西安电子科技大学

硕士学位论文

基于图像的信息隐藏分析技术研究

姓名：王剑

申请学位级别：硕士

专业：计算机系统结构

指导教师：田玉敏

20070101

摘要

数字图像信息隐藏的分析技术，也称为图像隐写分析(Image Steganalysis)。随着信息安全日益引起人们的关注，隐写分析已经成为信息隐藏领域一个重要的研究方向，它一方面可以促进信息隐藏算法安全性的提高，推动信息隐藏算法的实用化；另一方面可以防止信息隐藏技术的非法应用，防止不法分子利用信息隐藏技术进行犯罪活动，危害国家安全。本文主要研究被动攻击下的通用隐写分析技术，主要工作如下：

1、针对基于加性噪声的空域隐写技术，设计实现了一种通用隐写分析算法。算法分别利用差分直方图分布的高阶统计矩和差分直方图曲线下特殊区域之间面积的比重来量化隐写前后的差分分布差异，利用Fisher分类器检测是否存在隐藏信息。实验结果表明，该算法获得了较好的检测性能。

2、针对 JPEG 压缩域隐写算法，设计实现了一种通用隐写分析算法。算法以 DCT 系数直方图分布的均值、方差、偏度、峰度以及特殊系数值的直方图分布概率作为特征，并利用方差分析对特征进行有效性分析，实验结果证实了所提算法的有效性。

关键词： 信息隐藏　隐写技术　隐写分析技术　Fisher 线性判别

Abstract

Image steganalysis is an art of detecting hiding information in digital image. Recently, steganalysis has been paid a great deal of attention both from the academic research and law enforcement. On one hand, steganalysis can improve security of steganographic algorithm. On the other hand, it can be used to reveal the illegal communication with some steganographic tools. The universal steganalysis for images with passive adversaries are focused on in this thesis. The main contributions are summarized as follows.

1. A universal steganalysis method for the additive noise modelable steganography in the special domain is put forward. Higher order statistics moments of the differential histogram distributions and the ratios of special areas under the differential histogram curves are calculated to characterize the difference between the differential distribution of the cover image and that of the stego-image. Fisher classifier is used to detect whether there exists the hidden information in images. Experimental results show that our technique achieves better detection performance.

2. A universal steganalysis method is designed for the steganography techniques in the JPEG compression domain. The features for the blind classifier are calculated as the mean, variance, skewness, kurtosis of the histogram distribution of the DCT coefficients, and the histogram distribution probabilities of special coefficient values as well. The variance analysis is utilized to identify useful features in steganalysis. Experimental results prove the validity of the method.

Keyword: Information hiding steganography steganalysis FLD

创新性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

本人签名：丁金 日期：2007.1.25

关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。本人保证毕业离校后，发表论文或使用论文工作成果时署名单位仍然为西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。（保密的论文在解密后遵守此规定）

本学位论文属于保密在___年解密后适用本授权书。

本人签名：丁金 日期：2007.1.25

导师签名：田玉敏 日期：2007.1.26

第一章 绪论

信息安全是信息社会急需解决的最重要问题之一，它已成为信息科学领域中一个重要的新兴学科。虽然密码技术一直是保护信息机密性的最有效手段，但是随着网络化、数字化、信息化水平的不断提高，信息安全的内容也在不断地丰富，密码技术在某些方面已经不能够满足信息安全的要求，而且密码的不可破译度主要依靠密码算法的抗攻击能力和密钥的长度，随着计算机计算能力的提高，特别是量子计算和DNA计算时代的即将来临，一切建立在计算复杂性理论基础上的密码的安全性都将面临严峻的挑战。令人欣慰的是，信息隐藏技术的出现和发展，为信息安全的研究带来了新的曙光。近年来，由于各国政府出于国家安全方面的考虑，对密码的使用场合及密码强度都做了严格的限制，这就更加激发了人们对信息隐藏技术研究的热情。

1.1 隐写分析技术的研究背景和意义

随着多媒体技术和网络技术的迅猛发展，互联网上的数字媒体应用正在呈爆炸式的增长。数据压缩和多媒体技术的发展，使得人们能够方便快捷地对数字媒体（数字声音、文本、图像和视频）的原文件进行无限制的任意编辑、修改、拷贝和散布。由此引发出数字媒体的知识产权保护和信息安全的问题，这一问题日益突出，已经成为了数字世界的一个非常重要和紧迫的议题。

信息安全的内涵包括信息的保密性、完整性、真实性和不可否认性[1]。信息的保密性是指除授权者外任何不具备信息享有权的人对信息内容的不可获取性；信息的完整性是指信息在存储和交换等过程中不被篡改和破坏；信息的真实性是指所获取的信息应该来自该信息的真实发出者或拥有者，它是相对于信息的伪造和欺诈而存在的；信息的不可否认性则是指信息的不可抵赖性，它要求信息的真实发出者事后对自己的行为不能予以抵赖。在现代通信技术条件下，如何在获取、存储、交换和享有信息过程中保证信息的保密性、完整性、真实性和不可否认性等便是信息安全所需解决的基本问题。

解决信息安全问题的一个重要技术是密码技术。密码技术是利用密码算法按照一定的安全协议所构成的一种技术。密码算法的理论基础是密码学，密码学是专门设计和分析密码算法并为密码算法提供理论基础的学科。人们常常认为信息

安全的实现可以通过加密来完成。但是采用传统密码学理论开发出来的加解密系统，对于机密文件的处理都是将其加密成密文，使得在网络传递过程中出现的非法拦截者无法从中获取机密信息，从而达到保密的目的。但是这种方法有一个很明显的不足，经典加密技术往往把一段有意义的信息（明文）转换成看起来没有意义的东西（密文），它明确地提示攻击者哪些是重要的信息，容易引起攻击者的好奇和注意，从根本上造成了一种不安全性。并且密文有被破解的可能性，而一旦加密文件经过破解后其内容就完全透明了。即使攻击者破译失败，他们也可以将信息破坏，使得即使是合法的接收者也无法阅读信息内容。采用加密技术的另一个潜在缺点是，随着电脑硬件的迅速发展，以及基于网络实现的具有并行计算能力的破解技术的日益成熟，加密算法的安全性受到了严重挑战，仅仅通过增加密钥长度来增强安全性已经不再是唯一的可行方法。所以，仅仅依靠密码学已经不能完全解决信息安全问题。

近年来，国际上提出一种新的关于信息安全的概念，开发设计出一种不同于传统密码学的技术，即将机密资料信息秘密地隐藏于普通的文件中，然后再通过网络传递散发出去。这样非法拦截者从网络上拦截下来的伪装后的机密资料，并不像传统加密过的文件那样是一堆乱码，而是看起来和其他非机密性的一般资料无异，因此十分容易欺骗非法拦截者。这一点是传统加解密系统所欠缺的，也是信息隐藏(Information Hiding)基本的思想。近年来，出于对数字作品版权保护和个人隐私保护等方面的需求，掀起了信息隐藏技术研究的热潮。信息隐藏技术利用载体信息中具有随机特性的冗余部分，将机密重要信息嵌入到载体信息之中，使其不被其他人发现。在实际应用中，存在冗余信息的载体非常丰富，这一点也在客观上增强了信息隐藏技术的隐蔽性和可行性。这种通过把信息存在本身隐藏起来的技术使得攻击者无从获取秘密信息的位置，从而增强了安全性。

信息隐藏主要有数字水印(Digital Watermarking)和隐写术(Steganography)两大分支[2]。数字水印主要是为了保护知识产权，通过在原始媒体数据中嵌入信息来证实该媒体的所有权归属。数字水印的主要目的不是限制对媒体的访问，而是确保媒体中的水印不被篡改或消除。因此稳健性是数字水印的最基本要求之一。数字水印的稳健性是指水印图像经过一些常见的改变后，水印仍具有较好的可检测性。这些改变包括常见的图像处理（如数据压缩、低通滤波、图像增强、一次抽样、二次量化、A/D 和 D/A 转换等）、几何变换和几何失真（如裁剪、尺度拉伸、平移、旋转、扭曲等）、噪声干扰、多重水印(Multiple Watermarking)的重叠等。对不同的应用场合，要求有不同的稳健性。需要指出的是，存在另一种与稳健水印性质相反的水印，称为易损水印(Fragile Watermarking)，它们被用来证实原始媒体是否被改变过。稳健性在整个水印系统设计中具有非常重要的分量，这也是将隐写术和

数字水印区别对待的原因之一。

隐写术是信息隐藏技术的另一个重要分支。密码学研究如何保护信息的内容，而隐写术研究如何隐藏信息的存在。隐写术的英文单词 Steganography 是由 Trithemius 首先构造出来的，一般认为它来源于希腊文，其含义为“被掩盖的笔迹”(covered writing)。该词的现代含义通常理解为将秘密信息隐藏在载体文件中，在尽可能不引起第三方怀疑的情况下，通过公共通信网发送出去，它更注重的是不可察觉性。隐写术的目的是保护隐藏在载体中的秘密信息。如果第三方察觉出载体数据中含有额外的秘密信息，即使无法破译其具体内容，隐写也已经宣告失败，因为公共通信网的监控者可以中止通信双方的通信并追查隐蔽信息的来源。这种察觉数字媒体中秘密信息的存在性的攻击方法被称作隐写分析(Steganalysis)。由于信息安全日益引起人们的关注，隐写分析也就成为了研究热点之一。

隐写分析是对隐写术的攻击，目的是检测秘密信息的存在以至破坏隐秘通信。隐写分析是解决非法使用隐写术问题的关键技术。近几年来由于恐怖活动猖獗，隐写分析技术引起了很大的关注，获得了较快的发展。据报道[3]，“9.11”恐怖事件的恐怖分子就是利用了数字隐写技术，通过因特网相互传递隐藏的密谋信息。现在因特网上存在很多的隐写工具，而且这些工具大部分是免费的。隐写工具的易用性以及获得上的便利性，使得执法部门非常关心非法资料的流通，迫切地需要检测各种隐形信息的技术，以达到揭示非法消息，打击恐怖主义，预防灾难发生的目的，从而保证国家的安全和社会的稳定。

隐写分析不仅具有重要的应用价值，更具有重要的学术意义。隐写分析研究可以揭示当前隐写术的缺陷，对隐写术的安全性进行测试与评价，促进隐写算法安全性的提高。

1.2 隐写分析技术的发展趋势

随着信息隐藏技术的发展，越来越多的人开始关注隐写分析技术的研究[4]。目前已经有很多学者和研究人员在进行相关的研究工作，也取得了一些进展。然而，目前提出的隐写分析方案一般都是针对特定的隐写算法而言的。也就是说，如果怀疑一个文件中藏有秘密信息，又能大概估计出它可能采用的隐写方法，那么就可以根据相应的方法去进行检测。但是这样就引发了一系列问题，首先，目前提出的隐写算法如此之多，如果要对这些方法一一进行检测，所需花费的计算量是相当大的；其次，目前已经研究出的一些隐写效果比较好的算法的隐蔽性都非常高，如果不知道详细的隐藏算法和准确的隐藏位置，要对其进行检测非常困难。

因此，好的隐写分析方法应该可以脱离具体的隐写算法，从文件信息的特征上进行分析，分析其特征在隐藏信息前后的改变情况，并根据这些改变来判断该文件是否隐藏了信息。目前已有学者在从事这方面的研究，但这是一项艰难的工作，需要借鉴许多其他学科的研究成果，比如说数字图像中的信息隐写，其检测就需要很多图像处理方面的知识，而如果是对声音文件的信息隐藏，在考虑检测算法时就需要了解声音文件的特性。如果能够从统计的角度找到信息隐写对载体特性影响的模型，信息隐写的检测也就会相应降低难度。

信息隐写检测的进一步研究就是隐藏信息的提取和破解。当一个文件被检测出隐藏有秘密信息后，如何将被隐藏的信息提取恢复出来，了解秘密的内容，就是信息隐藏提取要做的工作。在知道了某文件中隐藏了信息后，首先要将秘密信息尽可能完整地得到，其次要从提取出来的信息中恢复秘密内容。要想准确地恢复保密信息的内容，目前还没有行之有效的方法，只能进行各种尝试，可以适当借鉴密码分析技术的一些思路和方法。

信息隐写的检测和提取是今后信息隐藏技术的研究热点，这方面的工作也将会对信息隐藏技术的发展起到积极的促进作用。信息隐藏检测模型的建立也将会促进具有更好隐蔽性的信息隐写算法的研究。

1.3 现阶段基于图像的隐写分析研究的不足及难点

数字图像是实现信息隐藏的优良载体，目前成为了Internet 上实现隐秘通信的主要载体，基于数字图像的隐写分析技术在近几年得到了很大的重视。目前国外已经有部分基于图像的隐写分析系统在军事、国防领域投入了使用，而且初见成效。国内在该方面的研究起步较晚，有一些公司、研究所也从事着这方面的研究，在全国信息隐藏学术研讨会上，有几篇基于图像的隐写分析的科技文章，但相对于国外来说还不够深入。

基于图像的隐写分析技术研究的难点主要表现在：

(1) 基于图像的隐写分析技术是一个综合性、跨学科的研究领域，涉及数字图像、数字信号处理、信息论、概率论与数理统计、小波分析、模式识别等多个学科的知识；

(2) 对于不同的研究客体，同一检测方法的检测效果不同。因此，每种隐写分析算法都具有不同的适用范围，无法找到一个完全通用的算法；

(3) 隐写技术与隐写分析技术是攻与防的关系。随着基于图像的隐写技术的推陈出新，相应的检测就需要不断的进步，提高检测能力；

(4) 隐写分析技术主要针对于隐秘通信的检测，涉及的应用领域大多集中在军事、情报和国防等敏感领域。各国在该技术的研究上都具有较强的保密性，许多先进的研究成果无法获得。即使应用在民用领域，由于涉及版权保护，也属于商业机密内容，技术细节也是无法获取的；

(5) 隐写技术在实际应用中往往与多种技术相互结合，如数据加密技术、数据压缩技术等，更是增加了隐写分析的难度。

现阶段基于图像的隐写分析技术的不足之处主要表现在：

(1) 现有的研究还没有形成一个成熟的理论来描述隐写技术和隐写分析技术，因此多数研究只是在隐写分析技术的局部进行；

(2) 现有的隐写分析技术研究多集中在检测算法的研究方面，或者单纯地从隐写工具和算法的分析入手，分析具体隐藏工具和算法的检测问题，还没有形成基于图像的隐写分析整体知识体系；

(3) 现有的隐写分析算法普遍表现为效率较低，误报率和漏报率相对较高。各检测算法都有很大的局限性，仅适用于较强的假设条件下，没有通用性的检测方法。算法的研究仍处于理论研究阶段，大多数都无法实际应用到检测系统中；

