

一种基于图像内容的最低有效位匹配隐写分析方法

熊钢* 平西建 张涛 孙兵

(解放军信息工程大学 郑州 450002)

摘要: 将隐藏信息检测与图像内容分析相结合是当前提高图像隐写分析性能的一个新方向。与基于图像整体内容的检测方法不同, 该文分析了最低有效位(Least Significant Bit LSB)匹配隐写对图像子区域统计特性的影响, 提出一种新的联合判决检测方法。首先依据图像内容复杂度将整体图像分割为若干类子区域, 其次采用两组不同的滤波器分类提取各子区域像素序列直方图频谱特征, 之后用各类子区域特征分别训练 Bayes 分类器以获得其权重, 最后对待测图像的每一个子区域进行分类检测, 并将结果加权融合得到最终判决。实验结果表明, 该方法对 LSB 匹配隐写的检测性能优于现有典型方法。

关键词: 隐写分析; 最低有效位匹配; 图像分割; 直方图频谱

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2012)06-1380-08

DOI: 10.3724/SP.J.1146.2011.00932

An Approach of Detecting Least Significant Bit Matching Based on Image Content

Xiong Gang Ping Xi-jian Zhang Tao Sun Bing

(PLA Information Engineering University, Zhengzhou 450002, China)

Abstract: Recently, it is a new direction to improve the performance of image steganalysis by combining the detection of information hidden with image content analysis. Relative to methods depending on entire image, this paper analyzes the effect of LSB (Least Significant Bit) matching steganography on image sub-areas, and presents a novel steganalyzer based on the combined discrimination. Firstly, the images are divided into several sub-areas according to the image content complexity. Secondly, the histogram spectral features of pixel sequence of each sub-area are extracted by using two different filters. Then, the Bayes classifiers are trained respectively by features of each class of sub-area in order to obtain its weights. Finally, each sub-area of a test image is detected depending on its class and the final discrimination result of the test image is achieved by weighted fusion of the results of its sub-areas. Experimental results show that the proposed method exhibits excellent performance for the detection of LSB matching, outperforms existing representative approaches.

Key words: Steganalysis; Least Significant Bit (LSB) matching; Image segmentation; Histogram spectrum

1 引言

近年来, 信息隐藏技术逐渐成为信息安全领域的一个研究热点, 其中以图像为载体的数字隐写及隐写分析是此领域的研究重点^[1]。空间域最低有效位(Least Significant Bit, LSB)匹配隐写算法是图像隐写的经典算法, 其克服了 LSB 替换的“值对效应”, 具有较高的安全性。现有针对 LSB 匹配隐写的检测大都采用基于机器学习的方法, 其实质是寻找有效分类特征对载体、载密图像进行统计分类。表 1 对现有主要检测方法所提特征进行了归纳。

文献[2-8]主要是分析 LSB 匹配隐写对图像直方

图产生的平滑效应提取分类特征。文献[9-12]主要依据 LSB 匹配隐写对自然图像相关性的影响提取分类特征。文献[13,14]是部分典型盲检测方法, 其通过分析数字隐写对图像统计性质(主要是小波域等变换域系数统计特性)的影响, 提取高阶统计量特征。此外一些基于图像统计特征的方法(如文献[15,16])也非常具有代表性。然而研究表明, 在未压缩自然图像上各检测算法的检测性能受图像内容影响较大, 且对低嵌入率隐写检测仍较为困难^[1]。为了克服图像内容的影响, 文献[17]从图像整体内容入手进行了探索。依据图像信源区域统计特性, 本文对此展开了进一步研究。

