# 第四章 数字隐写对抗技术

周满
15271802577
zhouman@hust.edu.cn
华中科技大学网络空间安全学院

# 数字隐写对抗技术

- 一、隐写分析概述
- 二、专用隐写分析算法
- 三、通用隐写分析算法
- 四、隐写分析性能评估

# 隐写分析概述

## 隐写对信息安全的挑战

- 网络与信息技术高速发展的同时，伴随着层出不穷的网络信息安全问题

- 科学技术是一把双刃剑，信息隐藏技术也不例外，对应的隐写分析技术近年也得到快速发展

# 隐写分析概述

## 隐写对信息安全的挑战

- 近年来有关安全部门掌握的情况表明，恐怖分子和间谍机构已经在利用隐秘通信技术从事危害国家和社会安全的活动
  - 2011年5月，德国联邦刑警的计算机刑侦专家成功解开了基地组织成员通过隐写技术隐藏在一个色情视频中的141个文本文档，该文档中包括大量基地组织的行动报告和未来行动规划
- 从目前互联网管理中内容安全的意义上讲，要求能够对隐藏信息进行检测和提取

# 隐写分析概述

## 隐写对信息安全的挑战

- 正向：隐写使得保密通信更安全
- 逆向：滥用隐写技术对社会安全造成威胁

- 因此利用隐写传递恐怖信息的宣传引起广泛关注，隐写和反隐写受到有关部门和研究者高度重视

# 隐写分析概述

## 隐写分析的目的

- 保证对互联网信息的监控
- 遏制隐写术非法应用
- 打击恐怖主义威胁
- 维护国家和社会安全

自现代隐写术从上世纪90年代初开始快速发展以来，对其反向研究的隐写分析技术一直是信息安全领域关注的热点之一

# 隐写分析概述

## 隐写分析技术原理和模型

- 所有的信息伪装和数字水印技术能描述成下面的简单公式:

