不是秘密的。两者的相同之处是它们都隐藏数据并且通常使用相似的算法。隐写术和数字水印的不同之处有三点：第一，载体是数字水印要保护的對象，对隐写术而言载体的功能只是起到诱骗的作用而没有实际价值；第二，数字水印的存在不需要隐藏，隐写术中隐藏信息的存在则不能为人所知；第三，数字水印通常需要 120 位信息容量，而隐写术因为通信的目的希望获得尽可能大的信息容量。

(3) 隐写术同信号处理、通信、游戏理论和密码学等领域相关

信号处理和通信领域的大量技术既可用于设计隐写术^{[37][38]},又可用于消除隐藏的信息。游戏理论已被用于一些特例研究^{[39]-[41]}。密码学使用密钥保护信息,然而密码学和隐写术对密钥的功能要求完全不一样:信息的保密性是密码学的主要目标,但它不是隐写术实现的必要条件;此外,密码学已伴随着 Shannon 的标志性论文^[42]在信息理论领域得到了高度重视,而隐写术在数学和实现技术两方面仍不成熟。

1.3 隐写术和隐写分析

没有隐藏信息的原始数字媒体称为“载体”(cover),包含隐藏信息的数字媒体称为“隐密媒体”(stego-media)。信息隐藏技术应满足三个要求:不可觉察性、鲁棒性和容量^[43]。不可觉察性,即隐藏信息不会导致数字媒体的明显降质,使人类视觉和听觉器官,甚至利用计算机统计分析也无法觉察到隐密媒体相对于载体的变化。鲁棒性,要求信息隐藏技术有一定的抗干扰能力,即对媒体的数字处理和变换不会导致隐藏信息无法恢复。容量,即在保证不可觉察性和鲁棒性的条件下载体可隐藏的最大信息量。

隐写分析的设计基于下面两个假设:①如果能够辨别隐藏信息的存在,那么可能确定用于隐藏它的算法名称。②如果能够识别隐写术,则可能使用该算法抽取出隐藏的信息。

按照隐写分析的目的和分析工作的难易程度,隐写分析可分为检测并确定可疑对象、判定隐写术采用的算法、恢复密钥、估计隐藏信息长度^[44]、定位和抽取隐藏信息等部分^[45]。从控制和切断隐写术实现的隐蔽通信渠道这一目的出发,可按照如下两条标准对隐写分析算法进行分类:隐写分析包含的不同功能成分和可攻击的隐写术多少。根据隐写分析包含的功能,可将它分为“被动”和“主动”两类。被动隐写分析只检测媒体中是否存在隐藏信息;主动隐写分析则还要获知隐写术算法相关信息和抽取隐藏信息。按照隐写分析能攻击的隐写术多少,可将它分为“专用”隐写分析和“盲”隐写分析两类。专用隐写分析算法针对特定隐写术隐藏信息的技术特点以及该隐写术对载体统计特性造成的特定变化而设计。盲隐写分析算法研究多种隐写术对载体统计特性造成的一般意义上的变化,因而可用于攻击多种隐写术。对隐写术的攻击方法有如下 3 种:

① 恶意攻击和破坏隐密媒体,使隐藏信息无法恢复。例如在隐密图像

的嵌入域再一次嵌入一些信息,达到覆盖和破坏原有隐藏信息的效果。

② 检测媒体中是否存在隐藏信息。

③ 抽取隐写术和隐藏数据的相关信息。从隐密媒体中获取所采用的隐写术名称和算法细节;分析隐藏数据的长度和位置,估计隐藏数据时采用的密钥,甚至抽取出隐藏数据。

其中,方法①不在本书的研究范围内;方法②和方法③分别是“被动”隐写分析和“主动”隐写分析。

根据不同的标准对隐写分析分类可为相关研究提供方向性指导。

1.3.1 根据已知信息分类的隐写分析

参照密码分析学中的分类方法,根据能获得的对象可将隐写分析分为6类:

① 唯隐密对象分析:只能获得隐密对象时的隐写分析。

② 已知载体分析:可同时获得载体和隐密对象时的隐写分析。

③ 已知隐藏信息分析:已知部分隐藏信息,而不知道其他具体的隐藏情况(如隐写术、机密信息以及该媒体中是否包含隐藏信息)时的隐写分析。

④ 选择媒体分析:已知隐密媒体和一些隐写术时的隐写分析。

⑤ 选择隐藏信息分析:将选择的隐藏信息通过一些隐写术嵌入载体,确定所得隐密对象中的模式特征,用于判定所怀疑隐密对象中采用的隐写术。

⑥ 已知隐藏分析:隐写术、载体和隐密对象均已知时的隐写分析。

隐写分析系统检测数字媒体中是否存在隐藏信息并抽取相关信息时,无法得到该媒体的载体,更无法假定隐密对象是采用哪一种隐写术隐藏信息得到的。因此,大多数隐写分析系统属于唯隐密对象攻击。

1.3.2 根据信息隐藏域分类的隐写分析

隐写术既可能将信息隐藏在空域,也可能将信息隐藏在变换域。隐写分析需要在信息嵌入域抽取特征以研究隐写术对数字媒体造成的影响,隐写分析因而可分为空域隐写分析和变换域隐写分析。利用“隐密媒体”中人工嵌入的随机比特平面和载体中自然形成的随机比特平面之间的差异性,可设计区分两种媒体的空域隐写分析算法。变换域主要包括离散余弦变换域(DCT)、离散傅里叶变换域^[46](DFT)和离散小波变换域^{[47][48]}(DWT)。根据变换域隐写术的特点对媒体做相应变换,取出隐藏信息过程

中受影响较大的频域系数,才能实现可靠的隐写分析。将隐写分析分为多个域,有可能在数据嵌入域得到容易被隐写术改变的特征,结合多个嵌入域的特征实现高效的隐写分析算法。

1.3.3 根据实现技术分类的隐写分析

按照隐写分析算法的实现技术,可将它分为感官检测、统计检测和特征检测^{[49][50]}三类。“感官检测”利用人类感觉器官感知和分辨噪音的能力来区分载体和“隐密”数字媒体。“统计检测”比较载体中固有的和隐密媒体中可检测到的统计分布,从中找出两者的差别以实现检测。实现统计检测的关键是获得可描述载体统计特性的完善统计模型;然而,至今仍没有图像等载体的通用数学模型,因此很难实现通用的统计检测算法。“特征检测”根据隐写术对数字媒体特征的改变进行检测,这种特征既可以是感官特征,也可以是统计特征。一般来说,感官特征比较明显而较易检测;统计特征则要根据隐写术隐藏信息过程中采用的变换操作进行数学推理分析,确定载体和隐密媒体之间的可度量特征差异。使用特征检测隐写术通常需要依赖对特征差异的统计分析。

1.3.4 根据载体类型分类的隐写分析

隐写术的载体既可以是文本、图像、音频和视频等数字媒体文件,也可以是可执行的二进制文件等非数字媒体。设计隐写分析算法时,需要研究载体特性和相关隐写术的技术特点。

隐写术的不同载体具有不同的特点。以文本为载体的隐写术通常使用文字间距、文字间的相对位置高低,甚至文字大小和颜色等格式信息实现信息隐藏;因而以文本为载体的隐写术具有信息容量小和鲁棒性差的特点(例如将文本打印出来,重新使用 OCR 软件识别文字,即可去除文本中隐藏的信息)。音频和视频文件占用的存储空间较大、传输所需时间多;传输较多音频和视频文件易引起网络监控的怀疑,事实上违背了采用隐写术实现“隐蔽”通信的目的。因此,网络环境中使用音频和视频为载体的隐写术相对较少。以图像为载体的隐写术兼有以上两类载体的特点:信息容量大并且占用的存储空间相对音频和视频文件较小。在网站上或者电子邮件附件中添加一幅图像很少引起怀疑,因此,以图像为载体的隐写术和隐写分析的研究和开发获得了军方、广大研究者和商业公司的青睐。从应用

的广泛性和维护安全可信的网络信息环境来说,图像隐写分析研究成为信息安全研究的一个重要方面。以软件为载体的隐写术也存在信息容量较小的问题,相关技术主要用于“软件水印”保护软件作者的版权而非隐蔽通信。