2 方法原理及流程

将静止灰度图像(如图1(a))视为一个像素灰度

2011-09-06 收到, 2012-03-05 改回

国家自然科学基金(60970142)资助课题

*通信作者: 熊钢 gangxiong1986@gmail.com

表 1 现有主要检测方法采用的分类特征

文献	分类特征
文献[2-4]	1 维、2 维灰度直方图特征函数质心
文献[5,6]	1 维、2 维灰度直方图局部极值点的邻域绝对差
文献[7,8]	1 维、2 维灰度直方图(共生矩阵)幅度值特征
文献[9,10]	差分直方图系数值及其各阶统计矩
文献[11,12]	差分像素 0 值的概率分布及差分值的 Markov 转移概率矩阵
文献[13,14]	小波子带系数及系数残差的概率函数矩及其特征函数矩

值随像素坐标变化的2维随机过程，则此过程具有统计上的区域平稳性和整体非平稳性。

图像中同一景物内部区域的数据统计特性具有一致性(如图1(b))，表现出区域平稳性。已有研究表明^[9-12]，图像信源的区域平稳性对数字隐写较为敏感，有利于提取隐写分析特征。然而，图像内容的变化使得不同区域的平滑、纹理、边缘等属性不同(如图1(b), 1(c), 1(d))，因此图像在统计上又具有整体非平稳性，图像整体非平稳性对隐写不敏感(分析见3.1节)，有利于分类处理。据此提出一种LSB匹配隐写分析方法，其具体流程如图2所示。

图2训练阶段，首先利用图像整体非平稳性，定

义内容复杂程度将图像分割成 K 类子区域；其次在各类子区域中，充分利用其区域平稳性提取隐写分析特征 Fea_1, \dots, Fea_K ，之后对 K 类特征分别进行训练以确定各类子区域的权重。测试阶段，按照与训练阶段相同的操作对待测图像进行分割和特征提取，将每一类特征送入相应的，已训练好的分类器进行判决，之后结合训练阶段确定的权重对各分类器的判决结果进行加权融合，做出最终判决。

3 图像分割

3.1 图像内容复杂度定义

为了刻画图像统计特性，下面定义图像像素灰度值的平均偏离度(记为 T)来度量图像内容复杂度。设图像基本单元 β 大小为 $M \times N$ ， x_{ij} 为坐标 (i, j) 处的像素灰度值， \bar{x} 为图像单元 β 的平均灰度值，则其平均偏离度 T 的计算为

$$T = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |x_{ij} - \bar{x}|$$

$$= \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \left| x_{ij} - \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N x_{mn} \right| \quad (1)$$