(4) Internet 是基于图像的信息隐写技术应用的重要舞台，因此面向大型网络应用，尤其是Internet 中应用的检测系统是未来研究的重点。

1.4 本文的主要工作和论文安排

本文主要研究被动攻击下的图像通用隐写分析技术，主要的研究内容可归纳如下：

1、针对基于加性噪声的空域隐写技术设计实现了一种通用隐写分析算法。算法分别利用差分直方图分布的高阶统计矩和差分直方图曲线下特殊区域之间面积的比重来量化隐写前后的差分分布差异，从而给出了设计通用隐写分析算法的一般方法。利用该算法对隐写后的图像进行检测的实验结果表明，该算法获得了较好的检测性能。

2、针对JPEG 压缩域隐写算法设计实现了一种通用隐写分析算法。算法以DCT系数直方图分布的均值、方差、偏度、峰度以及特殊系数值的直方图分布概率作为特征，并利用方差分析对特征进行有效性分析，对JSteg、F5 和OutGuess 隐写算法的实验检测结果证实了所提算法的有效性。

本文结构安排共分五章，各章节内容安排如下：

第一章概述了隐写分析技术的研究背景、意义以及研究现状，并分析了现阶段隐写分析研究的不足和难点。

第二章对数字图像隐写分析技术的基本理论和基本概念进行了分析和概括，给出通用隐写分析技术的系统模型和性能评价指标。

第三章针对一类基于加性噪声的空域隐写技术设计实现了一种通用隐写分析算法。

第四章设计实现了一种针对JPEG压缩域算法的通用隐写分析技术。

第五章主要对论文工作进行总结，并对进一步的研究工作提出了展望和设想。

第二章 隐写分析技术

随着数字媒体和通信技术的发展，如何保证通信的安全一直是研究人员关注的目标之一。

通信的安全可以通过以下三种方式来实现：

- 信道安全 通过安全的、不可被监听的特殊信道来实现通信的安全。
- 载体安全 通过某种方式把信息通过载体传输，而通过载体不被关注或者信息无法提取来实现通信的安全。
- 内容安全 通过加密直接把通信内容转换为密文，使得无法被非授权的破译。

一直以来，加密方式都是实现通信安全的主要方式，但是加密方式的弱点很明显，仅仅应用了加密的通信在传输密文的时候很容易引起监听者的注意，从而暴露秘密通信存在的事实。

因此，对通信安全的研究不仅仅停留在对通信内容的加密方面，同时也从通信的传输安全方面进行考虑，尤其是信息隐藏技术。信息隐藏是研究如何将敏感信息通过非敏感的形式进行传输，在信息安全方面的重要性也逐渐凸现。

信息隐藏是信息安全的一个重要领域，信息隐藏是在不对载体信号产生过分影响的条件下将信息嵌入到数字媒体中，以实现版权保护、隐蔽通信等功能。

信息隐藏分析技术是信息隐藏的对抗技术，用于检测秘密信息的存在或破坏隐秘的通信。本文主要研究针对数字图像的隐写分析技术。

2.1 隐写技术

2.1.1 隐写技术常用术语

在讨论隐写技术之前，先介绍一些信息隐藏技术相关的常用术语：

- 秘密信息M：嵌入到载体信息中需要保密的信息称为秘密信息M；
- 载体信息C：未嵌入秘密信息的原始载体如图像、声音、视频及文本等文件称为载体信息C；
- 掩蔽信息S：载体信息被嵌入秘密信息后称为掩蔽信息S，也称伪装消息。图像隐写及检测过程中载体信息和掩蔽信息也称作载体图像和掩蔽图像；

- 嵌入比例：载密信息的嵌入比例分绝对嵌入比例和相对嵌入比例两种情况，绝对嵌入比例是指单位载体中可嵌入的比特数。相对嵌入比例是指单位载体中嵌入的比特数与最大可嵌入的比特数之比。例如，在24位彩色图像中，单一位平面的最大绝对嵌入比例为3比特/像素。LSB嵌入中24位彩色图像中最大可嵌入比特数为3比特/像素，则某LSB嵌入0.3比特/象素的相对嵌入比例为10%，相对嵌入比例一般简称嵌入比例；
- 隐蔽性：这是信息隐藏的基本要求。要求不影响对载体信息的理解，即人的生理感官和统计检测等都无法发现载秘信息内包含了其他信息，同时不影响载秘信号的感觉效果和使用价值；
- 稳健性：要尽量保证隐藏了信息之后的掩蔽信息在经历可能的处理（如信号处理、有损压缩、滤波、调制等）、主动攻击（如非法攻击、篡改、删除等）或者信道中随机噪声的影响后，还可以提取出原始的秘密信息；
- 安全性：隐藏的具体位置应是安全的，要求至少不会因格式变换而遭到破坏；
- 对称性：通常，秘密信息的隐藏和提取过程具有对称性（包括编码、加密方式等），以减少存取难度；
- 可纠错性：为了保证隐藏信息的完整性，使其在经过各种操作和变换后仍能很好地恢复，通常可采用纠错编码方法；

2.1.2 隐写系统模型

隐写术的经典模型可由Simmons的“囚犯问题”[5]给出。Alice和Bob是关在同一监狱不同囚室的两个囚犯。他们想策划一个越狱方案，但是他们之间所有的通信都要接受看守人Wendy 的检查。Wendy不允许Alice和Bob之间使用加密通信，并且一旦发现任何可疑的通信，将禁止他们之间再交换任何信息。所以，为了不引起怀疑他们必须把秘密信息隐藏在表面看起来无关紧要的消息里来实现隐蔽通信。

如果Wendy仅仅检查Alice和Bob之间的通信，判断其中是否包含隐蔽通信的话，则称Wendy是被动看守人，她所做的攻击为被动攻击。如果Wendy修改Alice传给Bob的消息，甚至可能假装成其中一个囚犯伪造消息并传给另一囚犯，则称Wendy为主动看守人，其攻击为主动攻击。本文主要研究被动看守（或攻击）下的隐写系统，模型如图2.1所示[6,7]，Alice和Bob是隐写系统的使用者，Alice试图将她的秘密信息隐藏于表面上看起来无关紧要的文件中，并以不引起第三方怀疑的方式通过公开信道传给Bob，第三方即敌手Wendy具有该信道的完全可读权限。图中 X 为载体对象(cover object)，表示原始的未嵌消息的数据，一般假设其分布P_X为

Wendy所知：Y 为隐写对象(stego object)，表示嵌入秘密信息后的数据，分布为P_Y；M (embedded message)表示嵌入的秘密信息，是消息空间上的一个随机变量；K (secret key)是Alice和Bob共享的密钥；R 是Alice私有的随机信源。Alice通过R控制向Bob发送载体对象X还是隐写对象Y。

定义 2.1（隐写系统）：给定载体对象 X 的分布 P_X 和消息空间$\mathcal{M}$，定义嵌入算法 $f:\mathcal{X}\times\mathcal{M}\times\mathcal{K}\to\mathcal{Y}$ 和提取算法 $g:\mathcal{Y}\times\mathcal{K}\to\mathcal{M}$，则$(f,g)$被称为一个隐写系统，如果存在上述的 K 和 R，使得对于$\mathcal{M}$上的所有 $H(M)>0$ 的随机变量 M 都满足 $I(\hat{M};M)>0$。

定义中 $I(\hat{M};M)>0$ 的条件暗示了隐写系统必须是一个“有用”的系统，即 Bob 至少可以从 $\hat{M}$ 中获得一些信息。

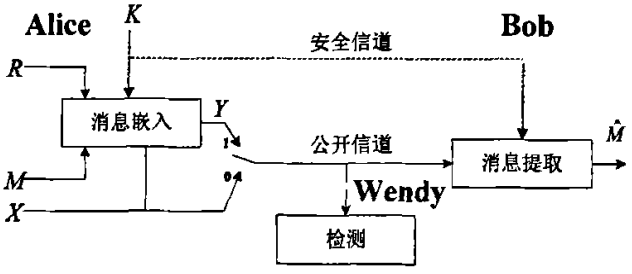


图 2.1 被动攻击下的隐写系统模型

评价隐写系统的主要指标有：

安全性：是隐写系统的首要指标，包括不可检测性和不可感知性两方面的要求。对于一个完善安全的隐写系统，Wendy应该无法检测隐写操作对载体对象分布PX的改变，并且对秘密信息M的存在性一无所知。这一概念平行于Shannon对密码系统的完善保密性概念[8,9]。另一方面，如果因秘密信息嵌入而造成的隐写对象异常能被人类视听觉系统所感知，则隐写系统也是不安全的。当然，不可感知性的要求具有一定的主观性。

隐写容量[10,11]：指隐写系统在一定的安全性要求下所能隐藏的最大信息量，它表征一个隐写系统的隐写能力。隐写过程实际上是建立了一个隐蔽通信信道，而通信信道必须保证通信的有效性。对于隐写系统来说，秘密信息通常仅占载体对象总数据量的很小一部分，有效性比较低，因此隐写容量问题应受到关注。

稳健性[12]：是主动攻击下隐写系统考虑的重点。如果对攻击后的隐写对象中仍能提取或恢复所嵌入的信息，则称这种隐写系统具有稳健性。

上述各种指标通常是互相制约的，例如安全性要求高的隐写系统容量会降低，而增加系统容量又会降低安全性或稳健性，所以应该根据实际应用要求，合理地对这些指标进行折衷选择。

2.2 隐写分析技术

2.2.1 隐写分析攻击类型分类

如同密码和密码分析的关系一样，隐写的目的是隐蔽信息，而隐写分析的目的在于揭示隐蔽信息的存在性，甚至只是指出隐蔽信息的可疑性。和密码分析类似，隐写分析也有着一些相应的攻击类型[13]：

- 唯隐写对象攻击(stego-only attack)：只能获得隐写对象，对可能使用的隐写算法和隐写内容全然不知。
- 已知载体攻击(known cover attack)：可以获得原始的载体和隐写对象。
- 已知消息攻击(known message attack)：在某种意义上，攻击者可以获得隐藏的消息，这可能有助于分析。但即使已知消息，同样是非常困难，甚至可以认为难度等同于唯隐写对象攻击。
- 选择隐写对象攻击(chosen stego attack)：知道隐写工具（算法）和隐写对象。
- 选择消息攻击(chosen message attack)：隐写分析专家用某个隐写工具或算法对一个选择的消息产生隐写对象。这个攻击的目标是确定隐写对象中相应的模式特征，它可以用来指出具体使用的隐写工具或算法。
- 已知载体和隐写对象攻击(known stego attack)：已知隐写工具（算法），并且可获得原始载体和隐写对象。

显然，第一种攻击在技术上最具挑战性，是隐写分析的重要研究内容。不妨说，成功地实现针对任何对象、任何隐写方法的盲检测是分析者要达到的终极目标。然而对隐写算法和隐写内容一无所知的全盲检测往往非常困难，因此，迄今为止人们常针对一些有效的隐写方法和特定的对象研究有针对性的分析技术。

2.2.2 常用的隐写分析方法

现有隐写分析技术可分为两大类：专用隐写分析和通用隐写分析。

1、专用隐写分析

由于盲检测的困难性，所以在隐写分析技术发展初期，人们首先从已知的隐写算法入手，分析在隐写算法已知的情况下隐写检测的可能性，即试图实施选择隐写对象攻击(chosen stego attack)。大量的研究和实践表明，对于现有的多数隐写算法，成功检测不但是可能的，而且可以估计嵌入消息的长度和位置，甚至可以提取消息。

LSB算法是人们最早提出的隐写算法，具有编译码简单，隐藏容量大的特点。

但是该算法一经提出，就遭到很多隐写分析算法的攻击。例如，Westfeld[14]根据在LSB隐藏信息前后载体图像中的值对(pairs of value)的统计特性差异设计了χ^2检验方法；Fridrich[15]利用图像像素值的空间相关性提出了RS(regular singular)统计检测方法；张涛[16]利用自然图像LSB平面和其他比特平面之间存在的弱相关性，设计了基于直方图差异对比的检测方法。这些算法不但可以有效检测嵌入消息的存在性，而且可以估计嵌入消息的长度。

JPEG压缩域上的隐写算法也遭到了各种各样的隐写分析攻击。Westfeld的χ^2检验和Fridrich的RS分析对JSteg同样有效；Fridrich根据F5算法对直方图的改变设计了直方图攻击[17]；而OutGuess和MB(model-based)算法，可以被基于分块特性的隐写分析方法攻破[18]。

对于扩频隐写算法，R. Chandramouli[19]利用信号处理中的盲信号分离技术(blind signal separation)提出了进行主动攻击的数学方法，不但可以检测到隐藏消息的存在，而且对于顺序嵌入的隐写还可以提取出隐藏信息。

然而，成功实现对特定算法的隐写分析并不意味着在现实中就一定能够成功地挫败隐写行为。实际上，这反而为敌手使用非公开的隐写方案以避开检测创造了可能性。所以，在考虑一个隐写分析方案的实用性的时候，必须考虑使用未公开的隐写算法的可能性。此外，当无法得到隐写算法的具体细节（例如只能得到隐写软件的可执行码）时，如果分别对每一个可能的隐写算法进行排除分析，既耗时费力，又不能保证分析的全面性。为此，人们开始从另一个角度寻求更切合实际的目标，即从隐写技术对载体对象造成分布或统计特性的差异这一根本属性上建立一个隐写分析框架，使得对于全部隐写算法或者至少对某一类隐写算法可以实施有效检测，这就是通用隐写分析技术。

2. 通用隐写分析

通用隐写分析技术[20,21]并不关注或者剖析隐写算法的细节，重点研究数字化载体信号中通常出现的模式是否被隐写算法所破坏，例如，以图像为载体的隐写算法通常会改变图像的某些特征（如差分直方图、DCT系数直方图等）所遵循的分布或者统计特性。隐写引起的各种模式的变化往往潜藏在载体数据中，分析者的任务正是要找到这些改变，准确地描述并利用这些改变以达到检测目的，捕捉到的对隐写行为的改变越敏感将越有助于隐写分析。可以说通用隐写分析是一门艺术也是一门科学，其艺术性体现在选择能够暴露隐藏消息存在性的特征或属性上，而科学性则是利用各种数学方法或手段来有效地测试这些选中的特征，从而判断隐藏信息存在与否。