C = P + t

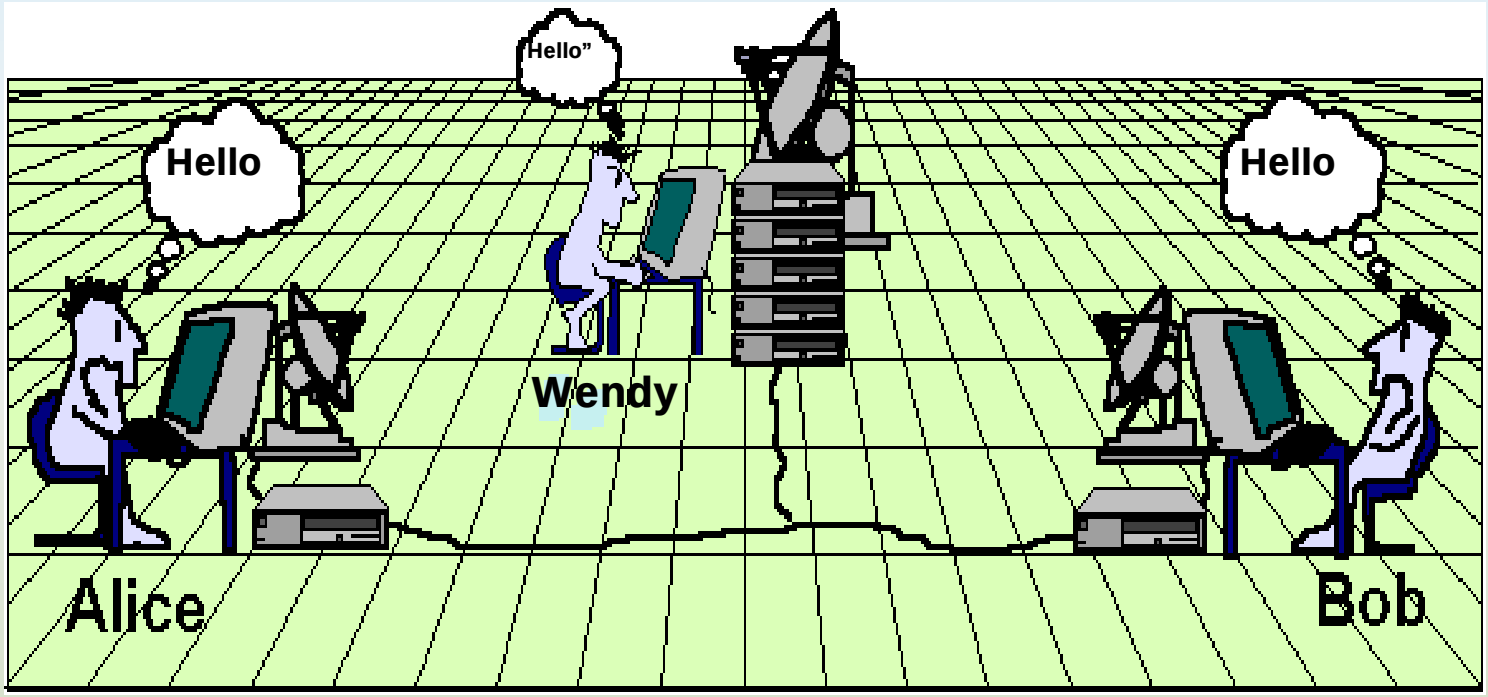
C：可用于信息伪装的载体

P：对该部分操作时会引起可感知降质

t：对该部分信息量处理时不会引起感觉上降质

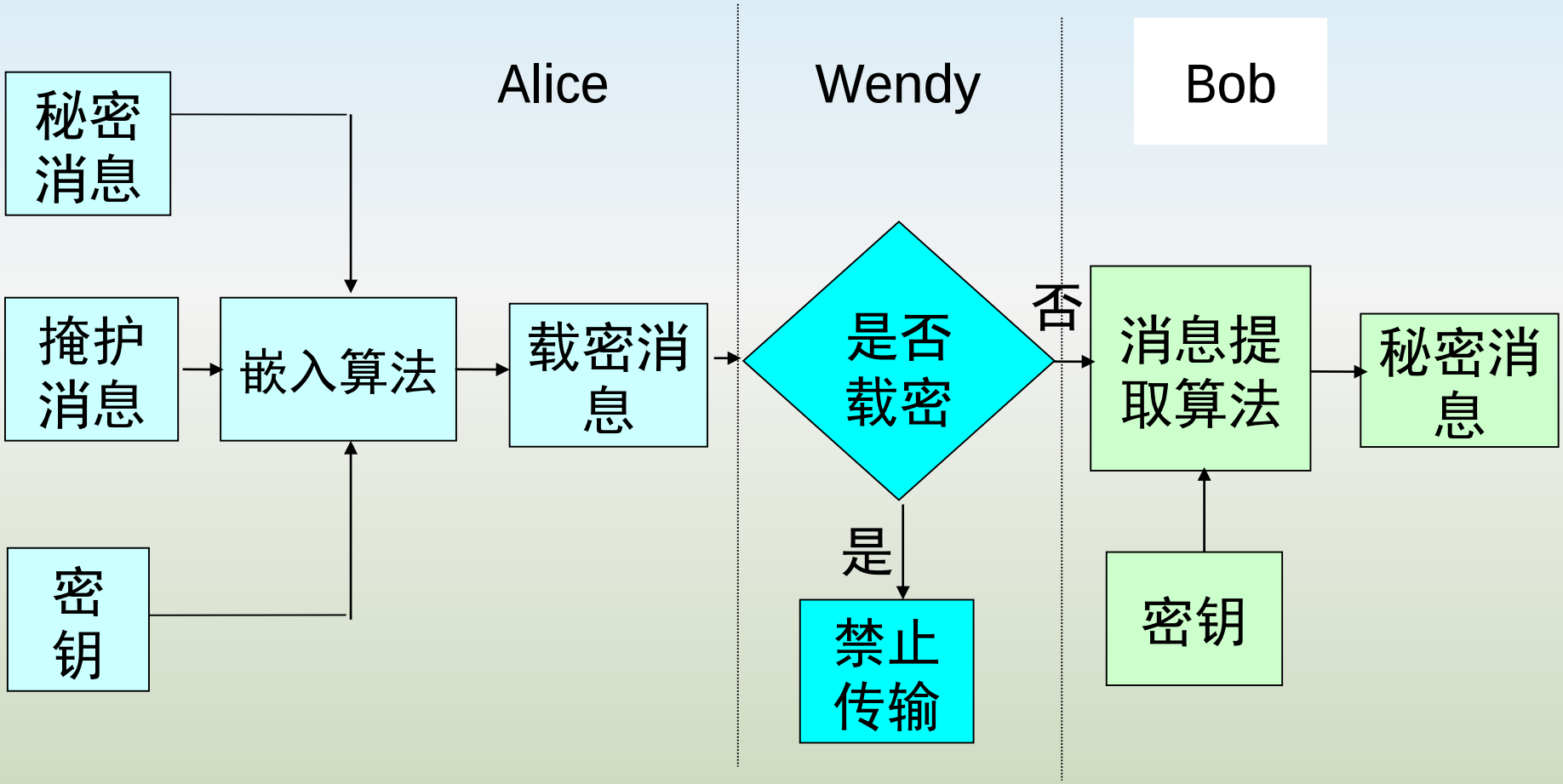
# 隐写分析概述

## 隐写分析技术原理和模型



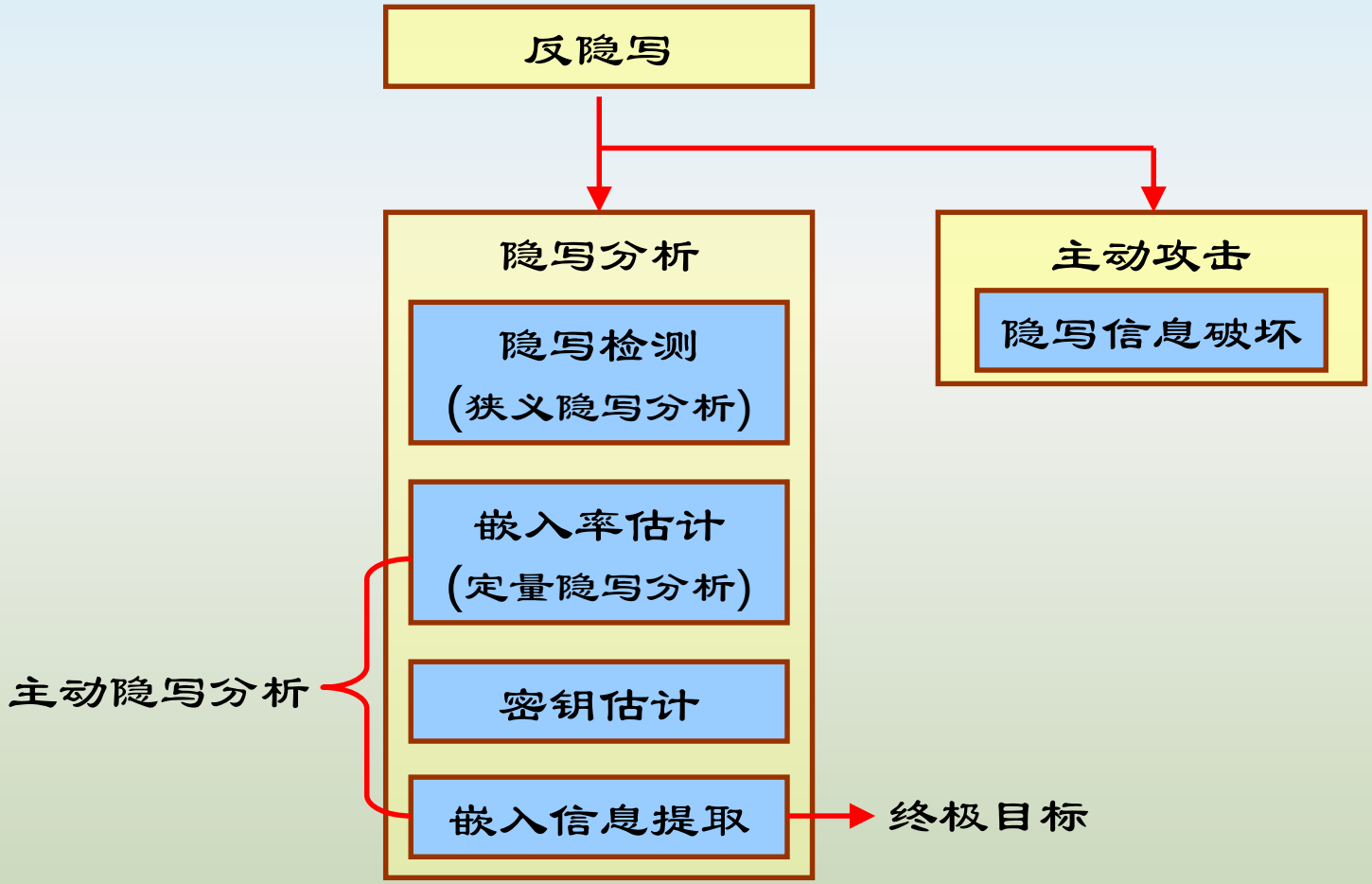
# 隐写分析概述

## 隐写分析技术原理和模型



# 隐写分析概述

## 隐写分析技术原理和模型



# 隐写分析概述

## 隐写分析手段

1.主动攻击

可以通过一次简单的压缩处理（鲁棒性攻击）来尝试对图像隐秘消息的破坏攻击

2.隐写信息检测

对经过的可疑图像载体进行检测，确定是否存在隐秘消息

3.隐写信息提取

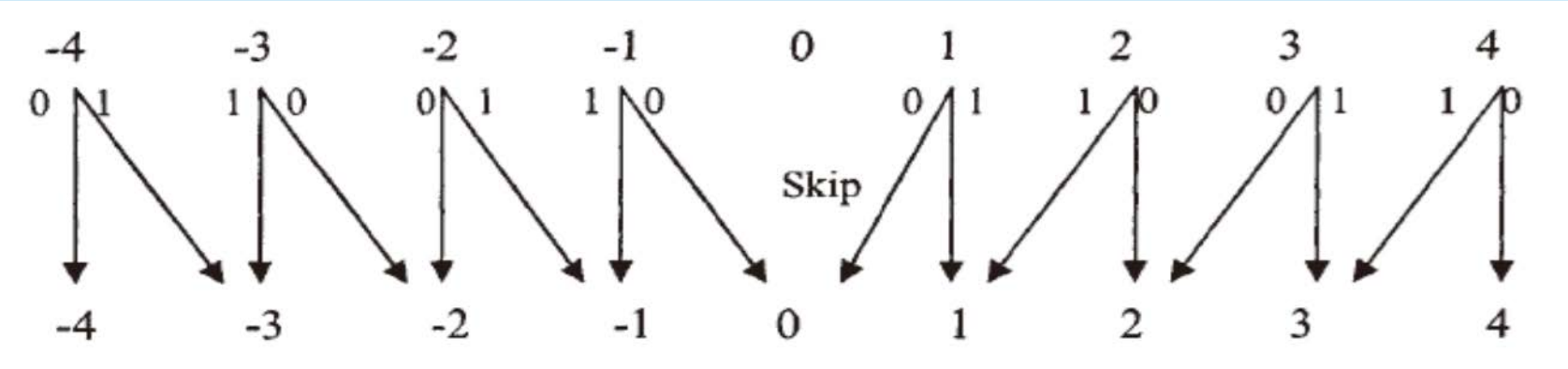
在检测到隐秘消息存在的前提下，尝试提取攻击（估计嵌入的秘密信息的长度、嵌入的位置、以及嵌入算法使用的密钥和某些参数）来破译出隐秘信息的内容

# 隐写分析概述

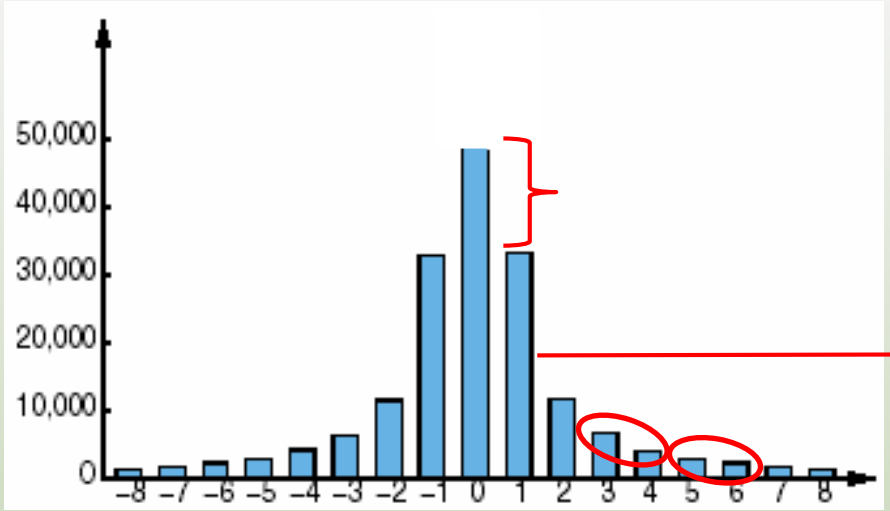
## 隐写信息提取技术

- 提取攻击的目的是从隐写载体中获取全部精确信息；
- 对于基于密钥的隐写，提取攻击的终极任务是找到恢复隐写密钥的方法，这在本质上是一种密码分析；
- 隐写与密码间的差异又使得传统的密码分析手段并不直接适用；
- 在提取攻击上的每一点突破都很困难；
- 现有相关研究凤毛麟角。