平均偏离度 T 反映了各像素取值相对于本单元

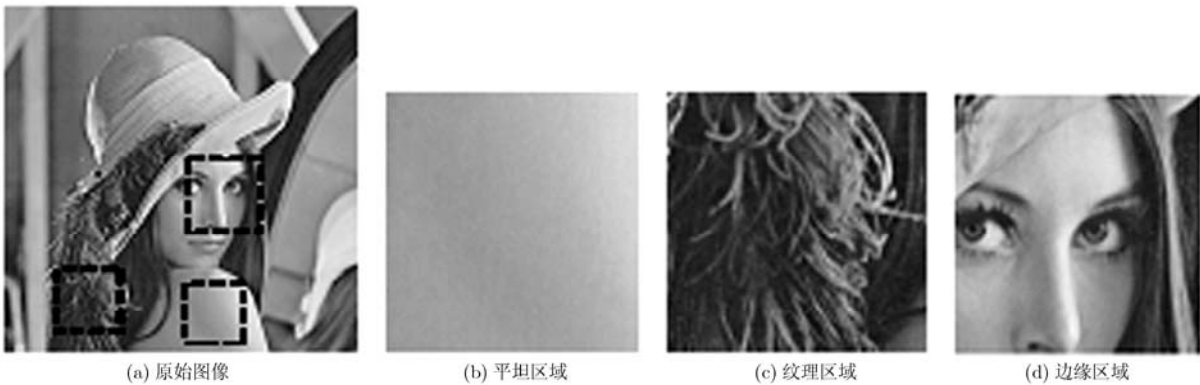


图 1 图像整体和区域统计特性示意图

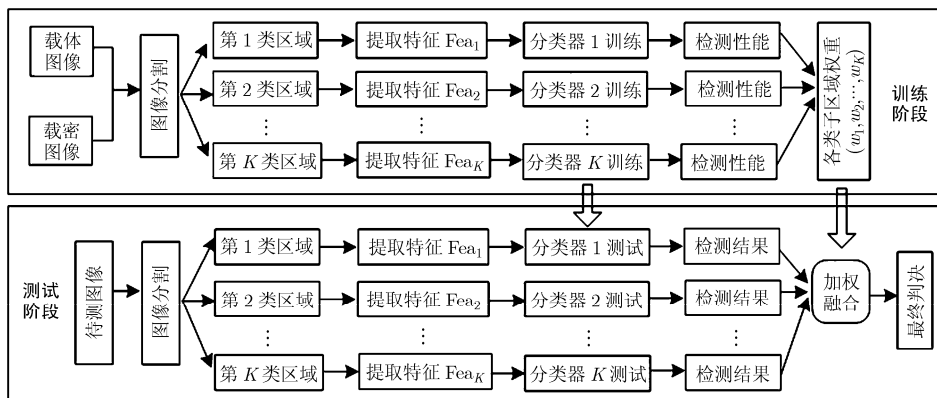


图 2 基于图像统计特性的隐写分析方法流程图

灰度均值的平均偏离程度, 图像内容越复杂, 其取值越大, 反之越小, 故可用 T 度量图像单元的复杂度。记载体、载密图像单元的平均偏离度为 T^c , T^s , 两者满足 $T^c \approx T^s$ 。分析如下:

设载体、载密图像素灰度值为 x_{ij}^c , x_{ij}^s , 基于 LSB 匹配隐写的加性噪声模型^[2], 可得

$$x_{ij}^s = x_{ij}^c + \Delta x \quad (2)$$

其中隐写噪声 Δx 的概率分布为

$$p(\Delta x) = \begin{cases} \rho/4, & \Delta x = \pm 1 \\ 1 - \rho/2, & \Delta x = 0 \end{cases} \quad (3)$$

其中 ρ ($0 \leq \rho \leq 1$) 为秘密消息嵌入率, 隐写噪声 Δx 的期望 $\overline{\Delta x} = E(\Delta x) = 0$, 由此可得

$$\begin{aligned} \bar{x}^s &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N x_{ij}^s = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij}^c + \Delta x) \\ &= \bar{x}^c + \overline{\Delta x} = \bar{x}^c \end{aligned} \quad (4)$$

据式(1), 式(4)可得

$$\begin{aligned} T^s &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x_{ij}^s - \bar{x}^s| \\ &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |(x_{ij}^c + \Delta x) - \bar{x}^c| \end{aligned} \quad (5)$$

由绝对值三角不等式性质可得

$$\begin{aligned} \frac{1}{MN} \left(\sum_{i=1}^M \sum_{j=1}^N |x_{ij}^c - \bar{x}^c| - |\Delta x| \right) &\leq T^s \\ &\leq \frac{1}{MN} \left(\sum_{i=1}^M \sum_{j=1}^N |x_{ij}^c - \bar{x}^c| + |\Delta x| \right) \end{aligned} \quad (6)$$

因为 $T^c = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x_{ij}^c - \bar{x}^c|$, $\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |\Delta x| \approx$

$E(|\Delta x|) = \frac{\rho}{2}$, 式(6)可化为

$$T^c - \frac{\rho}{2} \leq T^s \leq T^c + \frac{\rho}{2} \quad (7)$$

通常情况下, $T^c \gg \rho/2$, 所以 $T^c \approx T^s$ 。由此表明, 图像单元复杂度 T 在隐写前后基本不变, 因此将其用于图像分割能够保证图像基本单元在隐写前后其所属类别一致, 从而保证同一载体、载密图像对经分割后, 各子区域所包含的图像基本单元相一致。