2.3 隐写分析原理和性能评价

2.3.1 隐写分析原理

隐写术通用的隐写过程可表示为：

$$S' = S + f(S, M) \tag{2-1}$$

式(2-1)中，S 和 S' 分别代表载体消息和嵌入秘密信息后的隐藏消息，M 为待嵌入的秘密信息。隐写分析的过程就是从 S' 中检测出 M 以至提取 M。

用阐述隐写术的“囚犯”问题从攻击的角度将隐写分析分为被动攻击和主动攻击两类。进行秘密通信的囚犯的来往信件都要经过看守的检查，被动攻击指看守先判断是否存在秘密信息再做处理；主动攻击指看守不经判断就对消息进行修改的攻击。被动攻击又分被动隐写分析与主动隐写分析两种。被动隐写分析是指检测秘密信息的存在与否，并确定隐藏嵌入算法。主动隐写分析是指估计嵌入的秘密信息的长度，嵌入的位置以及嵌入算法中使用的密钥和某些参数，最终提取秘密信息。主动隐写分析相对于被动隐写分析要困难得多。目前国内外主要研究被动隐写分析。

不可感知和信息隐藏容量大是隐写术的基本要求。由于图像在互联网上应用广泛，且本身冗余度大，可嵌入较大的信息量，因此成为隐写术的主要载体。

2.3.2 隐写分析性能评价

对图象隐写分析方法的评价，一般采用4 个评价指标：准确性、适用性、实用性和复杂度。

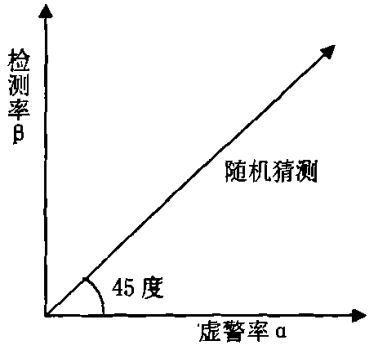


图 2.2 检测器 ROC 平面

准确性指检测的准确程度，是隐写分析最重要的一个评价指标，一般采用虚警率和检测率表示，两者的概率关系可描绘成图2. 2所示的检测器接收操作特性(detector's receiver operating characteristic, ROC)二维平面。虚警率是把非隐藏消息

误判为隐藏消息的概率，表示为$\alpha = P$(隐藏消息|非隐藏消息)；检测率是把隐藏消息正确判为隐藏消息的概率，表示为$\beta = P$(隐藏消息|隐藏消息)。还需要考虑漏报率，即把隐藏信息错误判为非隐藏信息的概率，表示为$\eta = 1-\beta = P$(非隐藏消息|隐藏消息)。隐写分析要求在尽量减少虚警率和漏报率的条件下取得最佳检测率。在虚警率和漏报率无法同时减少的情况下，着重减少漏报率。

全面衡量隐写分析准确性的一个量是全局检测率$P_r = 1-P_e$，其中P_e为平均错误概率：

$$\begin{aligned} P_e &= (1-\beta)P(\text{隐藏消息}) + \alpha P(\text{非隐藏消息}) \\ &= \eta P(\text{隐藏消息}) + \alpha P(\text{非隐藏消息}) \end{aligned} \tag{2-2}$$

当$\alpha = \beta$时，属于随机猜测，此时隐写分析检测器无效。当全局检测率达到85%或以上，可以认为检测器性能良好。

适用性指检测算法对不同嵌入算法的有效性，可由检测算法能够有效检测多少种、多少类隐写术或嵌入算法来衡量。

实用性指检测算法可实际应用的程度，可由现实条件允许与否、检测结果稳定与否、自动化程度和实时性等来衡量。

复杂度针对检测算法本身而言，可由检测算法实现所需要的资源开销、软硬件条件等来衡量。

到目前为止，还没有人给出适用性、实用性和复杂度的定量度量，只能通过比较不同检测算法之间的实现情况和检测效果得出一个相对的结论。

2.4 通用隐写分析技术的系统模型

与特定算法的隐写分析技术不同，通用隐写分析的重点在于区分隐写图像和自然图像，因此通用隐写分析技术一般采用模式识别的思想，即通过对比隐写前后载体模式（如统计特征）的改变或差异，合理选择特征并设计分类器，然后对待测对象进行分类。也有一些隐写分析方法并不直接用这些数据进行分类学习而是将训练数据用于计算一个所选特征集合的回归模型，然后利用该模型预测待测图像是否进行了隐写[20,21]。

通用隐写分析的过程一般分为学习和判决两个阶段[22]，如图2.3所示。学习过程首先对大量的训练图像（包括载体图像和隐写图像两类）进行数据分析，提取和选择有利于分类的特征向量，然后以此构造隐写分析分类器并对分类器进行训练，直到满足一定的精度要求。判决过程是利用学习过程中建立的分类器对被测

图像进行分类。在该模型中，选择分类特征和设计最优分类器是影响检测性能的两个重要因素。

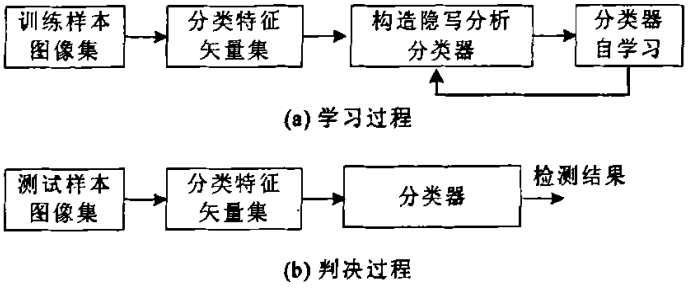


图 2.3 通用隐写分析系统模型

特征是决定分类的关键。确定分类目的之后，如何找到合适的特征就成为识别的核心问题。它严重影响分类器的设计与性能，因为如果对不同类别所选的特征差别很大，则比较容易设计出高性能的分类器，反之，则有可能使得分类器难于构造或分类效率不高。通常，好的特征应该满足准确性、单调性和一致性的要求。准确性可解释为特征以最小的平均误差检测隐藏消息存在的能力；单调性表示特征（理想情况下）必须和嵌入消息大小呈单调关系；一致性是指特征应具有一致性检测的能力，这暗示了特征必须独立于图像类型和种类。

设计最优分类器需要利用统计决策理论在特征空间中把被识别对象归为某一类别，一般要求已知各类别的先验概率和类条件概率密度。前者可以根据先验知识分析得到或者用训练样本中各类的比例进行估计；但后者却需要统计学中的一套复杂方法。事实上，分类器就是一个（或一系列）判别函数（或决策面），如果能够从要解决的问题和训练样本出发直接求出判别函数，就不必进行概率密度估计。在某些情况下，判别函数或决策面可以设为已知的形式，然后通过训练样本确定其参数，就能够更简便地设计出分类器。现有的通用隐写分析算法主要集中于寻找能反映载体图像和隐写图像本质差异的特征，然后利用这些特征训练已知形式的分类器，例如Fisher线性分类器(FLD)、神经网络、支持向量机(SVM)等。

2.5 本章小结

本章主要介绍了隐写分析技术的基础理论知识。首先介绍了信息隐藏技术的基本模型和常用的相关术语，给出了一个隐写系统的基本技术要求，对应于不同的应用，隐写系统需要满足的要求往往也是不同的，而且隐写系统的不可感知性、稳健性和嵌入容量也是相互制约的，因此实际的隐写系统经常需要在多个角度间

进行折衷。

作为隐写技术的对抗技术，隐写分析技术研究有着十分重要的实际意义。本章给出了隐写分析技术的原理、研究方法以及主要评价指标。由于秘密信息的发送者必须修改载体数据才能实现秘密信息的嵌入，因而载体数据的统计特性将不可避免地会发生一些变化，尽管隐写分析者并不知道原始载体数据，但可以利用其统计特性的异常觉察秘密信息的存在。最后，本章还给出通用隐写分析技术的系统模型。

第三章 空域隐写算法的通用隐写分析技术

不同的隐写技术往往采用不同的嵌入机制，包括嵌入域、嵌入位置以及嵌入法则的选择，这些不同必然导致不同类型的隐写技术对载体图像的模式或特征的改变也有差异。因此，对所有隐写技术实现全盲检测往往非常困难或者效率不高。为了提高通用检测的效率，必须区分不同类型的隐写算法，然后针对嵌入机制相同或相似的隐写技术分别设计隐写分析方法。

最常见的分类方法是根据秘密信息的嵌入域分为空域隐写和变换域隐写。嵌入域不同对图像特征的影响也不同，而最好（最敏感）的隐写分析特征应主要从嵌入域直接获取。例如对于 JPEG 图像隐写算法，特征必须在 DCT 系数上构建，而非它们的空域表示。这一原则可以从近期的几种不同的通用隐写分析方法的比较上得到支持[23,24]。接受这个原则意味着根据嵌入域划分通用隐写分析任务的必要性，即空域和 JPEG 压缩域隐写分析。本章我们关注空域隐写算法的通用隐写分析问题。

3.1 空域隐写算法的通用隐写分析技术研究现状

最早的隐写算法是在图像的空域进行的。LSB(Least Significant Bits 最不重要比特)替换算法[25]以其实现简单和隐藏数据量大的特点而广泛应用，但是随着隐写分析技术的发展，人们逐渐认识到，通过有效的统计分析攻击不但可以检测到 LSB 算法的使用，而且有可能提取出嵌入的秘密数据。之后，人们提出了大量的 LSB 改进算法（如自适应 LSB 算法、保持统计特性的 LSB 技术等[26,27]）和 $+-k$ 算法[28]、随机调制(SM)算法[29]等新型的空域隐写算法。

对这些空域隐写算法的隐写分析研究，最初都是停留在针对特定算法设计隐写分析方法上。例如，针对 LSB 算法的隐写分析就有 χ^2 统计分析[14]、RS 攻击[15]、准素集分析、差分直方图分析[16]等；对于 $+-k$ 算法，近年来提出了基于小波域去噪的分析技术[30]、高频分量分布特性分析[31]等；SM 算法（高斯噪声随机调制）也可以从隐写前后图像的差分分布特性差异上被检测到[32]。这些方法虽然能够达到较准确的分析结果，甚至可以提取秘密信息的长度信息，但是它们的适应性和可扩展性都很差。被分析的隐写算法只要稍做修改，就能使原有的分析技术失效。此外，由于隐藏学科的特点，隐写算法可能并不公开，常常无法得知要攻击算法

的具体细节。因此，迫切要求找到一种通用的分析算法，使得不但对于上述的每一种隐写算法，而且对可能出现的空域隐写算法都可以实施有效检测。

通用隐写分析的概念首先出现在 Avcibas 等人的文章中[20,21]。之后，Farid[33,34]提出了基于小波系数及其预测误差的统计矩的通用分析技术。但是由于检测性能不高，这些通用技术在开始时并未受到足够的重视。直到近两年，越来越多的学者开始关注通用隐写分析技术，并发现要提高通用隐写分析的检测性能，必须区分不同嵌入机制的隐写算法。这样，针对不同嵌入域算法的通用隐写分析技术才逐渐发展起来。

目前专门针对空域隐写算法的通用隐写分析技术还比较少。Harmsen[35]把隐写过程建模为加性噪声的叠加过程，并在此框架下对 LSB 方法、扩频方法和 DCT 域隐写方法进行了隐写分析。他将图像的 3 维直方图特性函数(HCF)的质心作为特征，分类器采用多元正态概率模型下的最小错误率贝叶斯分类器。该方法对于低噪声的彩色图像（例如曾经压缩过的图像）具有很好的分析结果，但是对于灰度图像和从数码照相机（或扫描仪）获得的原始的未经压缩的图像，检测性能却显著降低。Ker[36]通过引入校准的概念提高了该方法的性能。可以看出，上述的这些方法都是基于图像的直方图特性提出的，而直方图只反应了图像的一阶分布特性。如果秘密信息在嵌入时保持了直方图，例如在 $+-1$ 算法（$k=1$ 的 $+-k$ 算法）中，可以调制 $+1$ 和 -1 的概率使得直方图被保持，此时，这些方法都无法实施有效检测，所以它们的适用范围有限。

Celik 等人[37]基于隐写算法会引入失真的事实，通过测试隐写前后图像经过有损压缩后的率失真特性(rate-distortion characteristics)改变，对 LSB 和 SM 算法进行了隐写分析。该算法实质上和 Avcibas 的基于图像质量度量(Image Quality Measures: IQM)的分析方法具有相同的思路。

Holotyak 等人[38]采用和 Farid 类似的方法，利用小波高阶统计量对空域加性隐写实施通用隐写分析，但是他们是从小波域的噪声分量计算特征的，而且还提倡使用高阶统计矩，指出考虑高阶矩能够获得很大的性能增益。Goljan[39]考虑了图像的非平稳特性，提出在一级小波分解域上基于 MAP 估计载体图像的局部方差并利用 Wiener 滤波去噪，然后分别计算垂直、水平和对角 3 个子带上残留噪声的 9 阶绝对值中心矩作为特征。该算法获得了比 Holotyak 算法更好的检测性能，特别是显著降低了虚警概率。

本章将遵循 Holotyak 和 Goljan 的观点：特征的选择必须对嵌入改变敏感而对图像内容相当不敏感，提出基于图像差分分布特性选取特征，并采用高阶统计矩来提高隐写分析的性能。

3.2 空域隐写算法建模

数字信号在形成和传输过程中，总是面临着由传感器、采样、量化和信道所引入的噪声干扰。现有的许多图像隐写技术就是利用这些噪声源的存在，试图将秘密信息伪装成自然的噪声，然后将其叠加在载体图像上进行隐蔽通信。然而原始图像的自然统计特性以及采样设备特性，使得数字图像信号上具有大量的相关性。如果在嵌入数据时未考虑这些相关性，则嵌入的数据将成为破坏图像的外部干扰。这一现象使得我们可以对许多依赖于加性噪声的隐写方法建模，并实施隐写分析。

3.2.1 加性噪声隐写系统模型

加性噪声隐写系统模型如图 3.1 所示，隐写噪声是嵌入秘密信息时所引入的等效噪声。设载体图像(i,j)位置上的像素值用随机变量X_{ij}表示，隐写图像(i,j)位置上的像素值用随机变量Y_{ij}表示。(i,j)位置上的隐写噪声用随机变量η_{ij}表示，则隐写过程可表示为：$Y_{ij}=X_{ij}+\eta_{ij}$，$0\le i\le M-1, 0\le j\le N-1$，$M\times N$是图像的大小。

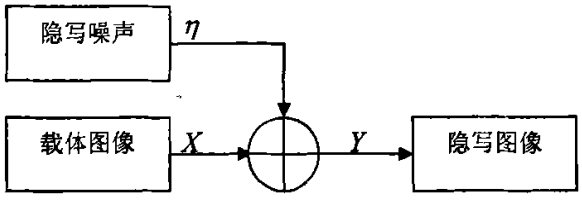


图 3.1 加性噪声隐写模型

在加性噪声模型中，假设隐写噪声在每个像素上是独立的，即η_{ij}，$0\le i\le M-1$，$0\le j\le N-1$是独立同分布(i.i.d.)的随机变量。η_{ij}的分布定义如式(3-1)：

$$\Pr(\eta_{ij}=n_{ij})=p(Y_{ij}-X_{ij}=n_{ij}),\ n_{ij}=0,\pm1,\pm2,... \tag{3-1}$$