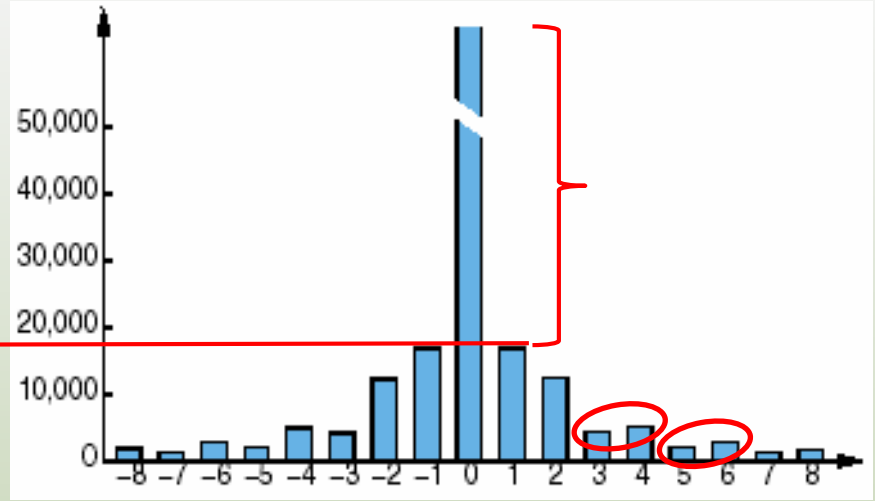
# 隐写分析技术：寻找不同点




$F_3$隐写前后图像DCT系数直方图的变化

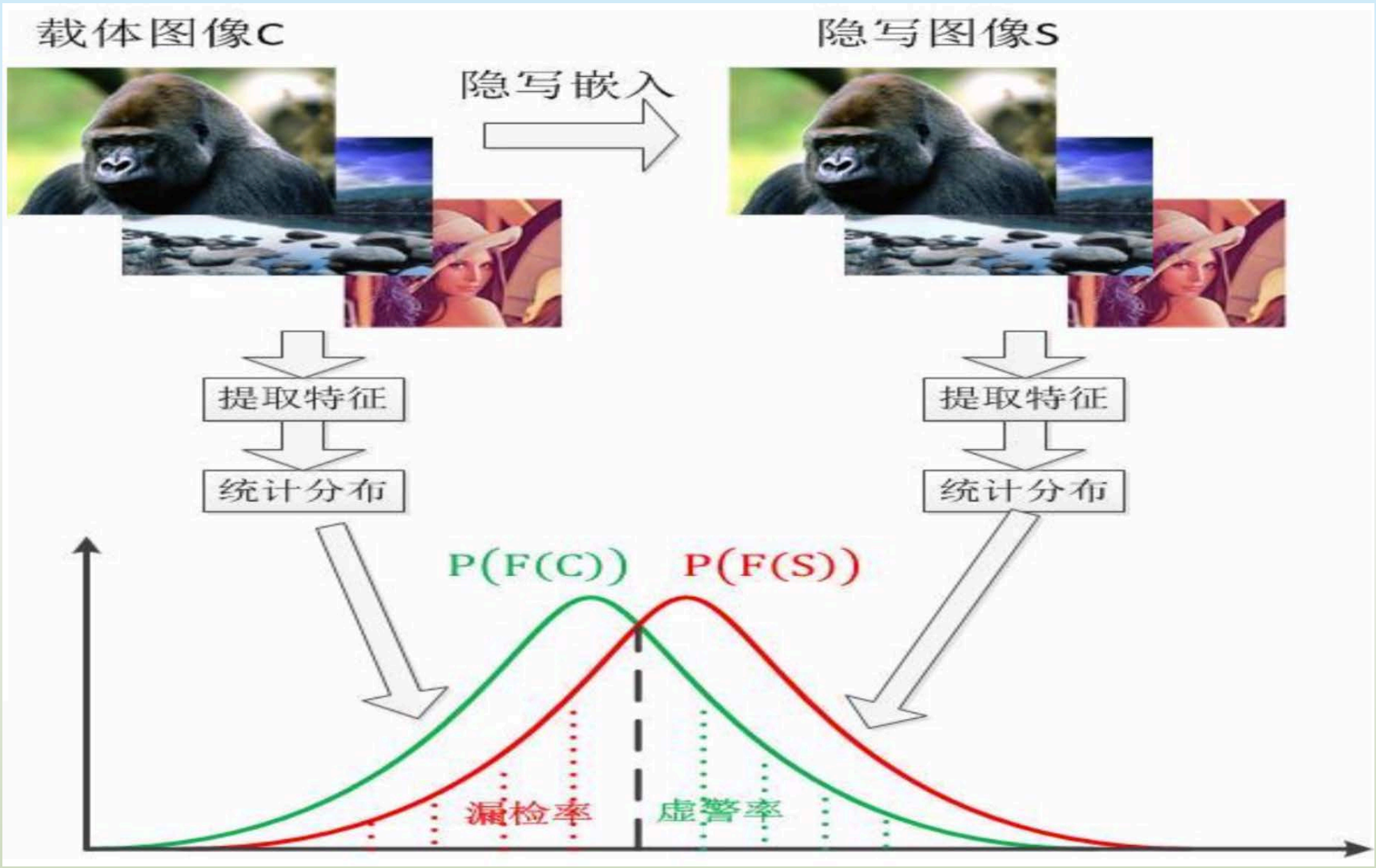



隐写前




隐写后

# 隐写分析技术：寻找不同点



# 隐写分析算法

## 隐写分析分类（1）

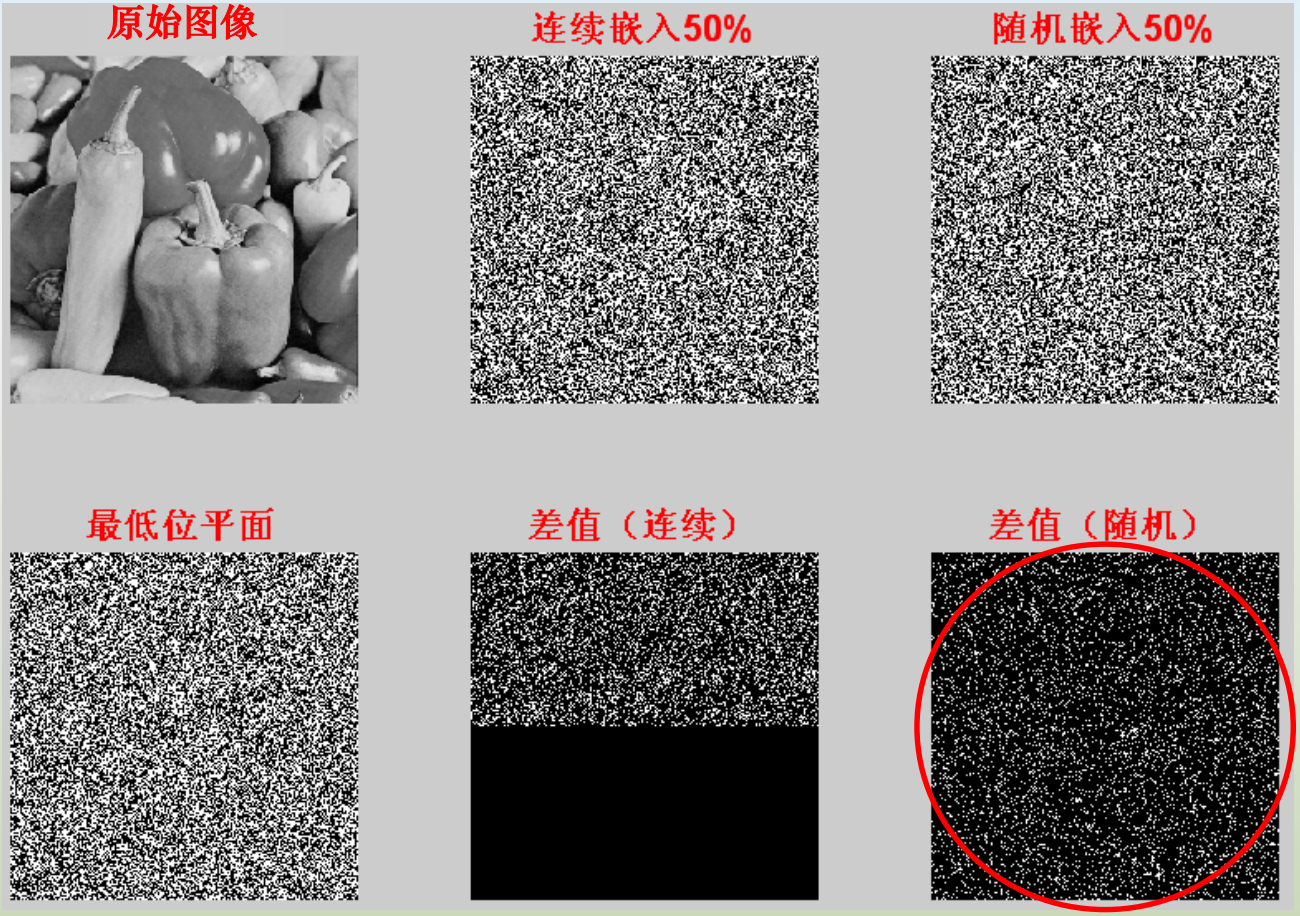
按照隐写分析算法的实现技术，分为3类：

- 感官检测法：比较原始载体和隐写载体，注意人眼可见的差异，但是如果没有原始载体，这种噪声就会作为载体的一个有机部分而不被注意到
- 标识特征检测法：某些隐写软件在隐写图像中留下标识特征，通过分析待检测对象中是否出现该类标识特征来实现检测
- 统计检测法：将原始载体的理论期望频率分布和待检测载体中的样本分布进行比较，从而找出差别的一种检测方法，主要包括特定隐写检测与通用盲检测

# 隐写分析算法

## 感官检测法

LSB替换隐写前后图像LSB平面的变化



# 隐写分析算法

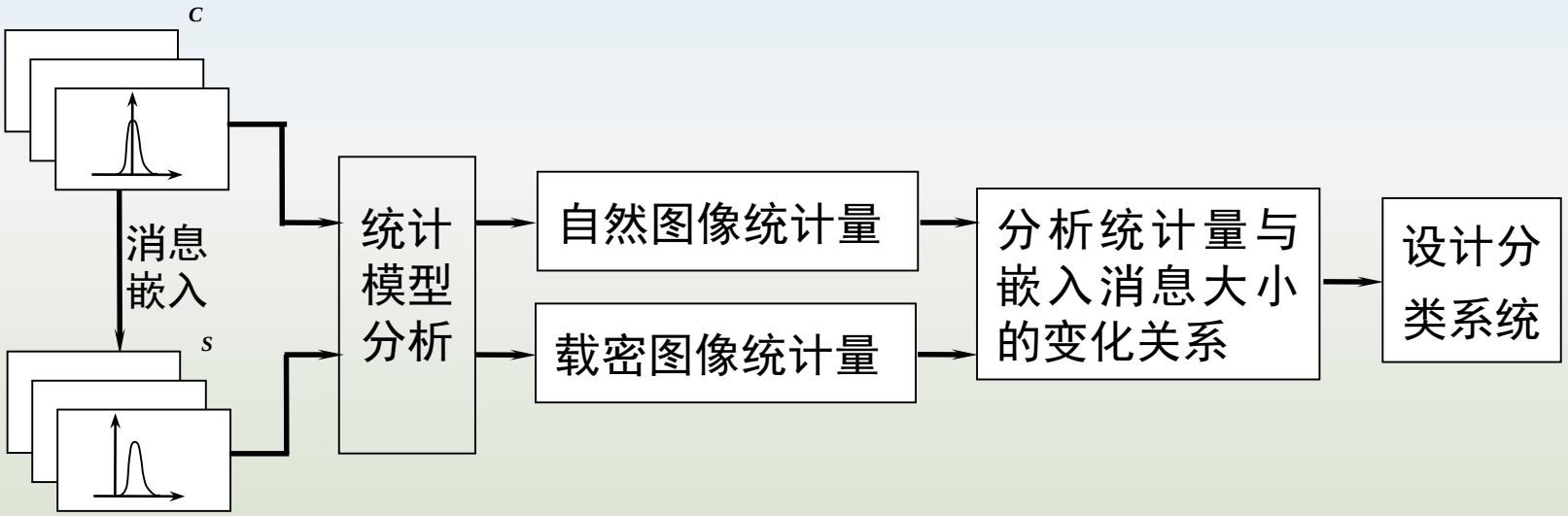
## 标识特征检测法

- 分析待检测对象是否出现大量隐写软件的标识特征:
- 隐写软件In the picture在嵌入消息结构头部存在标识特征“ITP”；
- 隐写软件The third eye在文件中存在标识特征“www.binary-techNologies.com”；
- 但是标识特征检测法只适用于已知的算法和工具，对于未知的隐写算法不能奏效。

# 隐写分析算法

## 统计检测法

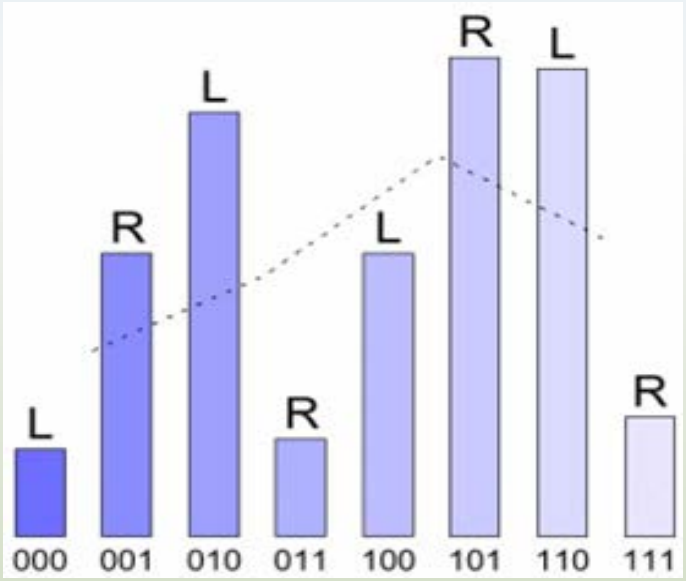
自然图像统计假设



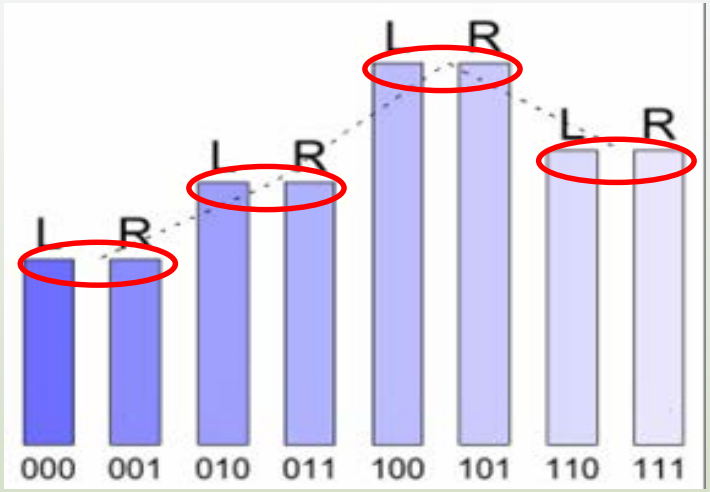
# 隐写分析算法

## 统计检测法

LSB替换隐写前后图像直方图的变化




载体图像直方图 (部分)




隐写图像直方图(部分)

# 隐写分析概述

## 隐写分析分类（2）

根据隐写分析工作的程度，可分为：

- 破坏隐秘消息
- 检测并确定可疑对象
- 判定隐写术采用的算法、恢复密钥、估计隐藏信息的长度
- 定位/提取/伪造隐秘信息

# 隐写分析概述

## 隐写分析分类（3）

根据隐写分析适用隐写算法的程度，分为：

- 专用隐写分析
  - 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
- 通用隐写分析
  - 利用多种隐写术对载体统计特性造成的一般意义上的变化而设计，因此可用于攻击多种隐写术

# 隐写分析概述

## 隐写分析分类（4）

根据信息隐藏域，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，分为：

1）空域隐写分析

LSB

2）变换域隐写分析

DCT，DWT，DFT

# 隐写分析概述

## 隐写分析分类（5）

按照载体类型的隐写分析技术分类，包括：

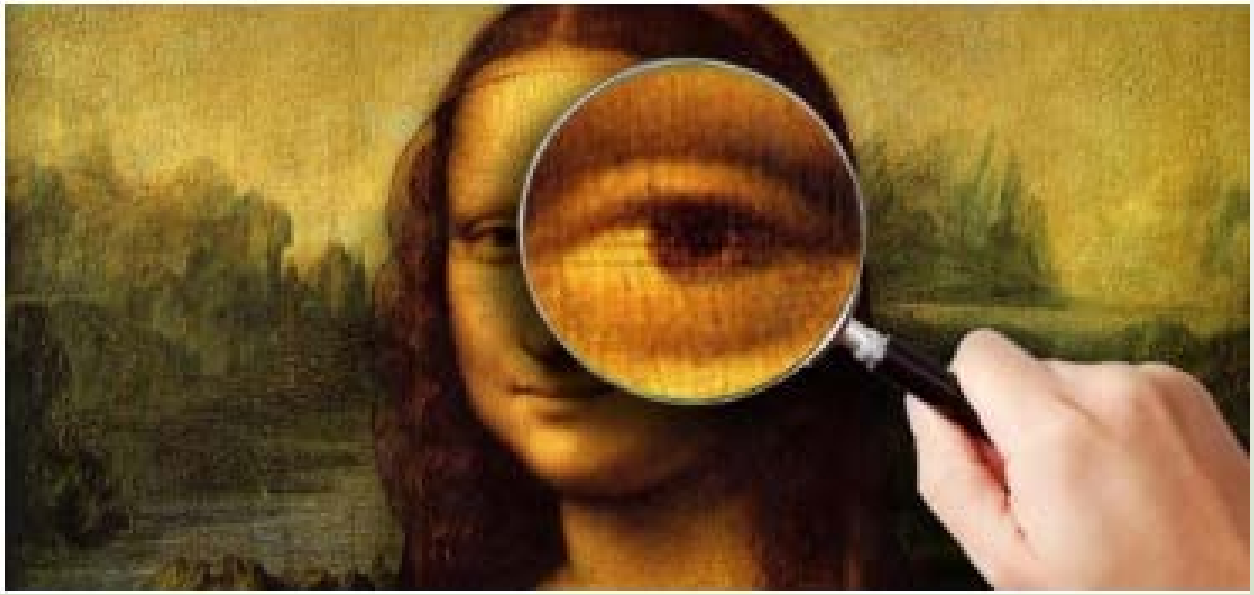
- 1）文本
- 2）图像
  - 图像隐写分析研究成为信息安全研究的一个重要方面
- 3）音频
- 4）视频
- 5）二进制文件

# 数字隐写对抗技术

- 一、隐写分析概述
- 二、专用隐写分析算法
- 三、通用隐写分析算法
- 四、隐写分析性能评估

# 专用隐写分析算法

- 近年来，隐写分析研究取得较大的进展，分析者利用隐写产生的统计特性不对称、直方图异常、调色板异常等现象，进行了成功的隐写检测或嵌入率估计



# 专用隐写分析算法

- 定义：针对某一种或者某一类隐写的有效分析方法

- 工作原理：在已知隐写算法的前提下，通过分析与测试隐写性质，得到专门用于识别该隐写的特征，并构造识别方法
  - 一般基于假设检验等统计推断法构造识别方法，需要选定阈值
  - 包括空域隐写分析和变换域隐写分析