为使嵌入信息对载体统计特性的影响尽可能小,不同格式媒体上的隐写术利用了多种格式相关的信息。因此可根据隐写术针对的载体类型对隐写分析算法进行分类,从而在隐写分析算法设计过程中引入载体格式相关的特征信息。隐写分析算法可分为文本隐写分析、图像隐写分析、音频隐写分析^[51]和视频隐写分析等多种类型。对于每种载体,还可根据其文件存储方式分类。例如,存在3种基本的图像文件格式:矢量类(vector)、光栅类(raster)和元数据类(meta format)图像。

① 矢量类图像,以矢量格式存储的图像被定义为数学上的直线、曲线和圆等,用数学公式存储矢量类图像,如EPS和EMF格式。

② 光栅类图像,由许多小点组成,即用颜色不同的像素组成。照片、扫描图像或绘画等均可存储为光栅类图像。

③ 元数据类图像,是矢量类图像和光栅类图像的复合体。一幅元图像包含矢量和光栅两方面信息。图像会以光栅形式存储,而描述性数字、文字或箭头等信息均以矢量信息存储。扫描类图像即元图像的例子。

光栅类图像又可分为3种:

① 原始类图像,即未经压缩的位图类图像,如BMP和TIFF格式的图像。

② 调色板类图像,将像素对应颜色存放在调色板中,图像数据是调色板中颜色索引值的一类图像,如GIF、PNG和索引类BMP图像。

③ 有损压缩类图像,以降低图像表示精确度为代价,达到减少表示图像所需数据的目标的一类图像,如JPEG和JPEG 2000格式。

矢量类图像不适合隐藏信息。本书主要研究原始类、调色板类和无损压缩类这3类图像上的隐写术和隐写分析算法。

1.4 主要工作

本书研究以图像为载体的隐写分析算法。在研究隐写术、隐写分析理论和算法的基础上,应用“载体相关”的隐写分析思想,选择具有代表性的有损压缩类图像中JPEG、原始类图像中BMP和有限调色板类图像中GIF三种格式图像为具体载体对象,研究和提出了三类隐写分析算法。本书主要研究成果和工作为:

(1) 提出了基于“多域特征”的 MDF_C 和 MDF_G 隐写分析算法

多域特征由代表原始图像及其校准图像函数关系的 DCT 域和空域统计量计算得到。MDF_C (Multi-Domain Features in Color images) 算法的 DCT 域统计量包括所有 DCT 交流系数的全局直方图、几个特定 DCT 位置交流系数的直方图、特定 DCT 系数的直方图相邻 DCT 块间同一 DCT 位置两个系数差值的直方图, 以及相邻 DCT 块中同一 DCT 位置两个系数的共生矩阵。MDF_C 算法的空域统计量包括解压图像每个颜色信道中 DCT 块边界两侧像素两个颜色之差的直方图、颜色信道两两之间 DCT 块边界两侧像素的两个颜色之差的直方图, 每个颜色信道中相邻像素位置的“颜色差值对”的共生矩阵, 以及颜色信道两两之间相邻像素位置的“颜色差值对”的共生矩阵。采用相同算法从校准图像抽取如上统计量, 并使用曼哈顿公式分别从两类图像的相同统计量计算得到 36 维的特征向量。MDF_G (Multi-Domain Features in Gray images) 算法的 DCT 域统计量包括原图像中 DCT 交流系数和校准图像中同一 DCT 位置交流系数的全局共生矩阵、原图像中几个特定 DCT 位置的交流系数和校准图像中同一 DCT 位置交流系数的共生矩阵、两种图像中两相邻 DCT 块内同一位置交流系数的共生矩阵。MDF_G 算法的空域统计量包括原图像及其校准图像两者的解压图像中相同像素位置的亮度共生矩阵和一阶偏微分共生矩阵。求统计量的“直方图特征函数”的“块中心”得到 26 维特征向量。