LSB、$+-k$和 SM 算法（高斯噪声随机调制）都可以看作是基于加性噪声的隐写方法，文献[27]还提出了根据任意给定的噪声分布嵌入数据的广义加性噪声隐写方案。

3.2.2 LSB 算法

LSB 算法是用秘密信息比特直接替换载体图像像素值的 LSB，即如果秘密信息比特和载体图像像素值的 LSB 相同，则对该像素值不做任何改动；反之，将该像素值的 LSB 翻转。可用式(3-2)来描述：

$$y_{ij} = LSB\_Stego(x_{ij}) = \begin{cases} x_{ij}, & if\ m_k = LSB(x_{ij}) \\ x_{ij} + 1 - 2 \times LSB(x_{ij}), & if\ m_k \neq LSB(x_{ij}) \end{cases} \tag{3-2}$$

其中，x_{ij}、y_{ij}和m_k分别是随机变量X_{ij}、Y_{ij}和M的具体分布。设秘密信息的嵌入比例为p。由于嵌入的消息往往经过加密，所以可看作是0、1随机分布的比特流，而载体图像的LSB也可近似为独立同分布(i.i.d.)的随机变量且$\Pr(X_{i,j}^{LSB}=0)=\Pr(X_{i,j}^{LSB}=1)=1/2$，则

$$\Pr(\eta_{ij}=-1)=p(Y_{ij}-X_{ij}=-1)=\Pr(X_{i,j}^{LSB}=1)\Pr(M=0)=p/4 \tag{3-3}$$

$$\Pr(\eta_{ij}=0)=p(Y_{ij}-X_{ij}=0)=\Pr(X_{i,j}^{LSB}=0)\Pr(M=0)+\Pr(X_{i,j}^{LSB}=1)\Pr(M=1)=1-p/2 \tag{3-4}$$

$$\Pr(\eta_{ij}=1)=p(Y_{ij}-X_{ij}=1)=\Pr(X_{i,j}^{LSB}=0)\Pr(M=1)=p/4 \tag{3-5}$$

η_{ij}的数学期望和方差分别为：

$$E(\eta_{ij})=\sum_{n_{ij}} n_{ij}\Pr(\eta_{ij}=n_{ij})=0 \tag{3-6}$$

$$\sigma_{n_{ij}}^2=\sum_{n_{ij}}(n_{ij}-E(\eta_{ij}))^2\Pr(\eta_{ij}=n_{ij})=p/2 \tag{3-7}$$

3.2.3 随机$+-k$算法

随机$+-k$算法是通过对载体图像像素值增减k嵌入秘密信息，具体地可用式(3-8)来描述：

$$y_{ij} = PMk\_Stego(x_{ij}) = \begin{cases} x_{ij}-k, if(r_{ij}>0\ or\ x_{ij}=255)\ and\ m_k \neq LSB(x_{ij}) \\ x_{ij}, \quad if\ m_k = LSB(x_{ij}) \\ x_{ij}+k, \ if(r_{ij}<0\ or\ x_{ij}=0)\ and\ m_k \neq LSB(x_{ij}) \end{cases} \tag{3-8}$$

其中，r_{ij}服从$\{-1,+1\}$上的均匀分布。随机$+-k$算法的“随机”强调了嵌入数据时，像素值的增加或减少是由共享密钥生成的伪随机数r_{ij}来控制的。该算法所引入的等效的隐写噪声η_{ij}的分布为：

$$\Pr(\eta_{ij}=0)=1-p/2 \tag{3-9}$$

$$\Pr(\eta_{ij}=k)=\Pr(\eta_{ij}=-k)=p/4 \tag{3-10}$$

η_{ij}的数学期望和方差分别为：

$$E(\eta_{ij})=\sum_{n_{ij}} n_{ij}\Pr(\eta_{ij}=n_{ij})=0 \tag{3-11}$$

$$\sigma_{\eta_{ij}}^2=\sum_{n_{ij}}(n_{ij}-E(\eta_{ij}))^2\Pr(\eta_{ij}=n_{ij})=k^2p/2 \tag{3-12}$$

当 $k=1$ 时，对于给定的嵌入比例 p，$+-k$ 算法所引入的隐写噪声能量和 LSB 算法相同。但是，由于它是在两个方向上（加 1 或减 1）对像素值做改变，不只改动 LSB 平面，而且还向其它平面扩散，因此对它的隐写分析要比 LSB 算法难得多。

3.2.4 SM 算法

SM（高斯噪声随机调制）算法就是通过在载体图像中加入服从某种概率分布的隐写噪声隐藏秘密信息。为了实现零误码地提取消息比特，算法巧妙设计了一个具有反对称性的参数奇偶函数，使得将该函数应用于隐写图像像素值就可以恢复消息比特。

具体地，对于 x_{ij} 的所有可能的值，定义参数奇偶函数 $P(x_{ij},r_{ij})\in\{-1,1\}$ 满足：$P(x_{ij}+r_{ij},r_{ij})=-P(x_{ij}-r_{ij},r_{ij})$（$r_{ij}>0$ 是一个整数参数），当 $r_{ij}=0$ 时，$P(x_{ij},r_{ij})=0$。出于对嵌入溢出处理的考虑，算法所采用的奇偶函数形式为：

$$\text{当 } x_{ij}\in[1,2r_{ij}]\text{ 时，}P(x_{ij},r_{ij})=\begin{cases}(-1)^{x_{ij}+r_{ij}}, & r_{ij}>0\\ 0, & r_{ij}=0\end{cases} \tag{3-13}$$

当 $x_{ij}\notin[1,2r_{ij}]$ 时，利用反对称性和式(3-13)计算 $P(x_{ij},r_{ij})$。

隐藏消息时，对嵌入行程上的每个像素灰度值 x_{ij} 生成服从某种分布的噪声，取整后得到 r_{ij}，如果 $r_{ij}=0$，则不改动 x_{ij}，并移向下一像素；如果 $r_{ij}\neq 0$，则判断是否有 $P(x_{ij}+r_{ij},r_{ij})=m_k$（该算法假设 m_k 取自+1、−1 的伪随机序列）。如果是，则将 x_{ij} 改为 $x_{ij}+r_{ij}$ 并移向下一像素，反之，将 x_{ij} 改为 $x_{ij}-r_{ij}$ 并移向下一像素。该嵌入规则可表示为式(3-14)：

$$y_{ij}=SM\_Stego(x_{ij})=x_{ij}+m_kP(x_{ij}+r_{ij},r_{ij})r_{ij} \tag{3-14}$$

如果像素灰度值产生溢出，则需要做截断处理。整个隐写过程可看作是用噪声对秘密信息进行随机调制，然后叠加在载体图像上。使用服从 $N(0,\sigma^2)$ 的高斯噪声进行随机调制隐写时，量化后参数的分布为：

$$\Pr(R=r_{ij})=\int_{r_{ij}-1/2}^{r_{ij}+1/2}\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{t^2}{2\sigma^2})dt \tag{3-15}$$

嵌入比例 p 和高斯噪声方差 σ^2 之间存在如下关系：

$$p=1-\Pr(R=0)=1-\int_{-1/2}^{1/2}\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{t^2}{2\sigma^2})dt=1-erf(\frac{1}{2\sqrt{2\sigma^2}}) \tag{3-16}$$

其中 $erf(\cdot)$ 是误差函数：$erf(x)=\frac{2}{\sqrt{\pi}}\int_0^x\exp(-t^2)dt$。因此有：

$$\sigma=\frac{\sqrt{2}}{4erf^{-1}(1-p)} \tag{3-17}$$

令 $n_{ij}=m_kP(c_{ij}+r_{ij},r_{ij})r_{ij}$，由于 m_k 取自+1、−1 的伪随机序列，且 x_{ij} 和 r_{ij} 独立于 m_k，所以 $m_kP(x_{ij}+r_{ij},r_{ij})$ 也是+1、−1 的伪随机序列，而 r_{ij} 关于 0 对称，则 n_{ij} 和 r_{ij} 具有相同的分布特性：

$$\begin{aligned}\Pr(\eta_{ij}=n_{ij})&=\int_{n_{ij}-1/2}^{n_{ij}+1/2}\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{t^2}{2\sigma^2})dt\\&=\frac{1}{2}erf(\frac{n_{ij}+1/2}{\sqrt{2}\sigma})-\frac{1}{2}erf(\frac{n_{ij}-1/2}{\sqrt{2}\sigma})\end{aligned} \tag{3-18}$$

η_{ij} 的数学期望和方差分别为：

$$E(\eta_{ij})=\sum_{n_{ij}}n_{ij}\Pr(\eta_{ij}=n_{ij})=0 \tag{3-19}$$

$$\begin{aligned}\sigma_{\eta_{ij}}^2&=\sum_{n_{ij}}(n_{ij}-E(\eta_{ij}))^2\Pr(\eta_{ij}=n_{ij})=\sum_{n_{ij}}n_{ij}^2\Pr(\eta_{ij}=n_{ij})\\&=\frac{1}{2}\sum_{n_{ij}}n_{ij}^2[erf(\frac{n_{ij}+1/2}{\sqrt{2}\sigma})-erf(\frac{n_{ij}-1/2}{\sqrt{2}\sigma})]\end{aligned} \tag{3-20}$$

理论上，求和应取遍所有整数，但是给定 σ，随着 $|n_{ij}|$ 的增大，$\Pr(\eta_{ij}=n_{ij})$ 迅速降为 0。

综上所述，LSB、$+-k$ 和 SM 算法（高斯噪声随机调制）所引入的隐写噪声均值为零，并且方差和嵌入比例 p 之间具有某种确定性关系。基于这一事实，我们可以设计有效的通用隐写分析方案，从而对这类基于加性噪声模型的隐写算法实施有效检测。

3.3 基于图像差分分布特性的通用隐写分析技术

从3.2节的分析中，可以看到基于加性噪声的隐写系统在图像中加入了与图像独立的隐写噪声信号，这势必会破坏图像相邻像素值的相关性或连续性，从而导致差分分布改变。本节将利用这一规律实现一种通用隐写分析算法。

一般假设图像的差分分布服从广义高斯分布GGD[40]，可以采用GGD参数拟合的方法估计差分分布；但是更为简便地，我们直接利用差分直方图来估计。首先分析LSB、$+-k$和SM算法所引入的隐写噪声对图像差分直方图产生的影响，然后给出特征提取算法和分类器设计方案。

3.3.1 隐写噪声对图像差分直方图的改变

为了分析隐写噪声对图像差分直方图的改变规律，首先分别用LSB、$+-k$ ($k=1,3,5$)和SM算法在图像中嵌入信息。图3.2给出了标准图像Lena隐写前后的差分直方图对比，可以看出隐写噪声对差分分布的影响呈现一定的规律性：各隐写图像差分值为0的出现频次都比载体图像少，但是随着差分绝对值的增大，出现频次逐渐和载体图像的接近，到达某点后，它们和载体图像相应差分值的频次之差又呈现增大趋势。例如，LSB和$+-1$算法使得隐写差分直方图曲线在差分值为2和-2的点上和载体差分直方图曲线相交；$+-3$和SM算法的隐写差分直方图曲线在差分值为3和-3的点上和载体差分直方图曲线相交；$+-5$算法则在差分值为4和-4的点上和载体差分直方图近似相交（因为差分值是离散的，实际的交点并不在差分值上）。此外，隐写操作还增大了图像差分分布的方差。这一点可用理论证明如下：

采用水平差分进行分析，设$dx_{ij}=x_{ij}-x_{i,j+1}$，$dy_{ij}=y_{ij}-y_{i,j+1}$，$dn_{ij}=n_{ij}-n_{i,j+1}$，$0\le i\le M-1, 0\le j\le N-2$，则在加性噪声模型下有：

$$dy_{ij}=y_{ij}-y_{i,j+1}=(x_{ij}+n_{ij})-(x_{i,j+1}+n_{i,j+1})=dx_{ij}+dn_{ij} \tag{3-21}$$

用随机变量表示为：$dY_{ij}=dX_{ij}+d\eta_{ij}$，$0\le i\le M-1$，$0\le j\le N-2$。由于隐写噪声和载体图像独立，则它们的差分之间也相互独立，dY_{ij}、dX_{ij}和$d\eta_{ij}$的方差之间具有如下关系：

$$Var(dY_{ij})=Var(dX_{ij})+Var(d\eta_{ij}) \tag{3-22}$$

隐写噪声$\eta_{ij}\ (0\le i\le M-1, 0\le j\le N-1)$，是独立同分布(i.i.d.)的随机变量，因此

$$Var(dY_{ij})-Var(dX_{ij})=Var(d\eta_{ij})=Var(\eta_{ij}-\eta_{i,j+1})=2\sigma_{\eta_{ij}}^2 \tag{3-23}$$

根据 3.2 节导出的隐写噪声方差和嵌入比例 p 之间的确定性关系，可以得到隐写图像差分方差和载体图像差分方差之间也具有确定性关系。

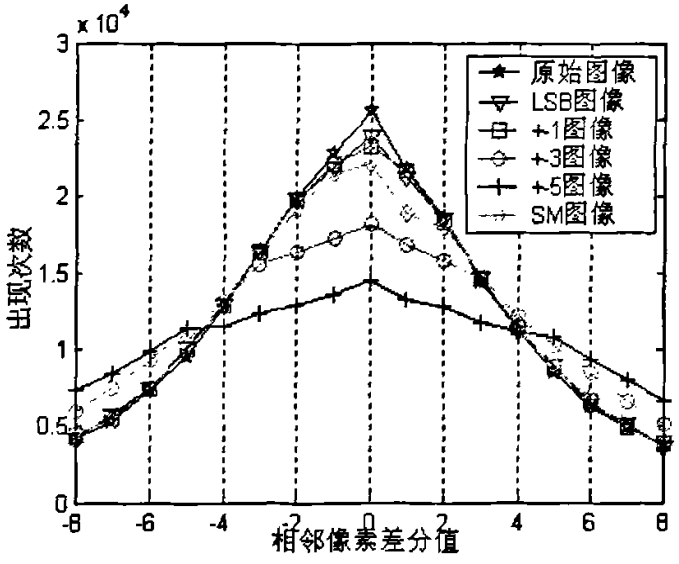


图 3.2 隐写图像和载体图像差分直方图对比

从图 3.2 中，还可以看出 LSB 和＋－1 算法对差分直方图的改变最小，这是因为它们所引入的隐写噪声方差最小。随着嵌入比例 p 的减小，它们对差分直方图的改变也会变小而不易于分析。此时，将载体图像和隐写图像的 LSB 平面翻转，再次计算差分直方图，如图 3.3 所示。可以发现，LSB 算法对 LSB 翻转后的差分直方图改变很大，这是因为 LSB 嵌入操作破坏了图像 LSB 平面和其它平面之间的相关性。