# 专用隐写分析算法

## 空域隐写分析算法

- 空域隐写分析的攻击对象主要是空域LSB隐写术，包括S-Tools、Stash、BPCS等，是隐写分析研究初期非常活跃的部分

- 较多围绕颜色值对进行研究，研究方法经历了从简单分析隐藏图像颜色值对，到采用较复杂的实验手段，如再次嵌入秘密消息、归类、划分集合等，来获得颜色值对变化量的过程。总体上来说，适用性与实用性较低

# 专用隐写分析算法

## 变换域隐写分析算法

- 变换域隐写分析的攻击对象主要是DCT域隐写术，包括JSteg、F3、F4 、F5等，是隐写分析研究现时活跃的部分

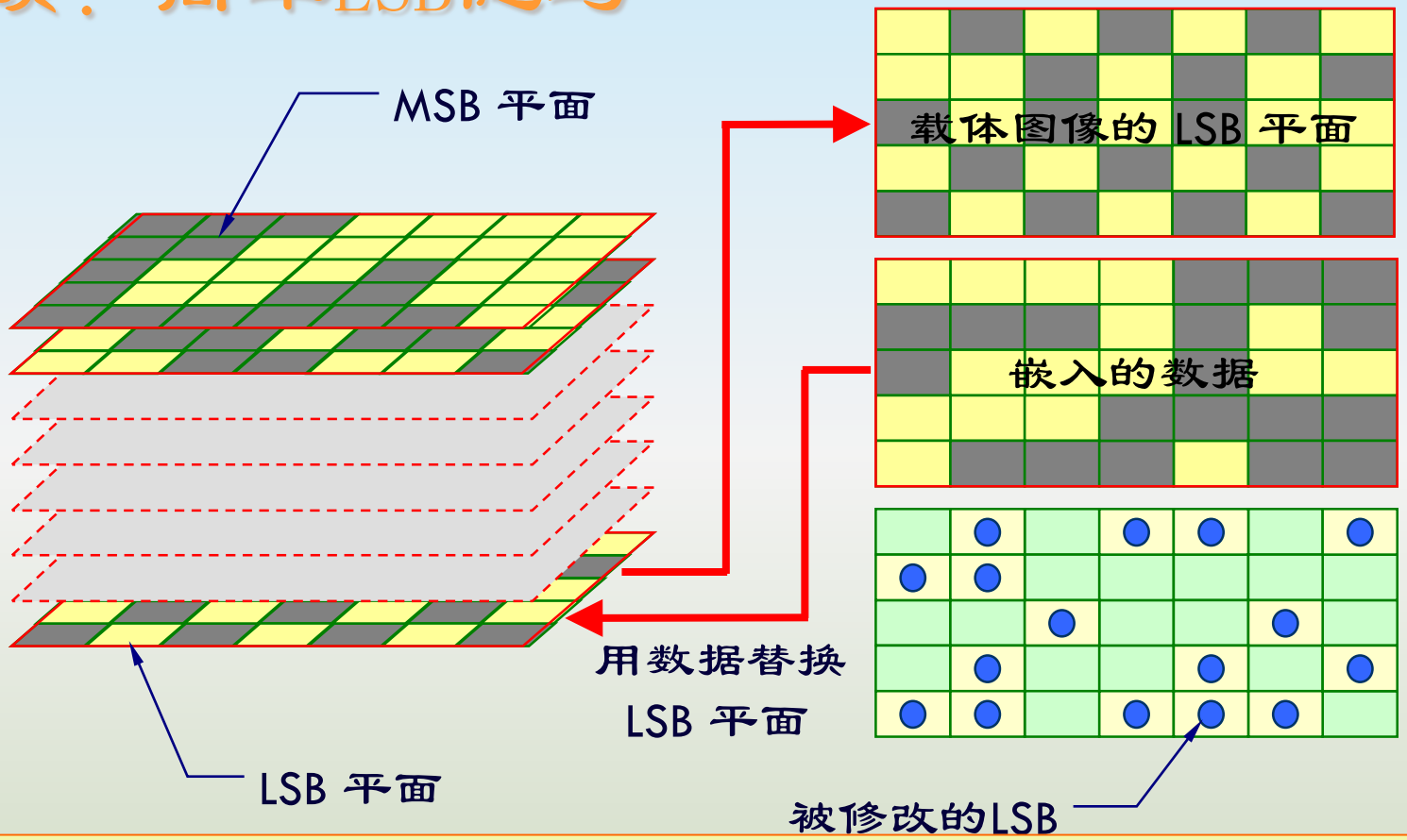
- DCT域隐写分析主要研究DCT系数的统计特性，及其对空域像素的影响，包括对载体图像DCT系数的估计，及空域像素块不连续性的计算。研究方法经历了从简单的一阶统计分析，到采用较复杂的实验手段，来获得相关变化量的过程

# 专用隐写分析算法

- 1. 卡方隐写分析算法
- 2. RS隐写分析算法

# 专用隐写分析算法

## 回顾：简单LSB隐写




方法简单，视觉失真小，数据容量大。

100%像素承载数据时，平均改变一半LSB，PSNR=51.1dB

# 专用隐写分析算法

## 回顾：简单LSB隐写

| 原灰度值 | | 嵌入 | | 嵌入后的灰度值 | | 提取 |
|---|---|---|---|---|---|---|
| 151 | 10010111 | 0 | 改为0 | 10010110 | 150 | 0 |
| 151 | 10010111 | 1 | 不变 | 10010111 | 151 | 1 |
| 152 | 10011000 | 0 | 不变 | 10011000 | 152 | 0 |
| 152 | 10011000 | 1 | 改为1 | 10011001 | 153 | 1 |
| 55 | 00110111 | 0 | 改为0 | 00110110 | 54 | 0 |
| 55 | 00110111 | 1 | 不变 | 00110111 | 55 | 1 |
| 54 | 00110110 | 0 | 不变 | 00110110 | 54 | 0 |
| 54 | 00110110 | 1 | 改为1 | 00110111 | 55 | 1 |

$$0 \leftrightarrow 1 \quad 2 \leftrightarrow 3 \quad \cdots \quad 2i \leftrightarrow 2i+1 \quad \cdots \quad 254 \leftrightarrow 255$$

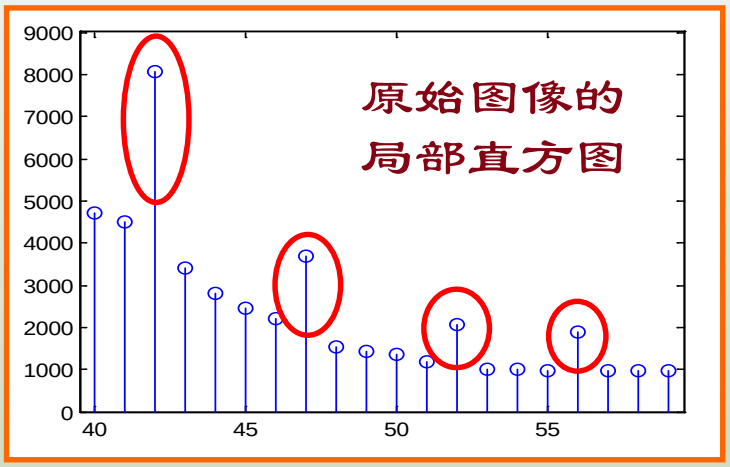
# 典型的空域隐写方法

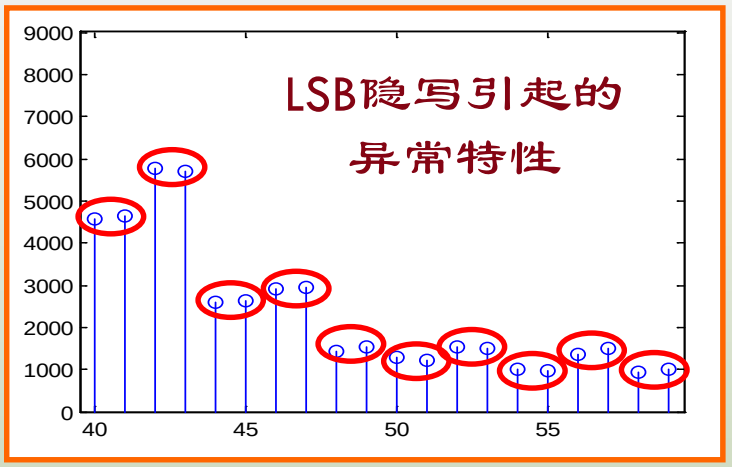
## LSB嵌入引起统计不对称性

- LSB嵌入对像素灰度的修改如下，记为$F^{(+)}$：

$$0 \leftrightarrow 1 \quad 2 \leftrightarrow 3 \quad \cdots \quad 2i \leftrightarrow 2i+1 \quad \cdots \quad 254 \leftrightarrow 255$$

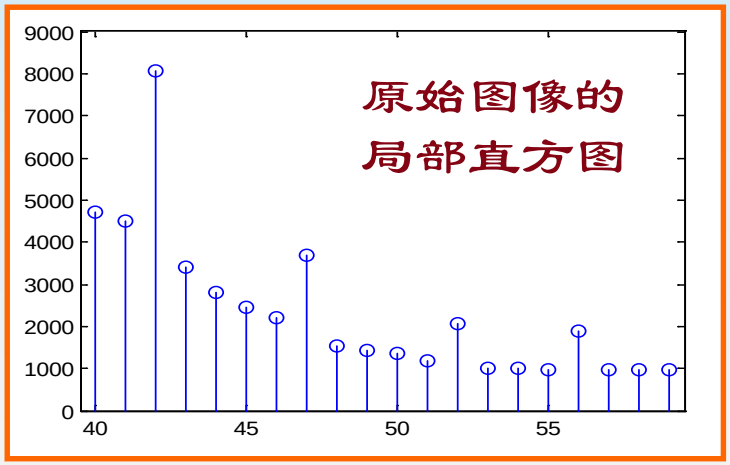
- $F^{(+)}$操作使灰度为$2i$和$2i+1$的像素数目趋于接近

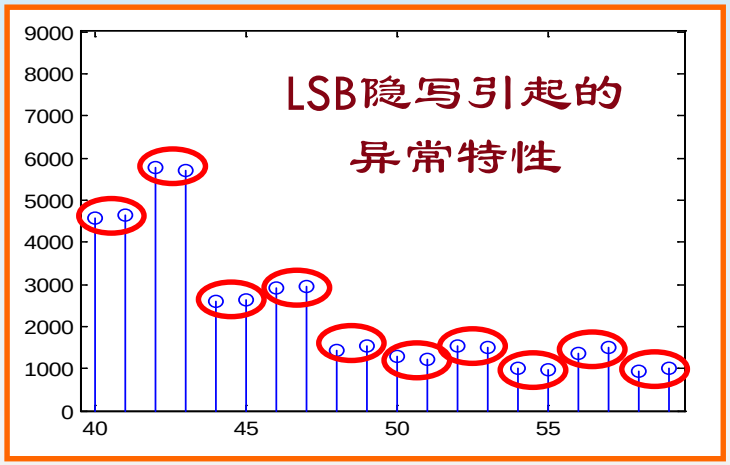


# 卡方隐写分析

## “值对”现象




图像像素灰度值转换的二进制比特序列如果看作是0和1出现概率相同的随机序列，利用LSB替换算法最终会导致最低有效位为0和1的个数趋于相等，这会使像素点灰度值前七位相同而最低有效位分别为0和为1的两个灰度值构成一个“值对”，每个“值对”中的灰度值出现的概率会趋于相同

# 卡方隐写分析

- 卡方隐写分析假设：
  - 1 嵌入的0，1比特概率各为50%；
  - 2 改动的规则：2i<->2i+1，不会有2i<->2i-1；
  - 3 如果秘密信息完全替代了最低位平面，则灰度值为2i和2i+1的像素点数量会比较接近，也就是产生了“值对”现象；
  - 4 如果图像未经隐写，则灰度值为2i和2i+1的像素点数量会相差较远。