在 Greenspun 和 CorelDraw 两类图像库上的实验表明该类算法可高效检测 JPEG 图像上多种隐写术。同传统隐写分析算法相比, MDF_C 算法的误检率大致降低了 0.54%~48.23%, 误检率大致相等时, 检测率提高了 0.47%~66%; MDF_G 算法的误检率大致降低了 0%~24.99%, 误检率大致相等时, 检测率提高了 1.24%~67.81%; MDF_C 算法和 MDF_G 算法对多种隐写术的总体性能更高。

(2) 提出了基于“高阶微分特征”的 HODF_G、HODF_C 和 PCF 隐写分析算法

HODF_G (High-Order Differential Features in Gray images) 算法和 HODF_C (High-Order Differential Features in Color images) 算法将微分运算引入隐写分析领域, HODF_G 算法由灰度 BMP 图像抽取 30 维特征, HODF_C 算法由彩色 BMP 图像抽取 136 维特征, 以支持向量机为分类算法构造隐写分析算法。HODF_G 算法抽取特征的过程如下, 首先计算像素亮度, 一阶和二阶全微分的直方图以及高阶偏微分的直方图; 其次统计像素亮度, 一阶、二阶和三阶偏微分, 以及一阶和二阶全微分等 6 个对象在

两相邻空间位置的共生矩阵；最后计算这些统计量的“直方图特征函数”的“块中心”得到特征，组成一个 30 维特征向量。HODF_C 算法首先从一幅彩色 BMP 图像的 R、G 和 B 三个颜色信道内得到 60 维特征；然后在颜色信道之间得到 76 维特征，组成一个 136 维的特征向量。PCF (Principal-Component Features) 算法使用“主元分析法”对 HODF_C 算法的 136 维特征向量降维得到主元特征。

在 CorelDraw 和 Washington 两种图像库对 Cox 和 Piva 两种扩频隐写术的实验说明了该类算法的高效性和鲁棒性。同传统隐写分析算法比较，HODF_G 算法的误检率大致降低了 2.13% ~ 29.63%，误检率大致相等时，检测率提高了 3.56% ~ 14.13%。在 CorelDraw 图像库上，HODF_C 算法对 Cox ($\alpha=0.05$) 和 Piva ($\alpha=0.1$) 隐写术的误检率更高，而检测率大致相等。对 CorelDraw 图像库上其他隐写术和 Washington 图像库，HODF_C 算法的误检率大致降低了 5.42% ~ 18.63%，检测率有不同程度的提高。PCF 算法的检测性能和对不同来源图像库的鲁棒性均胜过传统算法。

(3) 提出了基于“图像子块颜色复杂度”的 CCSB、CCSB_DP 和 CCSB_CNC 隐写分析算法

CCSB (Color Complexity in Sub-Block) 算法将“图像子块的颜色复杂度”定义为一个像素索引值与它的邻近像素索引值之差的绝对值的和。将“颜色复杂度”区分成正、负颜色复杂度，并将两者划分为主、次两类；然后从图像中抽取上述多种“颜色复杂度”的一维统计量和二维统计量，计算得到 28 维特征向量。CCSB_DP (Color Complexity in Sub-Block Differentiated with Parity) 算法根据“最佳奇偶匹配”法将 CCSB 算法中的统计量分为两类并计算得到 56 维特征向量。CCSB_CNC (Color Complexity in Sub-Block Classified with Number of Different Colors) 算法根据“图像子块中不同颜色数”将 CCSB 算法的统计量分布到更高维统计空间，得到 44 维特征向量。基于三类特征向量和支撑向量机构造了三种隐写分析算法。

对 GIF 图像上多种隐写术的实验表明，“图像子块的颜色复杂度”是描述 GIF 载体图像和隐密图像之间函数关系的有效工具。同传统隐写分析算法比较，CCSB 算法的误检率大致降低了 0.07% ~ 81.17%，误检率大致相等或更低时，检测率提高了 11.24% ~ 50.78%。CCSB_CNC 算法的检测性能略逊于 CCSB 算法，而 CCSB_CNC 算法表现更优。

(4) 对图像隐写分析算法的鲁棒性问题进行了研究

在不改变基本的特征抽取算法的基础上，以上述第 (1) 类算法为基