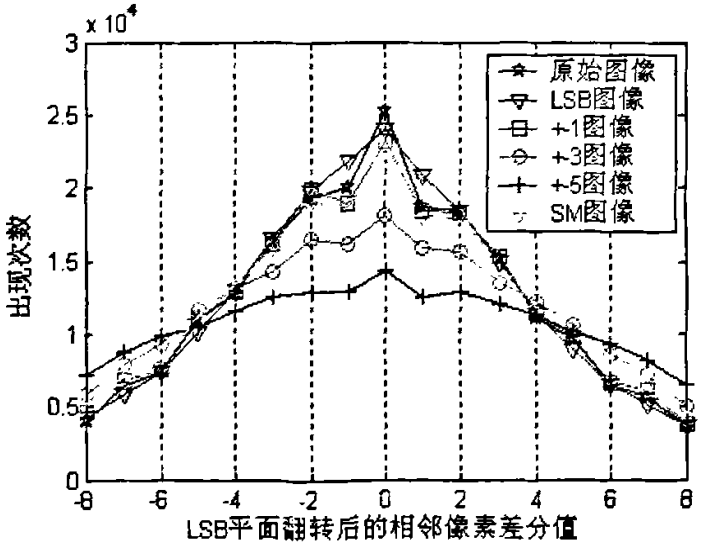


图 3.3 隐写图像和载体图像 LSB 平面翻转后差分直方图对比

但是，＋－1 算法对 LSB 平面翻转后的差分直方图的改变仍然较小，因此，＋－1 算法将是我们分析中的难点。需要找到能够量化差分直方图中心区域（两条差分直方图曲线交点以内区域）和尾部分布（交点以外）差异的量。对于中心区域，以差分值－1 到 1 之间的曲线下面积占该区域面积的比重来量化；对于尾部区域，借鉴 Holotyak 等人[38]的做法，使用高阶统计矩来捕捉分布差异。

3.3.2 特征选择

选择合适的特征空间对构建隐写分析分类器起着至关重要的作用。Holotyak 和 Goljan 提出：特征必须对嵌入改变敏感而对图像内容相当不敏感。也就是说，必须消除图像数据对于隐写分析的影响，他们采用了在小波域估计隐写噪声的方法。本章提取图像的差分分量相当于剔除了原始图像的大部分信息，使得隐秘信号（或被隐写噪声调制）和干扰信息（原始图像）的比率(SNR)增大，有助于提高隐写分类性能。

第一组统计特征是利用差分直方图和 LSB 翻转后的差分直方图分别估计图像差分分布和 LSB 翻转后的差分分布，然后计算这两个分布的高阶中心矩。由于载体图像和隐写图像的差分都呈近似对称分布，奇数阶中心矩近似为 0，所以使用偶数阶中心矩 m_2、m_4、m_6、m_8、m_{10}、m_{12}，其中大于 2 阶的中心矩用标准差 $\sqrt{m_2}$ 进行标准化，即 $m_i^{'} = m_i / \sqrt{m_2}$，$i = 4,6,8,10,12$。

另一组统计特征主要反应差分直方图和 LSB 翻转后的差分直方图的中心区域在隐写信息前后的变化。从图 3.2 和图 3.3 可以看出，几种算法的中心区域面积大体一致，因此，本文选取差分值－4 到 4 之间的曲线下面积作为中心区域面积。计算差分值－1 到 1 之间的曲线下面积和中心区域面积的比重 S 作为分析的特征。

这样，数字图像差分直方图分布和 LSB 翻转后的差分直方图分布的中心矩 m_2、$m_4^{'}$、$m_6^{'}$、$m_8^{'}$、$m_{10}^{'}$、$m_{12}^{'}$（12 个特征向量），以及差分直方图和 LSB 翻转后差分直方图的上述面积比重 S（2 个特征向量）两组统计量共 14 个构成最终的隐写分析特征向量。

3.3.3 分类器设计准则

模式识别分类问题是指根据待识别对象所呈现的观察值，将其分到某个类别中去。具体步骤如下[50]：

(1) 建立特征空间中的训练集，已知训练集里每个点所属类别。

(2) 从这些条件出发，寻求某种判别函数或判别准则，设计判决函数模型。

(3) 根据训练集中的样品确定模型中的参数。

(4) 将这一模型用于判决，利用判决函数或判别准则去判别每个未知类别的

点应该属于哪一类。

模式识别的基本框架——制定准则函数，实现准则函数极值化。常用的准则有以下几种。

1、最小错分率准则

完全以减少分类错误为原则，这是一个通用原则，它使错分类的样品数量为最小。

2、最小风险准则

当接触到实际问题时，可以发现使错误率最小不一定是一个普遍适用的最佳选择。有的分类系统对错误率大小并不看重，而是要考虑错分类的不同后果，为使总的损失最小，有时宁肯将错分率加大。因此引入风险、损失这些概念，以便在决策时兼顾不同后果的影响。在实际中计算损失与风险是复杂的，在使用数学公式计算时，往往用赋予不同权值来表示。在做出决策时，要考虑所承担的风险。基于最小风险的贝叶斯决策规则是为了体现这一点而产生的。

3、近邻准则

近邻准则是分段线形判别函数的一种典型方法。这种方法主要依据同类物体在特征空间具有聚类特性的原理。同类物体由于其性质相近，它们在特征空间中应具有聚类的现象，因此可以利用这种性质产生分类决策的规则。例如有两类样品，可以求出某一类的平均值，对于任何一个未知样品，先求出它到各个类的平均值距离，判断距离哪个类近就属于哪个类。

4、Fisher 准则

根据两类样品一般类内密集，类间分离的特点，寻找线性分类器最佳的法线向量，使两类样品在该方向上的投影满足类内尽可能密集，类间尽可能分开。相反如果把它们投影到任意一根直线上，有可能不同类别的样品就混在一起了，无法区分。如果把投影直线旋转一定的角度，就有可能找到一个方向，样品投影到这个方向直线上，各类样品就能很好的分开。如何找到一个最好方向及如何实现向最好方向投影的变换，这正是 Fisher 算法要解决的基本问题。

5、感知准则

感知准则函数以使错分类样品到分界面距离之和最小为原则。采用错误提供信息实现迭代修正的学习原理。用错分类提供的信息修正错误，这种思想对机器学习的发展以及人工神经元网络的发生发展产生深远影响。其优点是通过错分类样品提供的信息对分类器函数进行修正，这种准则是人工神经元网络多层感知器的基础。

3.3.4 Fisher 线性判别分类器

Fisher线性判别(FLD)是一种经典的两类判别技术[33]，它的目的是为两类输入数据找到一个最优的投影方向，使得数据的类内距离最小且类间距离最大，从而能正确的分类。

Fisher 线性判别分类器的基本思想是将 d 维特征空间的样本投影到一条直线上，形成一维空间，一般情况下，如果样本是线性可分的，则总能找到某个方向，在这个方向上样本的投影能最好的分开。

设列向量 $\boldsymbol{x}_i, i=1,...,N_x$ 和 $\boldsymbol{y}_j, j=1,...,N_y$ 是从载体图像和隐写图像两类训练集合中分别提取出的特征向量，分类器的具体实现如下。

1、分别计算两类训练集合的类内均值，如式(3-24)：

$$\boldsymbol{\mu}_x = \frac{1}{N_x}\sum_{i=1}^{N_x}\boldsymbol{x}_i, \quad \boldsymbol{\mu}_y = \frac{1}{N_y}\sum_{j=1}^{N_y}\boldsymbol{y}_j \tag{3-24}$$

2、计算两类训练集合的总类间均值，如式(3-25)：

$$\boldsymbol{\mu} = \frac{1}{N_x + N_y}\left(\sum_{i=1}^{N_x}\boldsymbol{x}_i + \sum_{j=1}^{N_y}\boldsymbol{y}_j\right) \tag{3-25}$$

3、计算训练集合的总类内离散度矩阵，如式(3-26)：

$$S_\omega = M_x M_x^T + M_y M_y^T \tag{3-26}$$

其中，M_x 的第 i 列包含第 i 个零均值样本即 $\boldsymbol{x}_i - \boldsymbol{\mu}_x$。同理，$M_y$ 的第 j 列包含第 j 个零均值样本即 $\boldsymbol{y}_j - \boldsymbol{\mu}_y$。

4、计算训练集合的总类间离散度矩阵，如式(3-27)：

$$S_b = N_x(\boldsymbol{\mu}_x - \boldsymbol{\mu})(\boldsymbol{\mu}_x - \boldsymbol{\mu})^T + N_y(\boldsymbol{\mu}_y - \boldsymbol{\mu})(\boldsymbol{\mu}_y - \boldsymbol{\mu})^T \tag{3-27}$$

5、求向量 w^*：

定义 Fisher 准则函数如式(3-28)：

$$J_F(\boldsymbol{w}) = \frac{\boldsymbol{w}^T S_b \boldsymbol{w}}{\boldsymbol{w}^T S_\omega \boldsymbol{w}} \tag{3-28}$$

其中，$\boldsymbol{w}$ 为特征向量。令 $\boldsymbol{w}^*$ 为使 $J_F(\boldsymbol{w})$ 最大的特征向量，即 $S_b\boldsymbol{w}^* = \lambda S_\omega \boldsymbol{w}^*$，则当训练样本 $\boldsymbol{x}_i$ 和 $\boldsymbol{y}_j$ 投影到由 $\boldsymbol{w}^*$ 定义的一维线性子空间（即 $\boldsymbol{x}_i^T\boldsymbol{w}^*$ 和 $\boldsymbol{y}_j^T\boldsymbol{w}^*$）时，类内离散度最小而类间离散度矩阵最大。因此，该投影不但降低了数据维数，而

且保持了可分性。

6、确定在投影空间上的分割阈值T。

7、对于测试集合中的某个待测样本z，计算出它在该子空间中的投影$z^T w^*$。

根据决策规则式(3-29)就可以分类：

$$\begin{cases} z^T w^* > T \Rightarrow z \in \omega_1 \\ z^T w^* < T \Rightarrow z \in \omega_2 \end{cases} \tag{3-29}$$

在两类 FLD 的情况下，必须确保特征向量能够投影到一维子空间（即至多存在一个非零特征值）。

3.3.5 实验与结果分析

对 LSB、$+-k(k=1,3,5)$和 SM 算法分别用本章的通用分析算法进行实验，实验用的图像为 BMP 灰度图像。实验时首先取 100 幅图像作为未嵌入秘密信息的载体图像，分别统计差分直方图特性并计算特征向量；然后用 LSB、$+-k(k=1,3,5)$和 SM 算法在每幅图像中嵌入秘密信息形成隐写图像，并计算每幅隐写图像的特征向量。这样，200 幅图像的特征向量构成了训练样本集。类似地，另取 200 幅图像构成测试样本集。

分别测试当嵌入比例p=1、0.5、0.25 bpp（比特/像素）时算法的检测性能，。表 3.1 给出了隐写分析测试结果。

表 3.1 通用隐写分析测试结果

隐写算法	嵌入比例 (bpp)	检测率(%)	虚警率(%)	漏检率(%)
LSB	1	98.0	3.0	1.0
	0.5	96.0	5.0	3.0
	0.25	83.5	23.0	10.0
+−1	1	96.0	4.0	4.0
	0.5	76.0	30.0	18.0
	0.25	57.0	50.0	36.0
+−3	1	99.5	1.0	0
	0.5	96.0	6.0	2.0
	0.25	90.0	12.0	8.0
+−5	1	100	0	0
	0.5	98.0	2.0	2.0
	0.25	96.0	5.0	3.0
SM	1	94.0	7.0	5.0
	0.5	81.5	26.0	11.0
	0.25	67.0	40.0	26.0

从实验结果表 3.1 可以看到，在秘密信息嵌入比例为 1bpp 的情况下，本章算法对 LSB、$+-k$ 和 SM 隐写算法的检测性能都在 90%以上，均高于第二章隐写分析性能评价指出的 85%，因此，此算法的检测性能良好。

在秘密信息嵌入比例降低为 0.5bpp 的情况下，本章算法除了对 $+-1$ 和 SM 算法的检测性能有所下降外，仍然具有良好的检测性能。随着秘密信息嵌入比例的降低，隐写图像对所选取特征的改变也越来越小。

3.4 本章小结

本章针对一类基于加性噪声的空域隐写技术设计了通用隐写分析算法。算法通过分析隐写操作所引起的图像差分直方图的各种变化，分别利用由直方图估计的差分分布的高阶统计矩和差分直方图曲线下特殊区域之间面积的比重来量化隐写信息前后的差分分布差异，从而给出了设计通用隐写分析方案的一般方法，为解决无法得知算法细节的新型隐写技术的隐写分析问题提供了途径。

通过本章的隐写分析实例，可以得到提高空域隐写技术安全性的启示：就是必须打破隐写噪声和载体图像的独立性假设，设计依赖于图像内容的自适应隐写技术。例如，Fridrich 在 SM 算法（高斯噪声随机调制）之后又提出叠加基于内容的隐写噪声的随机调制算法，这一改进使得秘密信息能够更好的藏匿于载体图像的空间分布中。

第四章 JPEG 压缩域算法的通用隐写分析技术

本章主要研究 JPEG 压缩域隐写算法的通用隐写分析技术。上一章已经指出，最敏感的隐写分析特征主要应该从嵌入域直接获取。因此，对于 JPEG 压缩域隐写算法，特征应该在 DCT 系数上构建。

4.1 JPEG 压缩域隐写算法的通用隐写分析技术研究现状

JPEG 是一种使用非常广泛的图像格式，由于获取便利和压缩性能优良等特点，使得以 JPEG 图像为载体的隐写技术越来越受到研究人员的青睐。目前已有多种 JPEG 图像隐写算法，如 JSteg[41]，F5 [42]和 OutGuess [43]等，它们都是在 JPEG 压缩域上实现隐写的。

和空域隐写分析技术类似，对 JPEG 压缩域隐写算法的隐写分析最初也是针对特定算法设计的。例如，对于 JSteg 隐写算法，χ^2 统计分析、RS 分析稍加修改就可以对它实施有效检测；对于 F5 算法，Fridrich 等人[17]提出直方图攻击；OutGuess 算法虽然保持了直方图特性，但是却不能抵抗 Fridrich[18]提出的基于分块特性的隐写分析。同样，这些专用的隐写分析方法虽然能够达到较准确的分析结果，但是它们的适应性和可扩展性较差。为此仍然寻求通用性更强的分析算法，使得不但对于上述的每一种隐写算法，而且对可能出现的 JPEG 压缩域隐写算法都可以实施有效检测。