# 卡方隐写分析

设 $N'_{2i}=(N_{2i}+N_{2i+1})/2$，因为LSB隐写时，灰度总在2i和2i+1之间转换，不会改变 $N_{2i}$ 与 $N_{2i+1}$ 的总量，所以可构造

$$\chi^2_{k-1}=\sum_{i=1}^{k}\frac{(N_{2i}-N'_{2i})^2}{N'_{2i}}$$

服从卡方分布，卡方值越小表示含有隐密信息的可能性越大

# 卡方隐写分析

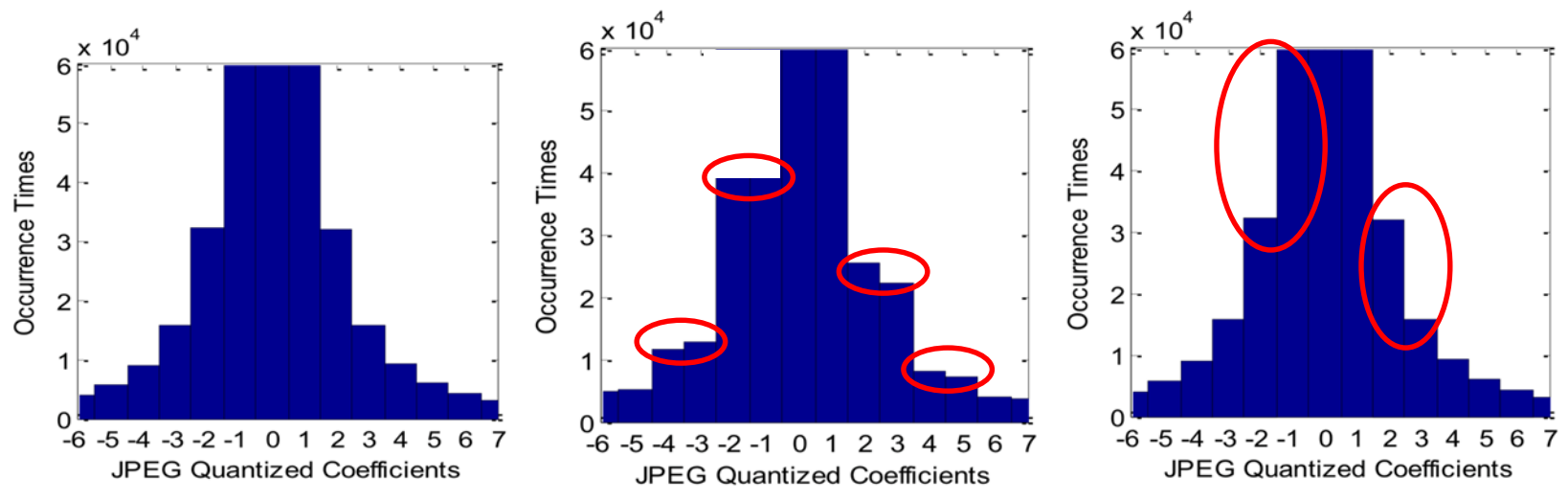
- 由此可总结出卡方隐写分析方法的特点：
  - 1 原理简单清晰，易于实现，检测效果较好；
  - 2 只能针对加密后密文0和1的概率各为50%的情况；
  - 3 只能针对特定的LSB替换隐写方法；
  - 4 只能针对高容量嵌入的情况；
  - 5 只能针对特定载体图像，即原始图像中灰度值为2i和2i+1的像素点数量相差较远的；
  - 6 对部分伪随机嵌入无效，因为不会引起值对现象。

# 思考

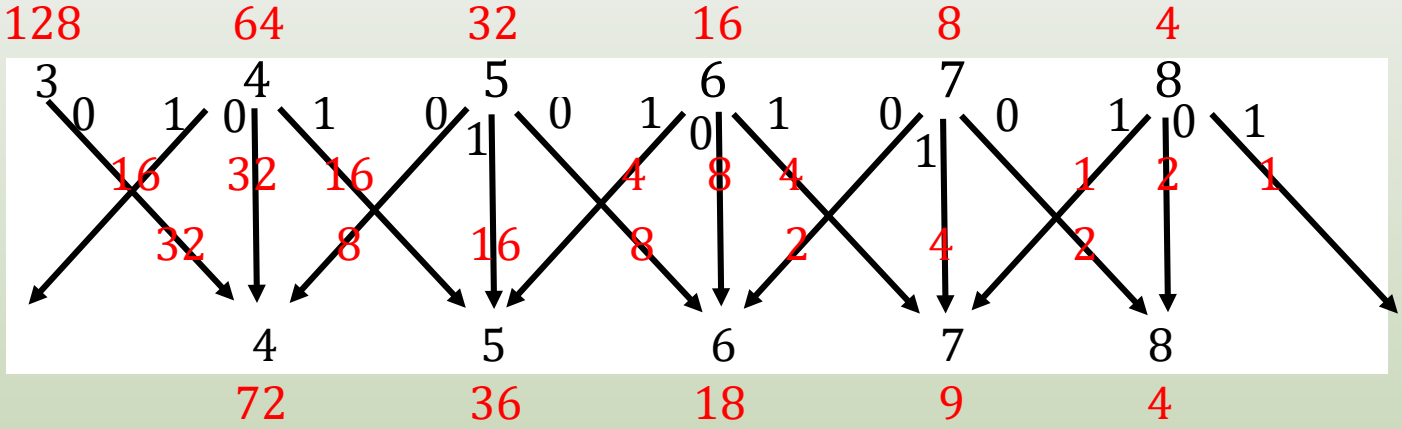
- 如何抗卡方隐写?

# 思考

## LSB改进为 LSBM




原直方图　　LSB隐写直方图　　LSBM隐写直方图

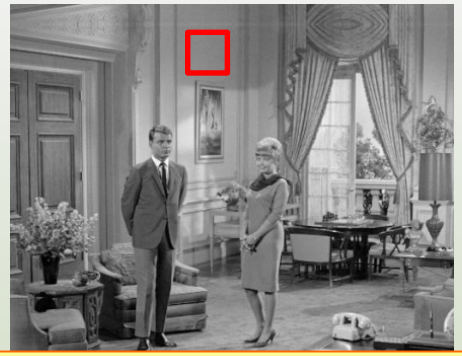


# RS隐写分析

- 提出者Fridrich
  - 纽约州立大学宾汉姆顿分校教授，水印鉴定、电力电气、数字影像法医鉴定等领域的专家;
  - 世界上最快的、使用最广泛的魔方解法—Fridrich法的发明人。

# RS隐写分析

绝大多数图像的相邻像素点之间具有较强的相关性，而秘密信息由于经过加密处理后，通常不具有相关性，所以当秘密信息被嵌入到载体图像数据的最低位后，像素灰度值之间的相关性会在一定程度上受到破坏



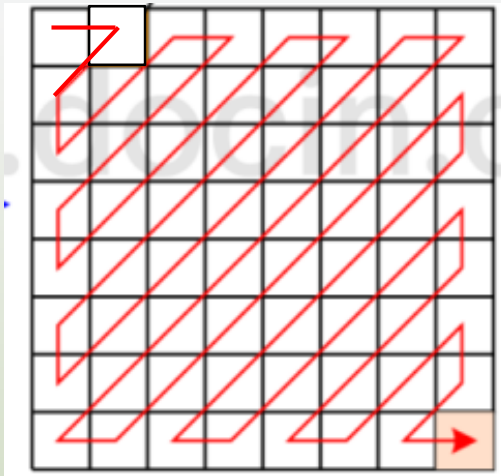
RS分析方法就是利用这种特性来检测数字媒体中是否含有秘密信息

# RS隐写分析

- RS隐写分析特点：
  - 适用于检测随机嵌入的LSB替换隐写，可以比较精确地估计隐写信息的长度，对彩色和灰度图像都适用
- RS隐写分析原理：
  - 基于隐写前后图像平滑度的变化来检测嵌入的秘密信息，利用翻转不对称性，考察空间相关性的偏离程度，也可以估计信息嵌入率

# RS隐写分析

- 算法原理：
  - 给定一个图像块，以“之”字型扫描排成一个像素向量 $\mathbf{G}=\{x_1 \quad x_2 \quad \cdots \quad x_n\}$



| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 5 | 6 | 14 | 15 | 27 | 28 |
| 2 | 4 | 7 | 13 | 16 | 26 | 29 | 42 |
| 3 | 8 | 12 | 17 | 25 | 30 | 41 | 43 |
| 9 | 11 | 18 | 24 | 31 | 40 | 44 | 53 |
| 10 | 19 | 23 | 32 | 39 | 45 | 52 | 54 |
| 20 | 22 | 33 | 38 | 46 | 51 | 55 | 60 |
| 21 | 34 | 37 | 47 | 50 | 56 | 59 | 61 |
| 35 | 36 | 48 | 49 | 57 | 58 | 62 | 63 |

# RS隐写分析

- 算法原理：
  - 用相邻像素灰度差值的绝对值总和来表示图像的空间相关性，即：

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n-1} |x_{i+1} - x_i|$$

  - f值越小说明图像块的混乱程度越小，也就是图像块的空间相关性越强

# RS隐写分析

- 算法原理：

  - $F^+$操作：

$$0 \leftrightarrow 1 \quad 2 \leftrightarrow 3 \quad \cdots \quad 2i \leftrightarrow 2i+1 \quad \cdots \quad 254 \leftrightarrow 255$$

  - $F^-$操作：

$$1 \leftrightarrow 2 \quad 3 \leftrightarrow 4 \quad \cdots \quad 2i-1 \leftrightarrow 2i \quad \cdots \quad 253 \leftrightarrow 254$$

LSB嵌入操作相当于对部分像素进行$F^+$翻转操作，但是不存在$F^-$翻转操作

# RS隐写分析

- **算法流程：**
  - 将待检图像分为很多大小相等的图像块，对每个小图像块随机抽取部分像素应用$F^+$或$F^-$操作，计算其混乱程度增加或减少块的比例
    - $R_M$：$F^+$操作，混乱程度增加的图像块比例
    - $S_M$：$F^+$操作，混乱程度减少的图像块比例
    - $R_{-M}$：$F^-$操作，混乱程度增加的图像块比例
    - $S_{-M}$：$F^-$操作，混乱程度减少的图像块比例

# RS隐写分析

- 算法流程：
  - 如果待检测图像没有经过LSB隐写，那么经过F+翻转和F-翻转，会等同增加图像块的混乱度，即：

$$R_M \approx R_{-M}, \quad S_M \approx S_{-M}, \quad \text{且} R_M > S_M, \quad R_{-M} > S_{-M}$$

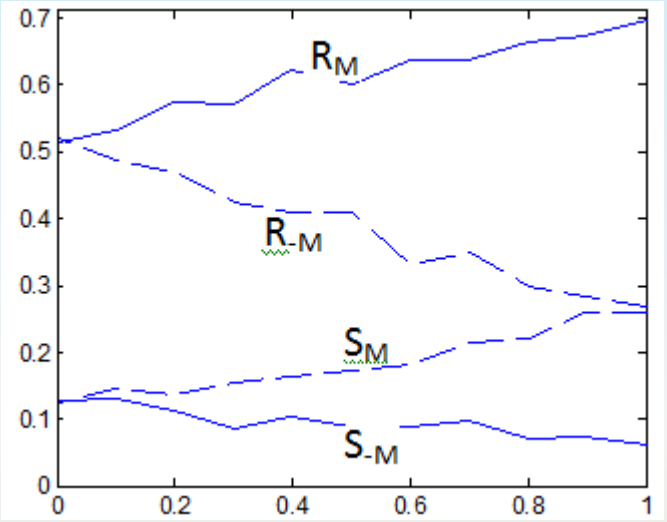
# RS隐写分析

- 算法流程：
  - 如果待检图像是经过LSB隐藏嵌入的，应用F-翻转对混乱度的增加要大于应用F+翻转，即：

$$R_M < R_{-M}, \quad S_M > S_{-M}$$

# RS隐写分析

- 对正常图像的像素块运用$F^{(+)}$和$F^{(-)}$两类翻转，检查起伏程度变化：
  - 经$F^{(+)}$和$F^{(-)}$两类翻转后起伏程度有同等增加；
  - 经$F^{(+)}$和$F^{(-)}$两类翻转后起伏程度有同等减少；
  - 增加的起伏程度明显大于减少的起伏程度。




正常图片嵌入率

- $R_M$：$F^+$操作，混乱程度增加的图像块比例
- $R_{-M}$：$F^-$操作，混乱程度增加的图像块比例
- $S_M$：$F^+$操作，混乱程度减少的图像块比例
- $S_{-M}$：$F^-$操作，混乱程度减少的图像块比例

$$R_M \approx R_{-M}, \quad S_M \approx S_{-M}, \quad \text{且} \; R_M > S_M, \quad R_{-M} > S_{-M}$$