除 Avcibas 和 Farid 的通用隐写分析技术之外，专门针对 JPEG 压缩域隐写的通用隐写分析技术还不是很多。经典算法是 Fridrich 提出的基于一阶和二阶分布特性的 JPEG 图像隐写分析技术[44]。她将“校准”图像(calibrated images)的概念引入基于特征分类的通用隐写分析技术中，其所有的特征都以如下方式构造，如图 4.1 所示：首先对 JPEG 图像 $\boldsymbol{J}_1$ 应用一个矢量函数 $\boldsymbol{F}$，该函数可以是 DCT 系数直方图、空间分块特性、共生矩阵(co-occurrence matrix)等特征提取算法。然后图像 $\boldsymbol{J}_1$ 被解压到空域，在每个方向上分别裁去 4 个像素，并以同样的量化表将 $\boldsymbol{J}_1$ 再次压缩为 $\boldsymbol{J}_2$。对 $\boldsymbol{J}_2$ 再次应用同样的矢量函数 $\boldsymbol{F}$。最后的特征 f 是二者之差的 L_1 范数。L_1 范数定义为所有矢量元素的绝对值之和。

这样选取特征的原因为：裁减并再次压缩所产生的“校准”图像所具有的大多数宏观特征都和原始图像近似。特别对于隐写图像，裁减后的隐写图像和载体

图像的感知特性相近，其 DCT 系数应该和载体图像具有几乎相同的统计特性。这里裁去 4 个像素非常重要，因为再次压缩（8×8 分块）后的 DCT 系数不会被以前的 DCT 域上的量化（和嵌入信息）所干扰。可以将裁减并再次压缩的图像作为载体图像的一个近似或者边信息(side-information)。将校准图像作为边信息已在 Fridrich 设计准确估计特定隐写算法嵌入消息长度中被证实非常有效[17,18]。

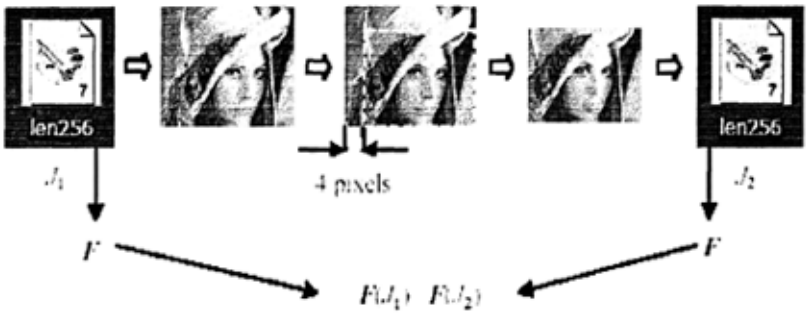


图 4.1 Jridrich 的基于“校准”的特征选取算法

Fridrich 分别选取了表征图像一阶分布的 DCT 总体系数直方图、低频系数直方图、二重直方图以及表征二阶分布的分块特性、共生矩阵等作为隐写分析的特征，其中多数特征直接在 DCT 域上提取，部分特征取自空域，如分块特性。文献[24]证明了该方法的有效性。

不同于 Fridrich 利用校准图像计算特征的方法，我们在分析现有 JPEG 压缩域隐写算法对图像 DCT 系数直方图产生改变的基础上，提出直接从图像的 DCT 系数直方图分布特性出发，以 DCT 系数直方图作为 DCT 系数的一阶分布估计，然后选取一阶分布的数字特征和特殊系数值分布概率作为隐写分析特征。

4.2 JPEG 压缩域隐写算法对直方图的改变

4.2.1 JPEG 压缩编码

JPEG(Joint Photographic Experts Group)压缩标准是 ISO/IEC 联合图像专家组制定的适用于连续色调（包括灰度和彩色）静止图像的压缩标准。图 4.2 给出了图像 JPEG 压缩编码框图。

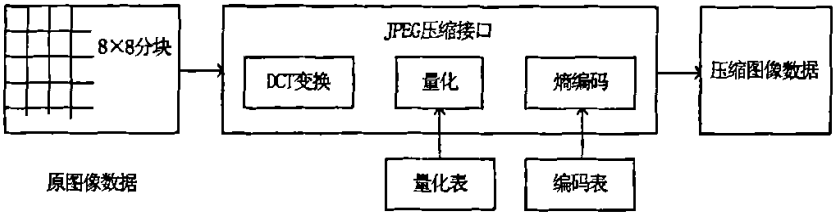


图 4.2 JPEG 图像压缩编码框图

1、离散余弦变换 DCT

JPEG 编码是对每个单独的颜色图像分量分别进行的。在进行正向离散余弦变换 FDCT(Forward Discrete Cosine Transform)之前，需要先将整个分量图像分成 8×8 像素的图像块（不足部分可以通过重复图像的最后一行/列来填充），这些图像块被作为二维正向离散余弦变换(FDCT)的输入。参见图 4.3。

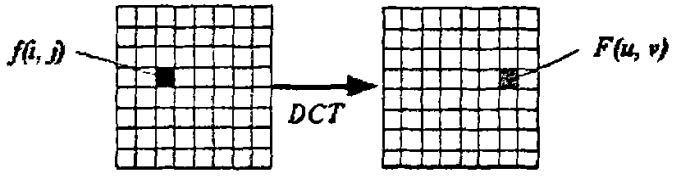


图 4.3 离散余弦变换

JPEG 编码的 DCT 变换使用式(4-1)计算：

$$F(u,v)=\frac{1}{4}C(u)C(v)\left[\sum_{i=0}^{7}\sum_{j=0}^{7}f(i,j)\cos\frac{(2i+1)u\pi}{16}\cos\frac{(2j+1)v\pi}{16}\right] \quad (4\text{-}1)$$

其中，$C(w)=\begin{cases}\frac{1}{\sqrt{2}}, w=0\\ 1, \quad w>0\end{cases}$

并称 $F(0,0)$ 为直流系数，称其他 $F(u,v)$ 为交流系数。

2、量化

量化(quantization)是对经过 FDCT 变换后的频率系数进行量化。量化的目的是减小非“0”系数的幅度以及增加“0”值系数的数目。量化是使图像质量下降的最主要原因。对于 JPEG 的有损压缩算法，使用的是如图 4.4 所示的均匀量化器进行量化，量化步距是按照系数所在的位置和每种颜色分量的色调值来确定的。

由于人眼对亮度信号比对色差信号更敏感，因此 JPEG 编码中使用了两种标准的量化表：亮度量化表和色差量化表。量化的具体计算为式(4-2)：

$$Sq(u,v)=round\left(\frac{F(u,v)}{Q(u,v)}\right) \tag{4-2}$$

其中，$Sq(u,v)$为量化后的结果，$F(u,v)$为 DCT 系数，$Q(u,v)$为量化表中的数值、$round(\ )$为舍入取整函数。

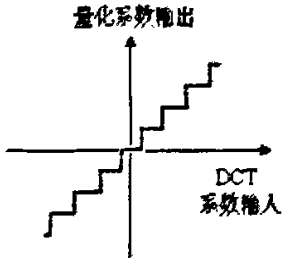


图 4.4 均匀量化器

量化后的二维系数要重新编排，并转换为一维系数，为了增加连续的“0”系数的个数，就是“0”的游程长度，JPEG 编码中采用的 Z 字形编排方法，如图 4.5 所示。

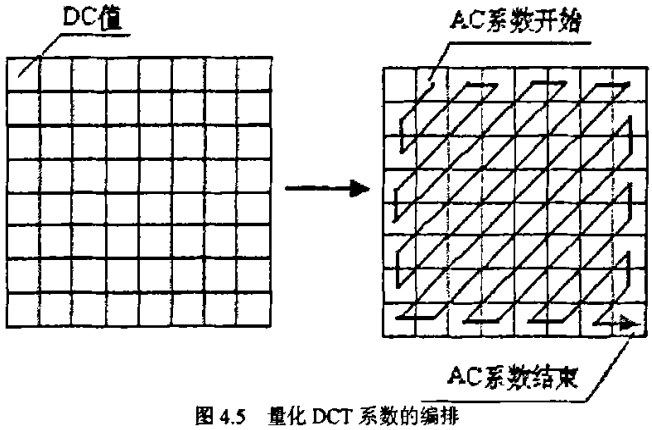


图 4.5 量化 DCT 系数的编排

这样就把一个 8×8 的矩阵变成一个 1×64 的矢量，频率较低的系数放在矢量的头部。

3、直流系数的编码

8×8 图像块经过 DCT 变换之后得到的 DC 直流系数有两个特点，一是系数的数值比较大，二是相邻 8×8 图像块的 DC 系数值变化不大。根据这些特点，JPEG 算法使用了差分脉冲编码调制(DPCM)技术，对相邻图像块之间的 DC 系数的差值 $\Delta=DC(0,0)_k-DC(0,0)_{k-1}$ 进行编码。

4、交流系数的编码

量化 AC 系数的特点是 1×63 矢量中包含有许多“0”系数，并且许多“0”是连续的，因此使用非常简单和直观的游程长度编码(RLE)对它们进行编码。

JPEG 使用 1 个字节的高 4 位来表示连续“0”的个数，而使用它的低 4 位来表示编码下一个非“0”系数所需要的位数，跟在它后面的是量化 AC 系数的数值。

JPEG 编码的最后一个步骤是把各种标记代码和编码后的图像数据组成一帧一帧的数据，这样做的目的是为了便于传输、存储和译码器进行译码，这样的组织的数据通常称为 JPEG 位数据流(JPEG bitstream)。

4.2.2 DCT 系数的分布特性

JPEG 压缩域隐写算法是通过修改图像 JPEG 压缩编码过程中量化后的 DCT 系数实现秘密信息嵌入的。如果在设计隐写算法时没有充分考虑 DCT 系数的分布特性，则必然导致隐写图像的 DCT 系数分布发生变化。

DCT 系数的统计分布问题一直为众多学者所关注[45-48]。一般认为，DCT 直流系数服从高斯分布，而对于交流系数，却有着不同的结论。基于中心极限定律，Pratt[45]猜想 AC 系数服从高斯分布(GD)。Reininger 和 Gibson[46]对图像的 DCT 系数做 Kolmogorov-Smirnov(KS)测试得出 AC 系数服从拉普拉斯分布(LD)的假设。Muller[47]将 DCT 系数看作广义高斯分布(GGD)获得了比拉普拉斯分布更小的 χ^2 统计量。Joshi 和 Fischer[48]通过 χ^2 拟合测试证实了 Muller 的结论，并对 DCT 系数的广义高斯模型和模型参数估计做了详细的讨论，指出高斯分布和拉普拉斯分布都是广义高斯分布的特殊情况，分别对应形状因子 $\nu=2$ 和 $\nu=1$。

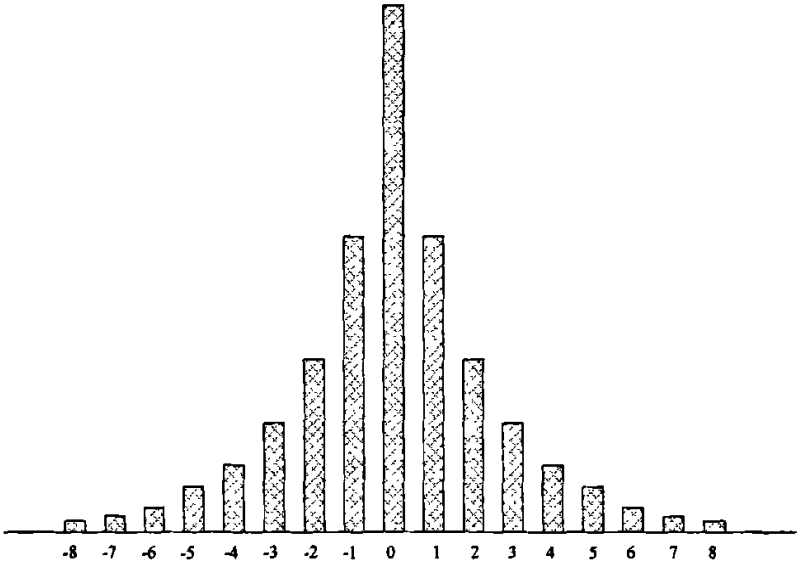


图 4.6 JPEG 图像 DCT 系数直方图特性示意

更为简便地，利用量化后的 DCT 系数直方图就可以估计 DCT 系数的分布特性，它反映了 DCT 系数的一阶分布特性。图 4.6 给出 JPEG 图像的 DCT 系数直方

图特性示意（如非特殊说明，本文出现的 DCT 系数直方图都是指量化后的 DCT 交流系数直方图），从图中我们可以总结量化后的 DCT 系数分布具有如下特点：

1) DCT 系数绝对值越大，其对应的直方图中的值越小，即出现的频率越低；

2) 随着 DCT 系数绝对值的升高，其出现次数下降的幅度减少；

3) DCT 系数直方图以 0 值为中心大致成对称分布。

勿庸置疑，一个安全的 JPEG 压缩域隐写算法必须尽可能的保持上述的 DCT 系数直方图的 3 个特点。然而现有的 JPEG 图像隐写算法，如 JSteg、F5 和 OutGuess 算法都会不同程度地改变 DCT 系数直方图特性，使得隐写分析者有迹可循。

4.2.3 JSteg 算法对直方图的改变

JSteg 算法由 D. Upham[41]提出，是将秘密信息顺序嵌入到 JPEG 图像 DCT 量化系数的 LSB 上，其实质是在 JPEG 压缩域上应用 LSB 替换技术。为了使隐写前后图像不发生明显视觉失真，不对值为 0 和 1 的系数嵌入信息。

JSteg 虽然实现简单，但是由于 LSB 算法的固有缺陷，导致隐写后 DCT 系数直方图发生了改变。图 4.7 显示了采用 JSteg 隐写前后图像 DCT 系数直方图的对比情况，可以很直观的看出 JSteg 算法对直方图的改变：它改变了直方图分布递减的趋势，使得相邻两系数的出现频率近似相等，正方向如系数 2 和 3、4 和 5……，负方向如－1 和－2、－3 和－4……依次类推。正是由于 JSteg 对 DCT 系数直方图的这一规律性的改变使得不少攻击方法都能够对其有效检测，如 χ^2 攻击，扩展的 χ^2 攻击以及各种直方图攻击。

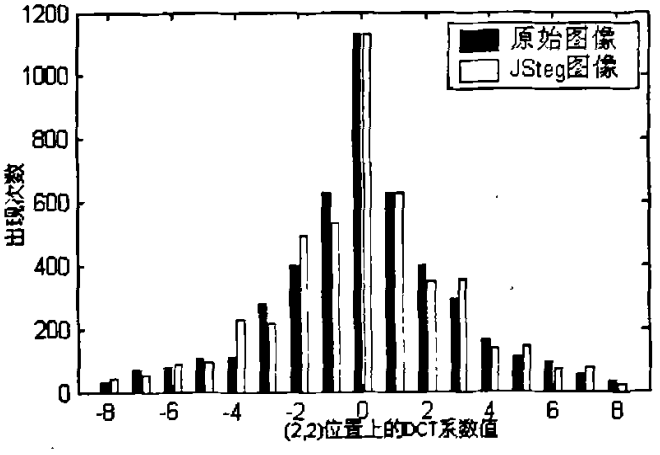


图 4.7 采用 JSteg 隐写前后 DCT 系数直方图对比

4.2.4 F5 算法对直方图的改变

F5 算法[42]由德国学者 Westfeld 于 2001 年提出，目标是为 JPEG 图像设计一种高隐藏容量并且安全的实用嵌入方法。F5 算法是将消息嵌入到随机选择的量化 DCT 系数中，并采用矩阵编码技术使得需要改变的系数数目最小化。矩阵编码的思想由 Crandall 提出，F5 算法实现了该思想，运用$(1, n, k)$编码大大提高了嵌入效率和隐藏容量。

F5 算法没有采用 LSB 替换的方法，而是通过对系数绝对值减 1 来嵌入信息。0 值系数不用于嵌入信息，如果嵌入时产生了一个新的值为 0 的系数，则需在下一个 DCT 系数上重新嵌入。Westfeld 证明了 F5 算法可有效保持 4.2.1 节指出的 DCT 系数的前两个特点，即维持了系数直方图递减的分布规律，但是它仍然会使直方图产生明显的改变，图 4.8 给出采用 F5 隐写前后图像 DCT 系数直方图的对比。首先由于缩减效应(Shrinkage)使得 0 值系数显著增加；其次除 0 外的其余系数都有不同程度减少，绝对值为 1 的系数显著减少。J. Fridrich 利用这两点改变设计了直方图攻击[13]，不但可以检测隐藏消息的存在而且能够得到嵌入消息的长度信息。

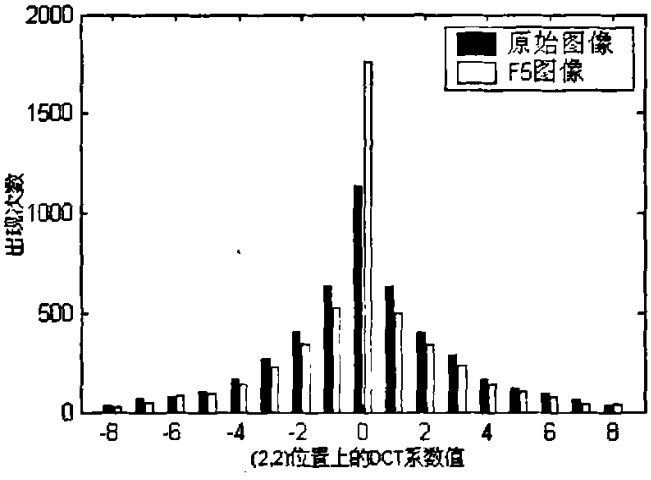
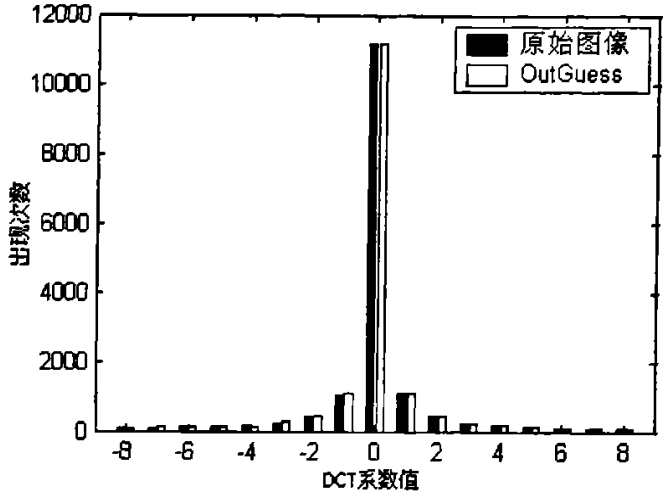


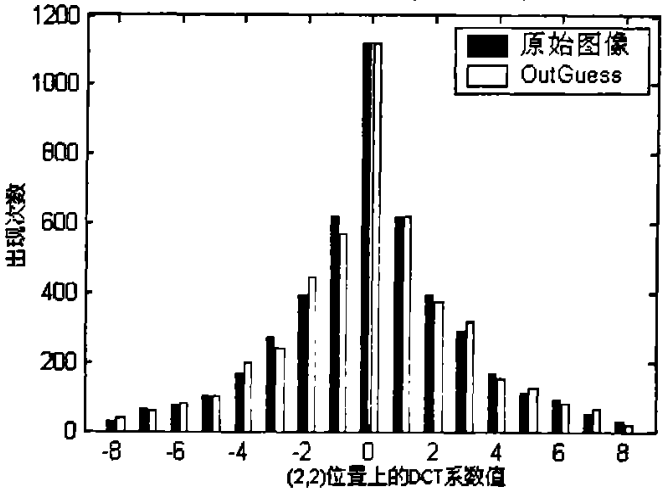
图 4.8 采用 F5 隐写前后 DCT 系数直方图对比

4.2.5 OutGuess 算法对直方图的改变

OutGuess 算法[43]由 Niels Provos 提出，它也采用了 LSB 嵌入技术，但是和 JSteg 不同的是，OutGuess 的隐写过程分为嵌入和矫正两个过程。嵌入过程包括识别冗余比特和选择比特嵌入两步，目的是通过识别不可察觉的分量，尽可能地减少对载体图像的改变，特别是那些具有高检测概率的分量。矫正过程是利用在嵌入信息时预留的一半系数进行直方图矫正，以减少统计失真。



a）总体系数直方图对比



b）低频系数直方图对比

图 4.9 采用 OutGuess 隐写前后的直方图对比

遗憾的是，OutGuess 的矫正步骤只能保证总体系数直方图特性，而各个频率上的系数直方图仍会产生较大差异，特别是低频系数直方图，如图 4.9 所示。其矫正直方图的代价却是放弃一半的系数嵌入信息，和 JSteg、F5 相比，相同条件下它的隐藏容量将减少一半。

综上所述，以上 3 种 JPEG 压缩域隐写算法都不同程度地改变了 JPEG 图像的 DCT 量化系数直方图。因此，可以利用直方图分布特性设计通用隐写分析技术。

4.3 基于直方图分布特性的通用隐写分析技术

图像 DCT 系数的概率密度参数需从大量图像中拟合得到，而且相对熵的计算包括各个频率系数的相对熵的叠加，运算量很大。如果以分块 DCT 代替 DCT，考虑块内不同频率位置上 DCT 系数之间的相互独立性，则可以认为相对熵等于块内不同频率系数分布的相对熵之和，而每个频率位置上的系数分布可由不同分块内该位置上的系数值拟合得到，或者直接利用直方图作为分布估计，这样可以大大减少计算量。

当然，利用分块 DCT 计算得到的相对熵并不等同于真正的隐写安全性定义，因为不同分块的相同频率位置上的系数之间并不是独立的，无论是用系数拟合还是直方图估计，都只能体现系数的一阶分布特性。但是，可以将它作为隐写安全性的必要条件，利用它设计通用隐写分析算法可以有效检测隐写算法的安全性。

4.3.1 基于直方图分布的特征提取

针对 JPEG 压缩域算法的隐写分析选取两组统计量作为特征。第一组为一阶分布的一、二、三、四阶累积量，用来量化隐写信息前后 DCT 系数直方图分布的差异。首先根据 DCT 系数直方图估计 DCT 系数的一阶分布。设图像大小为 $M\times N$，由 JPEG 压缩产生的 DCT 分块数为 $L=(M\times N)/(8\times 8)$，以 $H[k]$ 表示 DCT 总体系数（不包括直流系数）直方图，k 取遍所有可能出现的 DCT 交流系数值，则可以将 $H[k]/(MN-L)$ 作为 DCT 总体系数分布的一个估计。此外，考虑到一些算法能够保持总体系数直方图，但对每个频率位置上的直方图却不能保持，例如 OutGuess，因此，需要估计各个频率位置上的系数分布。以 $H_{u,v}[k]$ 表示 (u,v) 频率位置上的直方图，其中 $1\le u,v\le 8$ 且 $(u,v)\ne(1,1)$，k 取遍 (u,v) 位置上所有可能的系数值。将 $H_{u,v}[k]/L$ 作为 (u,v) 位置上 DCT 系数的一阶分布估计。这里，因为中高频上的非 0 系数太少，其在统计上不重要，因此只估计低频系数分布。为了将特征维数控制在一定范围内，只估计 (1,2)、(2,1) 和 (2,2) 频率位置上的系数的一阶分布。

然后，分别求取上述一阶分布的均值 (μ)、方差 (σ)、偏度 (γ^3) 和峰度 (γ^4) 作为隐写分析特征，偏度和峰度的计算如式(4-3)：

$$\gamma^3=\frac{E(x-\mu)^3}{\sigma^3},\quad \gamma^4=\frac{E(x-\mu)^4}{\sigma^4} \tag{4-3}$$

第二组是选取特殊值的分布概率。由 4.2 节的分析，可以发现各种 JPEG 压缩域隐写算法对于某些特殊值的一阶分布改变非常大，例如 JSteg 对值为 −1、−2、

1、2 的直方图分布改变；F5 对值为 0、−1、1 的直方图分布改变等。抓住这一特点，我们选取值为−2、−1、0、1、2 在总体系数直方图上的分布概率作为第二组隐写分析特征。

这样，数字图像的总体 DCT 系数一阶分布 $H[k]/(MN-L)$ 和在 (1,2)、(2,1) 和 (2,2) 频率位置上的 DCT 系数分布 $H_{1,2}[k]/L$、$H_{2,1}[k]/L$、$H_{2,2}[k]/L$ 的均值、方差、偏度、峰度（16 个特征向量）构成第一组特征向量；值为−2、−1、0、1、2 的 DCT 系数在总体 DCT 系数直方图上的分布概率（5 个特征向量）构成第二组特征向量。综上所述，两组共 21 个特征构成最终的隐写分析特征向量。

4.3.2 基于方差分析的特征筛选

由 4.2 节的分析，可以看到不同的 JPEG 压缩域隐写算法对直方图分布的改变不同，例如，F5 算法能够保持直方图的大体形状，即关于 0 大致呈对称分布，因而从载体图像和隐写图像计算得到的均值特征将近似为 0，其对于分类是没有意义的。不去除这些特征的影响，将不利于最终的检测。因此，在训练分类器之前，需要对特征进行有效性分析，剔除无效特征，以达到最好的分类效果。本文采用方差分析技术[49]实现这一目的。

考虑单因素方差分析，因素(factor)指影响实验指标的条件。因素水平(level of factor)指实验因素所处的某种特定状态或处理组，简称水平。

设因素 A 有 s 个水平 $A_1, A_2, ..., A_s$，在水平 $A_j (j=1,2,...,s)$ 下进行了 $n_j (n_j \geq 2)$ 次独立实验。假设各个水平 $A_j (j=1,2,...,s)$ 下的样本 $x_{1j}, x_{2j}, ...x_{n_j j}$ 来自具有相同方差 σ^2，而均值分别为 $\mu_j (j=1,2,...,s)$ 的正态总体，μ_j 与 σ^2 未知，且设不同水平 A_j 下的样本之间相互独立。方差分析的任务是对 s 个正态总体检验假设：

$$\begin{aligned} &H_0: \mu_1 = \mu_2 = ... = \mu_s, \\ &H_1: \mu_1, \mu_2, ..., \mu_s \text{不全相等} \end{aligned} \tag{4-4}$$

并对未知参数作出估计。这里，因素 A 就是我们要分析的特征，两类图像（载体图像和隐写图像）是它的两个水平。

为导出检验统计量，引入可以反映全部实验数据差异的总平方和：

$$\begin{aligned} S_T &= \sum_{j=1}^{s}\sum_{i=1}^{n_j}(x_{ij}-\bar{x})^2 \\ &= \sum_{j=1}^{s}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_{\cdot j})^2 + \sum_{j=1}^{s} n_j(\bar{x}_{\cdot j}-\bar{x})^2 \\ &= S_E + S_A \end{aligned} \tag{4-5}$$

S_T又称为总变差，它可以分解为 S_E、S_A之和。S_E表示由随机误差引起的样本观察值和样本均值的差异，称为误差平方和。S_A反应了样本均值和数据总平均的差异，称为因素 A 的效应平方和。可以证明：当 H_0 为真时，$S_A/\sigma^2 \sim \chi^2(s-1)$，$E(S_A/s-1)=\sigma^2$，而当 H_1 为真时，$E(S_A/s-1)>\sigma^2$；对于 S_E，不管 H_0 是否为真，都有 $S_E/\sigma^2 \sim \chi^2(n-s)$。又由 S_E 和 S_A 的独立性可知，当 H_0 为真时：

$$\frac{S_A/(s-1)}{S_E/(n-s)}=\frac{\frac{S_A}{\sigma^2}/(s-1)}{\frac{S_E}{\sigma^2}/(n-s)} \sim F(s-1,n-s) \tag{4-6}$$

由此得到检验问题（4-4）的拒绝域为：

$$F=\frac{S_A/(s-1)}{S_E/(n-s)} \geq F_\alpha(s-1,n-s) \tag{4-7}$$

其中，α 是显著性水平（一般取 0.05）。

利用上述检验假设对所选取的隐写分析特征进行有效的筛选，剔除无效的特征，可以进一步提高分类效果。

4.3.3 实验与结果分析

对 JSteg、F5 和 OutGuess 隐写算法分别用本章算法进行了实验测试。分类器仍然采用第三章提到的 Fisher 线性判别(FLD)分类器。实验用的图像为 JPEG 灰度图像。实验时首先取 100 幅 JPEG 图像作为未嵌入秘密信息的载体图像，分别统计直方图特性并计算特征向量；然后用 Jsteg (或 F5 或 OutGuess)软件在每幅图像中嵌入秘密信息形成隐写图像，并计算每幅隐写图像的特征向量。注意，为避免统计偏差，隐写图像和载体图像使用相同的压缩因子。这样，200 幅图像的特征向量构成了训练样本集。类似地，另取 200 幅图像构成测试样本集。