# RS隐写分析

- **对隐写图像的像素块运用$F^{(+)}$和$F^{(-)}$两类翻转，检查起伏程度变化：**
  - 经$F^{(+)}$翻转，大量像素点两次翻转又回到了原始值，起伏程度增加的少，减少的多；
  - 经$F^{(-)}$翻转，也就是先后经$F^{(+)}$和$F^{(-)}$两类翻转后大量像素点与原始值相差更大，起伏程度增加的多，减少的少；　2i+1<--> 2i <-->2i-1
    - $R_M$：$F^+$操作，混乱程度增加的图像块比例
    - $R_{-M}$：$F^-$操作，混乱程度增加的图像块比例
    - $S_M$：$F^+$操作，混乱程度减少的图像块比例
    - $S_{-M}$：$F^-$操作，混乱程度减少的图像块比例

$$R_M < R_{-M}, \quad S_M > S_{-M}$$

# RS隐写分析示例

- 原图像矩阵

163 163 162 163

163 162 164 161

163 164 161 161

163 162 159 159

f=0+0+0+1+0+1+1+0+1+1+1+0+0+2+0=8

- 隐写矩阵 0110101100010111

162 163 163 162

163 162 165 161

162 164 160 161

162 163 159 159

f=1+0+1+0+1+1+3+1+2+1+3+1+0+2+0=17

- $F^+$操作

162 163 163 162

163 163 164 161

163 164 160 161

162 163 159 158

f=1+0+0+0+0+1+2+0+2+1+3+1+0+2+1=14

- $F^-$操作

164 163 161 164

163 161 164 161

163 164 162 161

164 161 159 160

f=1+0+0+2+0+3+0+0+0+3+1+1+0+2+1=14

正常图像块执行$F^+$和$F^-$操作，混乱程度都会增加，增加的幅度近似

$R_M \approx R_{-M}$，$S_M \approx S_{-M}$，且 $R_M > S_M$，$R_{-M} > S_{-M}$

# RS隐写分析示例

- 原图像矩阵

163 163 162 163

163 162 164 161

163 164 161 161

163 162 159 159

f=0+0+0+1+0+1+1+0+1+1+1+0+0+2+0=8

- 隐写矩阵 0110101100010111

162 163 163 162

163 162 165 161

162 164 160 161

162 163 159 159

f=1+0+1+0+1+1+3+1+2+1+3+1+0+2+0=17

- $F^+$操作

163 163 162 163

163 163 165 161

162 164 161 161

163 162 159 158

f=0+0+1+1+1+1+2+1+1+1+1+0+0+2+1=13

隐写图像块执行$F^+$操作，大量像素点回到了原始值，混乱程度减少

$$R_M < R_{-M}, \quad S_M > S_{-M}$$

- $F^-$操作

161 163 164 161

163 161 165 161

162 164 159 161

161 164 159 160

f=2+0+1+1+3+3+4+1+3+3+5+2+0+2+1=31

隐写图像块执行F-操作，大量像素点与原始值相差更大，混乱程度进一步增加

# 练习

- 原图像矩阵

158 157 157 157

156 158 156 158

157 157 157 159

158 157 159 158

LSB替换嵌入信息为1010010110111001，分别写出隐写矩阵、对隐写矩阵进行$F^{+}$和$F^{-}$操作后的矩阵，并分别计算出原矩阵和这3个不同矩阵的f值。

# JPEG图像隐写分析

- JPEG图像是最重要的隐写载体之一
- 代表性隐写方法：Jsteg、F3、F4、F5
- 在分析中可利用的主要特征：
  - 量化DCT系数的最低有效位
  - 图片分块现象
  - DCT系数直方图
  - 量化表：量化步长的异常，从而暴露隐蔽信息存在

# 专用隐写分析算法

- 基本思想：
  - 利用统计不对称、直方图异常、调色板异常等分析方法使LSB替换、JSteg等典型隐写方法不再安全
- 面临的问题：
  - 新型抗隐写分析的隐写算法产生
  - 新的隐写算法、技术、工具也不断涌出

# 数字隐写对抗技术

- 一、隐写分析概述
- 二、专用隐写分析算法
- 三、通用隐写分析算法
- 四、隐写分析性能评估

# 通用隐写分析算法

- 据SARC公司统计，目前至少有800种隐写算法。通用隐写分析算法在没有任何先验知识的条件下，判断图像中是否隐藏有秘密信息，成为目前隐写分析研究的主流方向

# 互联网上常见的图像隐写软件

| 序号 | 工具 | 作者及网址 | 主要方法 | 图像格式 |
|---|---|---|---|---|
| 1 | BMP Secrets | http://www.pworlds.com | 空域替换方法 | JPEG, GIF, BMP等 |
| 3 | DCT-Steg (DCT-Jpeg) | Stefan Katzenbeisser http://www.dbai.tuwien.ac.at/staff/katzenb/stego/robust-st.html | 修改DCT系数 | JPEG |
| 4 | EzStego | Romana Machado http://www.stego.com | LSB方法 | GIF |
| 5 | F5 v0.9 | Andreas Westfeld (Dresden,Germany) http://www.inf.tu-dresden.de/~aw4 | 修改量化后的DCT系数 | BMP, GIF, JPEG |
| 6 | Hide and Seek | Colin Moroney | LSB方法 | BMP |
| 7 | JP Hide and Seek | Allan Latham, http://linux01.gwdg.de/~alatham/- stego.html | 修改量化后的DCT系数 | JPEG |
| 8 | JPHSWin | Allan Latham http://linux01.gwdg.de/~alatham/stego.html | 修改量化后的DCT系数 | JPEG |
| 9 | JSteg Shell | John Korejwa http://www.tiac.net/users/korejwa/jsteg.htm | 修改量化后的DCT系数 | 输出JPEG |
| 10 | JSteg-Jpeg | Derek Upham http://linkbeat.com/files/ | 修改量化后的DCT系数 | 输出JPEG |
| 11 | Mandelsteg | Henry Hastur | LSB方法 | GIF |
| 12 | OutGuess | Niels Provos http://www.outguess.org/ | 修改量化后的DCT系数 | JPEG, PNM |
| 13 | S-Tools v4.0 (stools) | Andy Brown, ftp://ftp.demon.net/pub/-mirrors/crypto/idea/code/s-tools4.zip | LSB方法 | BMP, GIF |
| 14 | White Noise Storm | Ray (Arsen) Arachelian | 扩频+LSB方法 | PCX |
| 15 | Steganos Security Suite III | DEMCOM, Fabian Hansmann http://www.steganography.com | LSB方法 | BMP, DIB |

# 公开隐写工具或算法

| | | | | |
|---|---|---|---|---|
| 1 | Blindside | BMP | 共享 | www.blindside.co.uk |
| 2 | BMP Secrets | BMP | 共享 | www.pworlds.com/products/i_secrets.html |
| 3 | Camouflage | Windows格式文件 | 共享 | www.camouflagesoftware.co.uk |
| 4 | StegMark | JPG,GIF,TIF, PNG,MIDI,WAV, AVI,MPEG | 商业 | www.datamark-tech.com |
| 5 | dc-Steganograph | PCX | 共享 | members.tripod.com/~Nikola_Injac /stegano/ |
| 6 | Digital Picture Envelope | BMP | 共享 | www.know.comp.kyutech.ac.jp/BPCSe /Dpenve/DPENVe-home.html |
| 7 | Empty Pic | GIF | 共享 | www.crtelco.com/~robertw/ |
| 8 | EZStego | GIF | 共享 | www.stego.com |
| 9 | F5 | JPG | 共享 | wwwrn.inf.tu-dresden.de/~westfeld /f5.html |

# 公开隐写工具或算法

| | | | | |
|---|---|---|---|---|
| 10 | FFEncode | 文本文件 | 共享 | www.rugeley.demon.co.uk /security/ffencode.zip |
| 11 | Gif-It-Up | GIF | 共享 | crypto.radiusnet.net/archive /steganography/gif-it-up/ |
| 12 | Gifshuffle | GIF | 共享 | www.darkside.com.au/gifshuffle/ |
| 13 | Hide and Seek | BMP | 共享 | ftp://ftp.hacktic.nl/pub/crypto/ incoming/hideseek95.zip |
| 14 | Hide In Picture | BMP | 共享 | www.brasil.terravista.pt/Jenipabu /2571/e_hip.htm |
| 15 | Invisible Secrets | JPG, BMP, PNG | 共享 | www.innovatools.com/software /isecrets/ |
| 16 | Invisible Secrets 3 | JPG, PNG, BMP, HTML, WAV | 付费 | www.neobytesolutions.com/invsecr /index.htm |
| 17 | JP Hide and Seek | JPG | 共享 | linux01.gwdg.de/~alatham/stego.html |
| 18 | JSteg | JPG | 共享 | zooid.org/~paul/crypto/jsteg/ |

# 公开隐写工具或算法

| | | | | |
|---|---|---|---|---|
| 19 | MP3Stego | MP3 | 共享 | www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/ |
| 20 | Steghide | BMP, WAV | 共享 | www.crosswinds.net/~shetzl/steghide/index.html |
| 21 | StegoWav | WAV | 共享 | www.geocities.com/SiliconValley/9210/stegowav.zip |
| 22 | Stella | GIF,BMP,JPG | 共享 | www.stella-steganography.de/ |
| 23 | S-Tools | BMP,GIF,WAV,软盘 | 共享 | ftp://ftp.ntua.gr/pub/crypt/mirrors/idea.sec.dsi.unimi.it/code/stools4.zip |
| 24 | TextHide | 文本文件 | 商业 | www.texthide.com |
| 25 | wbStego | BMP,TXT,HTML,XML,PDF | 付费 | wbstego.wbailer.com |
| 26 | 轻松图片加密 | BMP | 共享 | www.onlinedown.net/soft/30471.htm |
| 27 | 渗透 3.0 | FLASH SWF, BMP, JPG | 共享 | www.onlinedown.net/soft/4069.htm |
| 28 | BMP文件隐藏加密器 | BMP | 共享 | www.onlinedown.net/soft/26335.htm |

# 通用隐写分析算法

- 通用隐写分析算法是更适用的隐写分析技术，它在没有任何先验知识的条件下，判断图像中是否隐藏秘密信息
  - 通用：不针对某一种隐秘工具或者某一类隐秘方法的“盲”分析
  - 手段：不限定具体隐写算法，寻找具有普遍适用性的、高阶的、更加鲁棒的统计特征，以便适用更广的隐写算法

# 通用隐写分析算法

- 基本原理：
  - 对原始载体或隐写图像进行特征提取与选择；
  - 利用人工智能或模式识别的方法设计分类器并对分类器进行训练；
  - 通过样本比较和数据分析从特征空间的意义上区分原始载体和隐写图像。

# 通用隐写分析算法

- 特点：
  - 隐秘设计者很难在设计隐秘算法时兼顾不同的图像统计特征和模型；
  - 只要使用足够精确的统计模型，特征提取合理，就可能实现对不同隐秘技术的成功攻击。

# 通用隐写分析算法

## 早期通用隐写分析－Memon

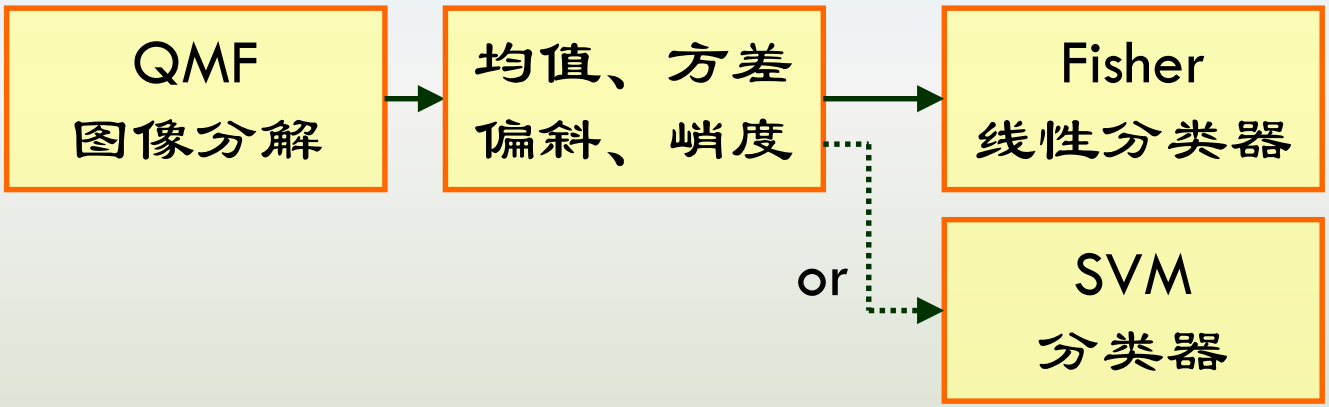
- 通用隐写分析最早在2001年由Memon提出；
- 隐写嵌入必定影响图像质量，使用某种图像质量测度可导出统计检验量用于分析；
- 利用方差分析选择适当的质量测度，然后利用多变量回归分析可实现通用隐写分析。
  - 质量测度如：平均绝对误差、均方误差、相关测度、图像保真度、谱距离、归一化均方HVS误差等。