分别测试当嵌入比例 p=0.5、0.25 bpc（比特/非零交流系数）时算法的检测性能。表 4.1 给出了实验测试结果。

表 4.1 通用隐写分析测试结果

隐写算法	嵌入比例 (bpc)	检测率(%)	虚警率(%)	漏检率(%)
JSteg	0.5	97.5	3.0	2.0
	0.25	84.0	17.0	15.0
F5	0.5	96.0	4.0	4.0
	0.25	80.5	21.0	18.0
OutGuess	0.5	94.0	7.0	5.0
	0.25	69.5	32.0	29.0

从表 4.1 可以看出，在秘密信息的嵌入比例为 0.5bpc（比特/非零交流系数）的情况下，本文算法对 JSteg、F5 以及 OutGuess 隐写算法的检测效率均在 90%以上，高于第二章提及的隐写分析评价指标中的 85%，表明检测性能良好；在秘密信息的嵌入比例为 0.25bpc 的情况下，检测效率有所下降，但漏检率均低于虚警率，符合算法设计原则，即在虚警率和漏报率无法同时减少的情况下，着重减少漏报率。这也正说明了隐写分析的难点所在，即随着嵌入比例的降低，隐写操作对图像特征的改变也随之减弱，提高了隐写分析的难度。

实验结果表明，本文给出的 JPEG 压缩域的通用隐写分析算法获得了较好的检测效果。

4.4本章小结

本章首先介绍JPEG 压缩域通用隐写分析技术的研究现状；分析现有的JPEG 压缩域隐写算法对DCT 系数直方图的改变；在此基础上，以直方图作为DCT 系数的一阶分布估计，提出将DCT 系数直方图分布的均值、方差、偏度、峰度以及特殊系数值的直方图分布概率作为特征，并基于方差分析对特征进行有效性分析，设计实现了一种通用隐写分析算法。对JSteg、F5 和OutGuess 软件的实验结果表明，该算法的检测性能较好。

第五章 总结与展望

本文主要对基于图像的通用隐写分析技术进行了研究：针对空域隐写算法，利用差分直方图分布的高阶统计矩和差分直方图曲线下特殊区域之间面积的比重来量化隐写前后的差分分布差异，利用 Fisher 线性分类器检测是否存在隐藏信息，设计实现了一种空域通用隐写分析方案；针对 JPEG 压缩域隐写算法，将 DCT 系数直方图分布的均值、方差、偏度、峰度以及特殊系数值的直方图分布概率作为特征，并利用方差分析对提取的特征进行有效性分析，设计实现了一种检测性能较好的 JPEG 压缩域通用隐写分析方案。

纵观国际上隐写分析领域的研究进展，结合本文的研究工作，对通用隐写分析技术中的关键问题和今后的研究方向得出了以下一些思考：

通用隐写分析技术从寻找隐藏数据对载体产生的影响这一本质特性出发，更有助于隐写检测这一目标的实现。通用隐写分析技术的关键问题如下：

1. 特征提取。需根据不同嵌入机制对隐写技术进行分类，以提高通用检测算法的性能；好的特征必须对嵌入改变敏感而对图像内容不敏感，因此必须消除图像数据对隐写检测的影响，可以利用滤波技术估计隐写噪声或去相关技术实现；使用累积量、原点矩、中心矩等高阶统计量以提高检测效率。
2. 特征的有效性分析。对提取的特征进一步分析其有效性和做去相关处理，包括方差分析、主成分分析等技术，以便进一步提高检测的效率和性能。
3. 分类器的设计。分类器的设计对于最终的检测性能有很大的影响，需要进一步研究与所选特征分布相关的分类器，如贝叶斯(Bayes)分类器、支持向量机(SVM)分类器等。

此外，通用隐写分析技术还存在着一些问题，将成为今后的研究方向：

1. 如何将通用隐写分析方法和专用隐写分析方法有机的融合，最有效的策略是什么；
2. 分析数字图像、嵌入信息与隐写算法之间的相关性，找到更多的数字图像特征描述量及数字图像的统计特征用于隐写检测；
3. 如何在网络环境下迅速、准确的实现隐写分析；
4. 基于图像的隐写分析技术作为一个系统工程，还包括如何提取隐写信息、破坏存在的隐写通信等内容，尤其在隐写信息提取方面，可能需要与密码技术、信息安全领域的许多方面相互结合共同完成。

综上所述，通用隐写分析是极富挑战性的研究内容，现有的成果无疑只是揭开了冰山一角，相信随着研究的不断深入，这一领域必将取得长足的进步和发展。

致 谢

深深感谢我的导师田玉敏教授！我取得的每一点成绩都离不开田老师的教诲和点拨，是她引领我进入信息隐藏这一神奇又充满活力的新兴领域，并以渊博的学识、严谨的学风、敏锐的洞察力给了我莫大的帮助和启发。这无疑也将对我今后的工作和学习产生积极而长远的影响。

感谢那些曾经和正在计算机外围设备研究所学习和工作的师兄弟、师姐妹们，他们分别是桑小川、司蓁、龚岩、郑伟、张淑斌、陈安、田广昊、陈琦光，和他们无拘无束的交流和讨论，使我在学业上深受裨益。

深深感谢我的父母和家人，他们为我的成长付出了很多心血，对我的理解和支持让我能够面对一切困难和挫折，我将永远的爱他们。

深深感谢我的妻子叶茂女士。感谢她不断地鼓励、开导和支持我，在我苦闷彷徨时，用爱激励我，给我信心和勇气；在我取得进步时，用心鼓舞我，与我分享成功的喜悦！

再次对关心我、爱护我和帮助我的师长、家人和朋友表示真挚的谢意！

参考文献

[1] 杨义先, 钮心忻, 任金强. 信息安全新技术. 北京邮电大学出版社, 2002

[2] 夏煜, 朗荣玲等. 基于图像的信息隐藏分析技术综述[J]. 计算机工程, 2003

[3] Jack Kelley. Terrorist instructions hidden online. USA TODAY, Feb 5, 2001. Available at: http://www.usatoday.com//tech/news/2001-02-05-binladen.htm

[4] 王朔中, 张新鹏, 张开文. 数字密写和密写分析. 清华大学出版社, 2005

[5] Simmons G J. The prisoners' problem and subliminal channel. In Advances in Cryptology, Proc. of CRYPTO'83, Plenum Press, 1984, pp.51-67

[6] Birgit Pfitzmann, Information hiding terminology, Proc. of 1st International Workshop: Information Hinding-IH'96, LNCS 1174, Springer- Verlag, 1996, pp.347-350

[7] Cachin C. An information-theoretic model for steganography, Proc. 2nd, International Workshop on Information Hiding, Portland, Oregon, Lecture Notes in Computer Sciences, Springer-Verlag, 1998, pp.306-318

[8] Shannon C E. A mathematical theory of communication, Bell System Technical Journal, Vol.27, No.4, 1948, pp. 397-423

[9] Shannon C E. Communication theory of secrecy systems, Bell System Technical Journal, Vol.28, No.4, 1948, pp. 656-715

[10] Chandramouli R, Memon N. Steganography capacity: A steganalysis perspective, SPIE Security and Watermarking of Multimedia Contents V, Vol. 5020, 2003

[11] Chandramouli R. Data hiding capacity in the presence of an imperfectly known channel, SPIE Proceedings of Security and Watermarking of Multimedia Contents II, vol. 4314, 2001, pp. 517-522

[12] Scott Craver. On Public-key steganography in the presence of an active warden, Proc. of 2nd International Workshop: Information Hinding-IH'98, LNCS 1525, Springer- Verlag, 1998, pp.355-368

[13] Johnson N. F, Jajodia S. Steganalysis of images created using current steganography software, in David Aucsmith (eds.): Information Hiding, LNCS 1525, Springer-Verlag Berlin Heidelberg, 1998, pp. 32-47

[14] Westfeld A and Pfitzmann A. Attacks on steganographic systems. Proc. 3rd Int'l Workshop Information Hiding. Springer-Verlag, 1999, Dresden, Germany, 61-76

[15] Fridrich J, Goljan M, Du R. Detecting LSB steganography in color and gray-scale

images. Magazine of IEEE Multimedia, Special Issue on Security, 2001, 8(4): 22-28

[16] 张涛, 平西建. 基于差分直方图实现LSB信息伪装的可靠检测 [J]. 软件学报. 2004, 15(1) :151-158 .

[17] Fridrich J, Goljan M, Hogea D. Steganalysis of JPEG images: breaking the F5 algorithm. Proc. 5th Int'l Workshop Information Hiding. Springer-Verlag, Noordwijkerhout, the Netherlands, 2002, 310-323

[18] Fridrich J, Goljan M, Hogea D. Attacking the OutGuess. Proc. of the ACM Workshop on Multimedia and Security 2002. France, Dec. 6, 2002

[19] Chandramouli R. A Mathematical Approach to Steganalysis. http://citeseer.nj.nec.com/chandramouli02mathematical.html

[20] Avcibas I, Memon N, sankur B. Steganalysis using image quality metrics. Security and Watermarking of Multimedia Contents, San Jose, Ca., Feruary 2001

[21] Avcibas I, Memon N, sankur B. Steganalysis using image quality metrics. IEEE transactions on Image Processing, January 2003

[22] 边肇祺, 张学工等编著. 模式识别（第二版）. 北京: 清华大学出版社, 2000

[23] Pevny T, Fridrich J. Towards multi-class blind steganalyzer for JPEG images, International Workshop on Digital Watermarking, LNCS vol. 3710, Springer-Verlag, 2005, pp. 39–53

[24] Kharrazi M, Sencar H T, Memon N D. Benchmarking steganographic and steganalytic techniques, Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VII, vol. 5681, San Jose, CA, January 16–20, 2005, pp. 252–263

[25] Aura T. Invisible communication. EET 1995, Technical Report. Helsinki University of Technology, Finland, Nov. 1995. http://deadlock.hut.fi/ste/

[26] Elke Franz. Steganography preserving statistical properties. Proc. of the 5th International Workshop on Information Hiding, Lecture Notes on Computer Science, vol. 2578, Noordwijkerhout, The Netherlands. Oct. 7-9, 2002

[27] Chandramouli R, Grace Li, Memon N. Adaptive steganography. Security and Watermarking of Multimedia Contents IV, Proc. of SPIE, vol. 4675, 2002

[28] Sharp T. An implementation of key-based digital signal steganography, in I. Moskowitz (ed.): Information Hiding. 4th International Workshop. Lecture Notes in Computer Science, vol. 2137, Springer-Verlag New York, 2001, pp. 13–26. Hide 2.1, 2001, Available at: http://www.sharpthoughts.org

[29] Fridrich J, Goljan M. Digital image steganography using stochastic modulation,

SPIE Symposium on Electronic Imaging, San Jose, CA, 2003

[30] Holotyak T, Fridrich J, Soukal D. Stochastic approach to secret message length estimation in ±k embedding steganography, in E. Delp et al. (eds.): Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VII, 2005

[31] Fridrich J, Soukal D, Goljan M. Maximum likelihood estimation of length of secret message embedded using ±k steganography in spatial domain. Available at: http://www.ws.binghamton.edu/fridrich

[32] Junhui He, Jiwu Huang, Guoping Qiu. A new approach to estimating hidden message length in stochastic modulation steganography. IWDW 2005, 2004. 25, pp. 1-14

[33] Farid H. Detecting hidden messages using higher-order statistical models. Proc. Int'l Conf. Image Processing, IEEE Press, Rochester, NY, 2002, 905-908

[34] Farid H, Lyu S. Detecting hidden messages using higher-order statistics and support vector machines. Proc. 5th Int'l Workshop on Information Hiding, Springer-Verlag, Noordwijkerhout, the Netherlands, 2002, 340-354

[35] Harmsen J J, Pearlman W A. Steganalysis of additive noise modelable information hiding, Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents V, 2003, pp. 131–142

[36] Ker A. General framework for structural steganalysis of LSB replacement, in M. Barni et al. (eds.): 7th International Workshop on Information Hiding, LNCS vol. 3727, Springer-Verlag, Berlin, 2005, pp. 296–311

[37] Celik M, Sharma G, Tekalp A. Universal image steganalysis using rate-distortion curves, in E. Delp et al. (eds.): Proc. SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VI, vol. 5306, 2004, pp. 467–476

[38] Holotyak T, Fridrich J, Voloshynovskiy S, Blind statistical steganalysis of additive steganography using wavelet higher order statistics, 9th IFIP TC-6 TC-11 Conference on Communications and Multimedia Security, LNCS vol. 3677, Springer-Verlag, Berlin, 2005, pp. 273–274

[39] Goljan M, Fridrich J, Holotyak T. New blind steganalysis and its implications, Proc. SPIE Electronic Imaging, Photonics West, January 2006. Available at: http://www.ws.binghamton. edu/fridrich/Research/EI06-6072-1.pdf

[40] Huang J, Mumford D. Statistics of natural images and models. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999, pp. 541-547

[41] Derek Upham, JPEG-JSteg-V4. Available at: http://www.funet.fi/pub/crypt/steganography/ jpeg-JSteg-v4.diff.gz

[42] Westfeld A. F5 -- a steganographic algorithm: High capacity despite better steganalysis, 4th International Workshop on Information Hiding, 2001

[43] Provos N. Defending against statistical steganalysis, 10th USENIX Security Symposium, 2001

[44] Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes, in J. Fridrich (ed.): Information Hiding, 6th International Workshop. Lecture Notes in Computer Science, vol. 3200, Springer-Verlag New York, 2005, pp. 67–81

[45] Pratt W K. Digital image processing. New York: Wiley, 1978

[46] Reininger R, Gibson J. Distributions of the two-dimensional DCT coefficients for images, IEEE Trans. Commun., vol. COM-31, June 1983, pp. 835–839

[47] Müller F. Distribution shape of two-dimensional DCT coefficients of natural images, Electron. Letter., vol. 29, no. 22, Oct. 1993, pp. 1935–1936

[48] Joshi R L, Fischer T R. Comparison of generalized gaussian and laplacian modeling in dct image coding, IEEE Signal Processing Letters, vol. 2, no. 5, 1995, pp. 81 – 82

[49] 魏宗舒等著. 概率论与数理统计教程. 北京：高等教育出版社, 1983, 241-253

[50] 杨淑莹编著. 图像模式识别——VC++技术实现. 北京：清华大学出版社；北京交通大学出版社，2005,7

研究成果

[1] 王剑，田玉敏．数字图像的信息隐藏分析技术研究．西安电子科技大学 2006 年研究生学术年会．2006,11