I. Avcibas, N. Memon, and B Sankur, Steganalysis Using Image Quality Metrics, IEEE Trans Image Processing, 12(2), 2003: 221-229.
Earlier version appeared in Proc. SPIE, 4314, 2001: 523-531

# 通用隐写分析算法

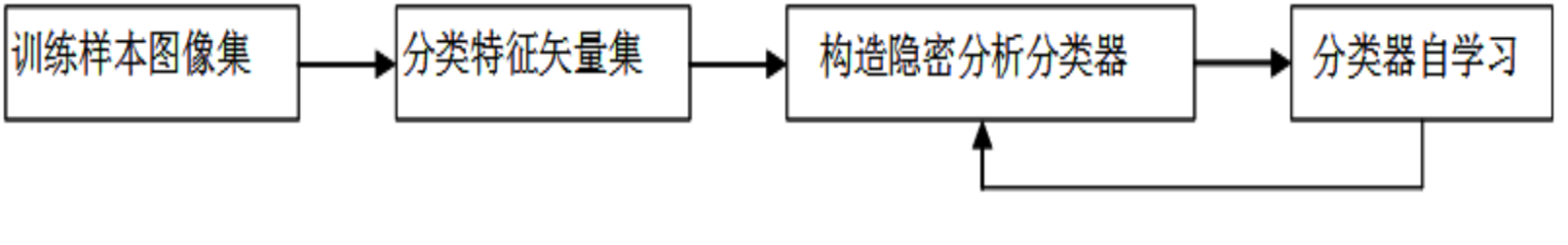
## 另一种早期通用隐写分析－Farid

- 使用类似小波的“正交镜像滤波器”（QMF）分解图像，建立自然图像的4阶统计模型

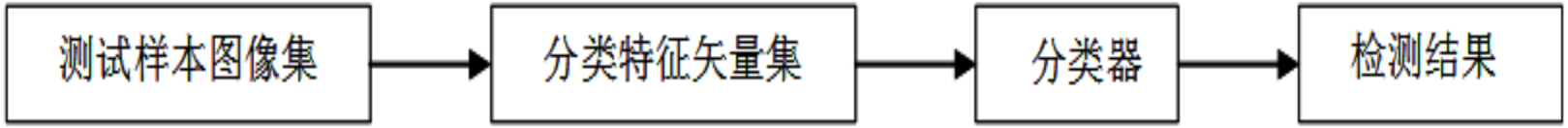



H. Farid, Detecting Hidden Messages Using Higher-Order Statistical Models, in Proc. of the 5th ICIP, vol. 2, 2002: 905-908.

# 通用隐写分析算法




（a）学习过程




（b）判决过程图

5-1 通用隐密分析系统模型

通用隐写分析方法的关键在于寻找对秘密信息嵌入敏感的统计量，设计合适的判别方案区分原始载体图像和隐写图像

# 通用隐写分析算法

- 核心问题：
  - 如何提取具有代表性、分类性能优良的特征？
  - 如何设计准确率高、复杂度偏低的分类器？

# 通用隐写分析算法

- 分类器
  - 一种机器学习的计算机程序，其设计目标是在通过学习后，可以自动对给定的数据进行分类
- 应用领域
  - 识别认证、数据挖掘、专家系统、模式识别、搜索引擎等
- 种类
  - 针对模型的不同，目前有多种分支，包括：卷积神经网络CNN、循环神经网络RNN、Bayes网络分类器、决策树算法、聚类算法、SVM算法等

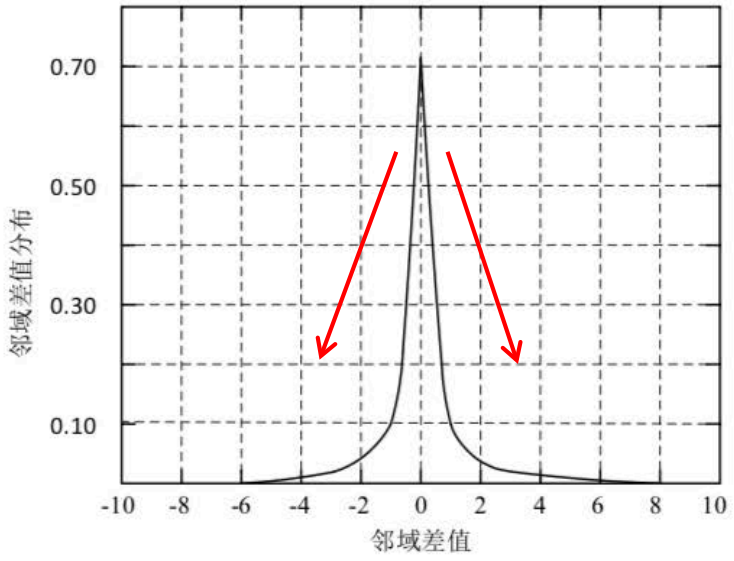
# 通用隐写分析算法

- 示例1：Markov图像隐写分析算法
  - Shi Yunqing
    - **施云庆，新泽西理工大学教授**，IEEE Fellow，上海交通大学本科，硕士毕业
  - 方法基本原理：
    - 将Markov矩阵作为图像的一个二阶统计特征，使用水平、垂直、对角和反对角四个方向的差分Markov方法来放大隐写造成的DCT系数的改变
    - 面向JPEG图像

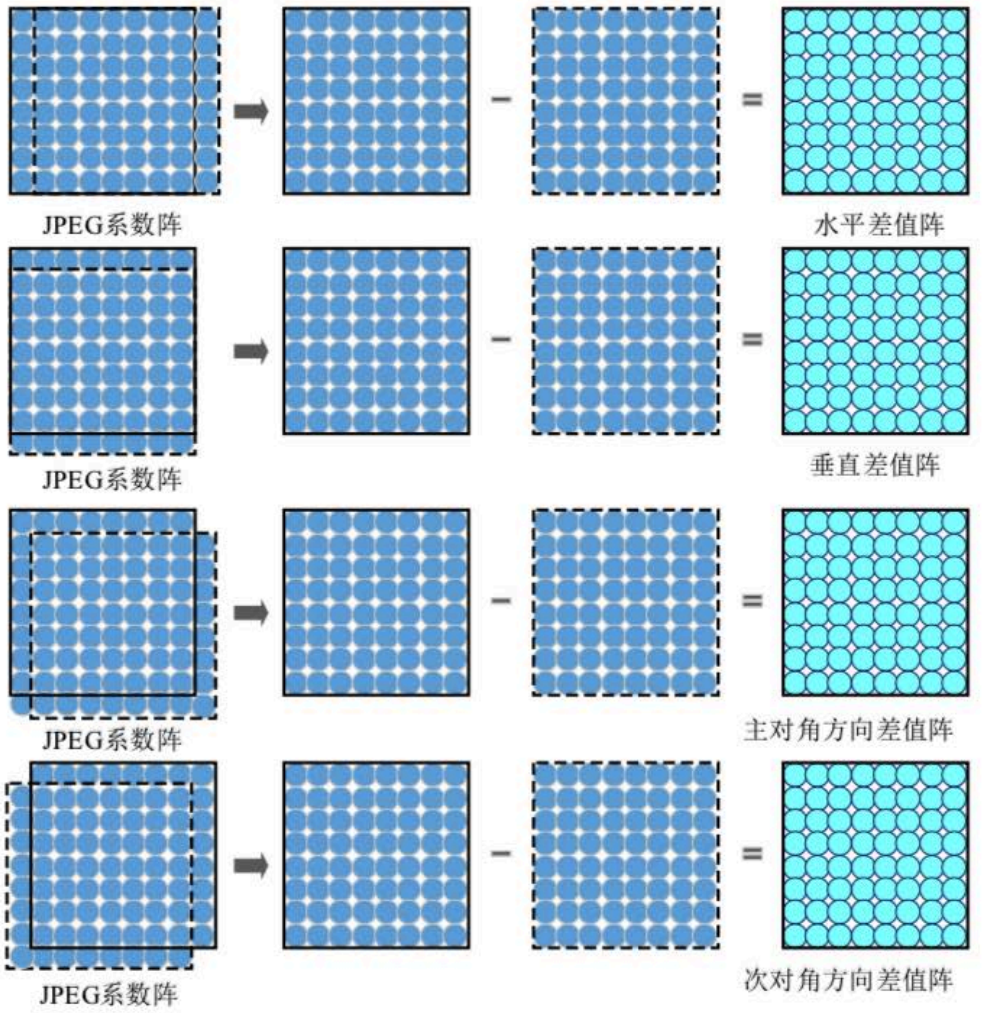
# 通用隐写分析算法

- Markov特征分析：相邻系数差值矩阵的相关性

$$F_h(u,v)= F(u,v) - F(u+1,v)$$
$$F_v(u,v)= F(u,v) - F(u,v+1)$$
$$F_d(u,v)= F(u,v) - F(u+1,v+1)$$
$$F_m(u,v)= F(u+1,v) - F(u,v+1)$$




相邻系数差值的分布

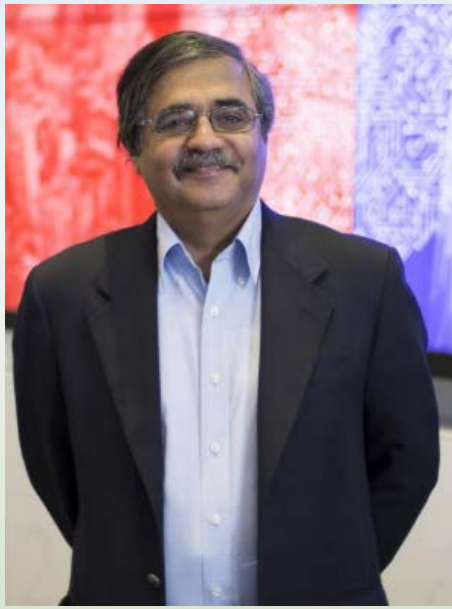


# 通用隐写分析算法

- 特点：引入阈值T，减少了转移概率矩阵的计算复杂度，是后续Markov模型减少计算量一直效仿的方法
  - 统计相邻数据之间的变化以及变化的转移关系
  - 采用SVM分类器进行分类识别
  - 被用在很多信号特征提取的工作中

# 通用隐写分析算法

- 示例2：IQM(Image Quality Metrics)方法
  - Nasir Memon
    - **纽约大学教授**，IEEE/SPIE Fellow，研究方向包括图像压缩、媒体安全和取证，曾担任IEEE TIFS主编
  - 方法基本原理：
    - 采用方差分析方法选取可用于区分载体图像和隐写图像质量的特征，然后根据选取的图像质量特征采用多元回归对图像进行分类



# 通用隐写分析算法

- 示例2：IQM(Image Quality Metrics)方法

- IQM可用来度量原始图像和受损图像间的相互关系，受损图像包括隐藏信息后的图像、加水印后的图像、按不同压缩方法和压缩比，压缩处理和加噪处理后的图像等；
- 在原始图像中嵌入隐藏信息后，图像质量会有所下降，利用合适的IQM可以捕捉这一质量下降的规律；
- 基本思想是相对一个公共的参考图像，载体图像和隐写图像之间存在一致的距离度量。由于高斯滤波器针对大多数隐写算法都能取得一定效果，因此使用高斯滤波产生参考图像。

# 数字隐写对抗技术

- 一、隐写分析概述
- 二、专用隐写分析算法
- 三、通用隐写分析算法
- 四、隐写分析性能评估

# 隐写分析性能评估

- 1）准确性
- 2）适用性
- 3）实用性
- 4）复杂度

# 隐写分析性能评估

## 隐写分析的准确性

- 隐写分析的目的是判断载体是否含有秘密信息;

- 因此可以将隐藏分析问题看作一个二分类问题，即判断待检测集合{I}中的任意载体是属于原始载体集合{C}还是属于载密载体集合{S}。

# 隐写分析性能评估

## 隐写分析的准确性

- 隐写分析算法判断某个待检载体（原始载体或者载密载体），会产生以下四种不同的结果:

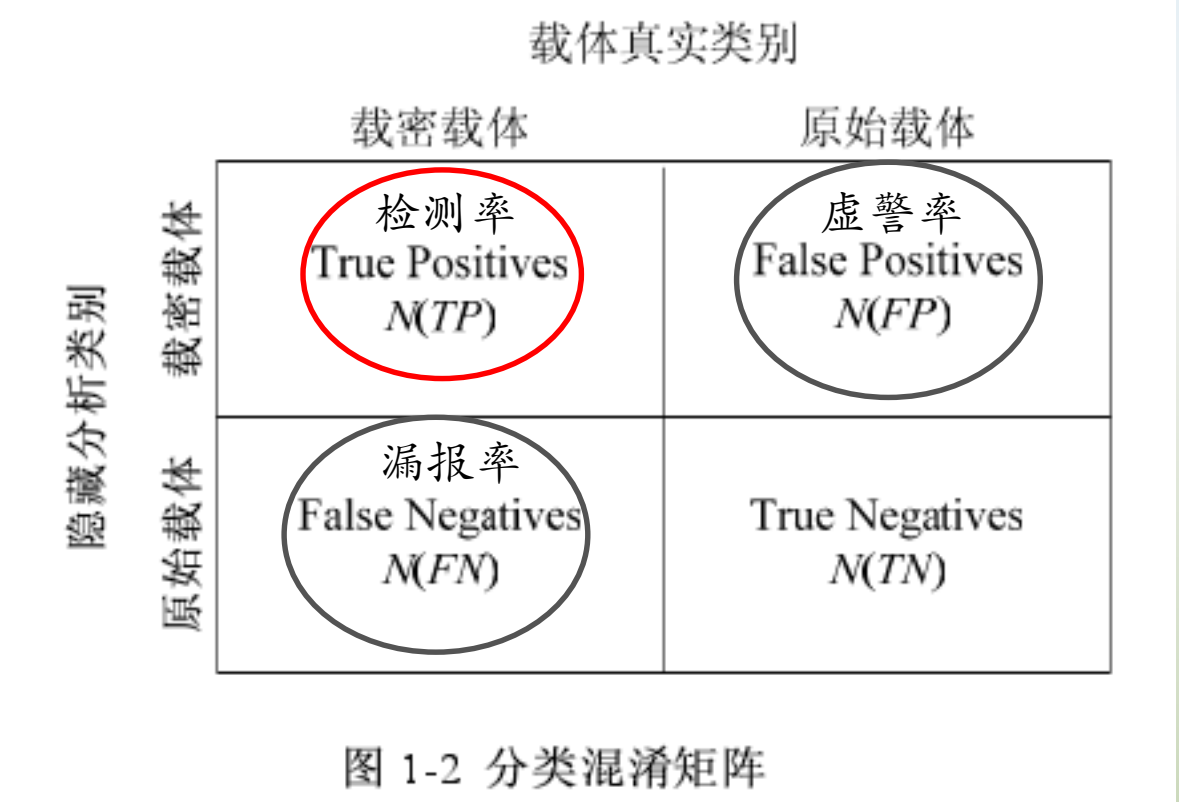



图 1-2 分类混淆矩阵

# 隐写分析性能评估

## 评定指标

- **检测率：** 表示算法能在隐写样本中正确地检测出隐秘消息的概率

$$\alpha = P(\text{肯定有隐秘消息} \mid \text{隐秘对象})$$

- **虚警率：** 表示算法在非隐写样本中错误地检测出隐秘消息的概率

$$\beta = P(\text{肯定有隐秘消息} \mid \text{非隐秘对象})$$

- **漏报率：** 表示算法在隐写样本中没有检测出隐秘消息的概率

$$\gamma = P(\text{没有隐秘消息} \mid \text{隐秘对象})$$

隐写分析的准确性是指在尽可能降低虚警和漏报的情况下，取得尽量高的检测率，且优先降低漏报率。

# 隐写分析性能评估

## 受试者操作特性ROC曲线

- 横坐标：虚警率
- 纵坐标：检测率

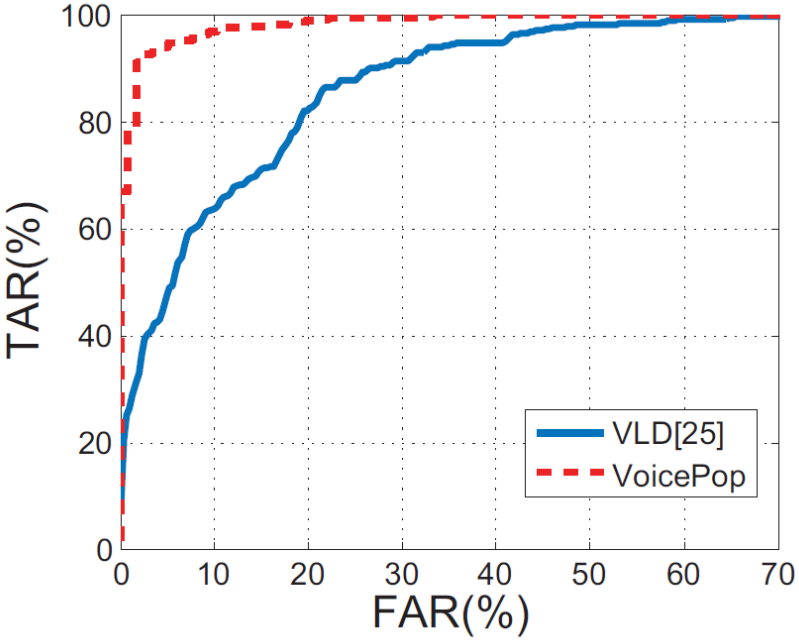



Fig. 9. Overall ROC curves.

# 隐写分析性能评估

## 受试者操作特性ROC曲线

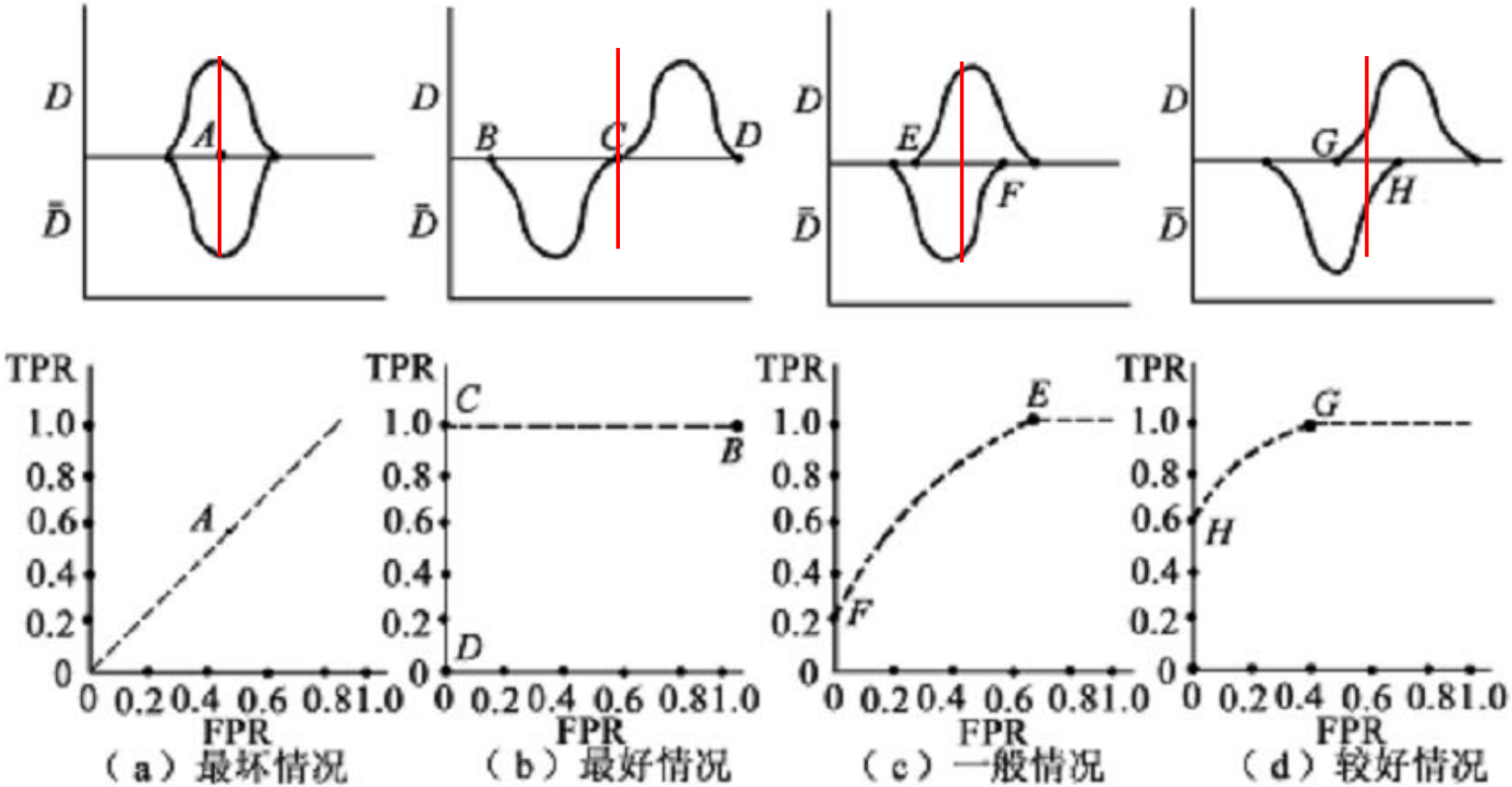



图 9.2　ROC 曲线的含义

# 隐写分析性能评估

## 隐写分析的适用性

- 检测算法对于不同隐写方法的有效性（可以理解为隐写分析算法的通用性程度）；

- 在研究和评价视频隐写分析时，可以将其看作从针对少数典型隐写方法，逐渐往更广泛的隐写方法集合的动态归纳和演绎过程。

# 隐写分析性能评估

## 隐写分析的实用性

- 检测算法可以实际应用的程度，衡量指标包括：
  - 现实条件是否允许
  - 检测结果是否稳定
  - 自动化程度有多高
  - 实时性是否达到

# 隐写分析性能评估

## 隐写分析的复杂度

- 针对检测算法本身而言的，由检测算法实现所需要的资源开销、软硬件条件等来衡量；

- 隐写分析算法的复杂度在一定意义上影响其应用的实用性。

# 隐写分析性能评估

- 隐写分析的几个评价指标之间存在矛盾折中的关系，例如：
  - 目前的专用隐写分析算法从检测的准确性角度明显高于通用分析算法。也就是说，适用性越强，对应的准确性一般会越低，反之亦然。
  - 采用高阶或更多统计特征进行检测的分析算法，复杂度显然更高，同时会更有效检测出隐秘消息，增加检测的准确性，但实用性有变差的趋势。

**在比较不同的隐写分析算法性能时，首先要综合几个方面的指标，其次要把它们放在具体的应用条件下来进行评估。**

# 隐写分析的困难所在

- 隐写不会改变视听觉效果
- 隐写分析需使文件属性保持不变
- 隐写的统计特性异常极其微小
- 隐写方法变化多端
- 含密媒体与大量正常媒体共存
- 高分辨率图像激增
- 隐写可用载体多样化

# 小结

## 隐写分析技术的发展

见招拆招

无招胜有招

# 小结

## 现代信息隐藏面临的困难

- 信息隐藏还没有找到自己的理论依据，尚未形成自己的理论体系。

  缺乏Shannon理论这样的密码学理论基础，缺乏对人类感知模型的充分理解，缺乏对信息隐藏方案的有效度量方法等。

- 实际应用道路上尚存在许多技术性问题需要解决。

  如何使用数字水印技术来实现各种数字媒体的产权保护和管理、消息的认证与用户的鉴权？如何正确提取隐写信息，以验证隐写检测的正确性？等等。

- 水印验证体系的建立、法律的保护等因素在信息隐藏技术迈向实用化过程中尚缺少应用环境。

# 小结

- 问题
  - 隐写分析技术的评测标准以及理论构架尚待完善;
  - 更实用的视频隐写分析原型系统有待提出;
  - 将统计分析和归类判断的方法相结合, 实现全自动检测是构建实用检测系统的方向。
- 挑战
  - 网络流媒体编码传输技术的发展
  - 难以提高小嵌入率检测的准确率
  - 每种算法都有其不可避免的局限性
  - 深度生成技术带来新的挑战