3.2 图像分割具体操作

(1)图像基本单元划分 将长为 $Z \times W$ 的图像 I 分成若干个大小为 $B \times B$ 的不重叠小块, 所得小块 Block_i 即为图像基本单元。将分块处理记为 blocking_B , 则此过程如式(8)所示。

$$\bigcup_{i=1}^{\text{NB}} \text{Block}_i = \text{blocking}_B(I) \quad (8)$$

其中 \cup 表示集合并运算符; 便于计算 $B = 2^n$,

$1 \leq n \leq \lfloor \log_2 \min(M, N) \rfloor$ ($\lfloor x \rfloor$ 表示对 x 向下取整); $\text{NB} = \lfloor Z/B \rfloor \times \lfloor W/B \rfloor$ 为分块总数目。

(2)图像基本单元排序 据式(1)计算每一图像基本单元 Block 的复杂度 T , 依据 T 值将全部图像基本单元升序排列, 得到分块序列 S^{BLK} 为

$$S^{BLK} = \{\text{Block}_1^A, \text{Block}_2^A, \dots, \text{Block}_{\text{NB}}^A\} \quad (9)$$

(3)图像基本单元序列 S^{BLK} 等量分组合并 按照预先设定的类别数 K , 将序列 S^{BLK} 等分为 K 组子序列, 每一组子序列即为图像的一个子区域, 如式(10):

$$\begin{aligned} \text{Area}_k &= \{\text{Block}_{(k-1) \times \text{NK} + 1}^A, \text{Block}_{(k-1) \times \text{NK} + 2}^A, \dots, \\ &\quad \text{Block}_{k \times \text{NK}}^A\}, \quad k = 1, 2, \dots, K \end{aligned} \quad (10)$$

其中 Area_k 表示分组后的第 k 组图像子序列(即子区域), $\text{NK} = \lfloor \text{NB}/K \rfloor$ 。

(4)计算图像子区域的平均复杂度 \bar{T} 依据子区域中各基本单元的 T 值, 计算 \bar{T} 如式(11):

$$\bar{T}_k = \frac{1}{\text{NK}} \sum_{j=1}^{\text{NK}} T_{(k-1) \times \text{NK} + j}, \quad k = 1, 2, \dots, K \quad (11)$$

其中 $T_{(k-1) \times \text{NK} + j}$ 表示第 k 组图像子区域 Area_k 中第 j 个图像基本单元的 T 值。

(5)依据图像库标定各子区域类别 设图像库图像数目为 R , 则整个图像可得 $R \times K$ 个子区域平均复杂度 \bar{T} , 依升序排列 \bar{T} 得序列 $S^{\bar{T}} = \{\bar{T}_1^A, \bar{T}_2^A, \dots, \bar{T}_{R \times K}^A\}$, 将 $S^{\bar{T}}$ 中 $\bar{T}_{(k-1) \times R + 1}^A \sim \bar{T}_{k \times R}^A$ 所对应的子区域标定为第 k 类, 最终整个图像库得到 K 类图像子区域。再将 $R \times K$ 个子区域还原到每一幅图像, 则单幅图像的第 k 个子区域被标定为类别 L_k ($1 \leq L_k \leq K$)。

上述图像分割方法特点是不需要设定 T 阈值且保证了各类子区域的数据量基本一致。下面将一幅实验图像分割为 3 类 ($K=3$), 结果如图 3 所示(分块大小 8×8)。

图 3 所示分割结果表明, 原始图像经过分割处理后得到了 3 类具有不同内容复杂度的子区域, 此分割方法打乱了原图中各图像基本单元的物理位置, 使得分散在整幅图像中的具有相似统计特性的图像基本单元得到集中, 从而有利于提高隐写分析特征的灵敏度。

4 特征分析和提取

4.1 子区域像素扫描

像素点 (x, y) 与其周围 8 个像素存在相邻关系如图 4(a)所示。为了分析隐写对图像不同方向相关性的影响, 拟采用 5 种方式对图像基本单元进行像素扫描。记扫描方式为 f , 则行扫描 ($f=1$, 扫描矩阵 M_1

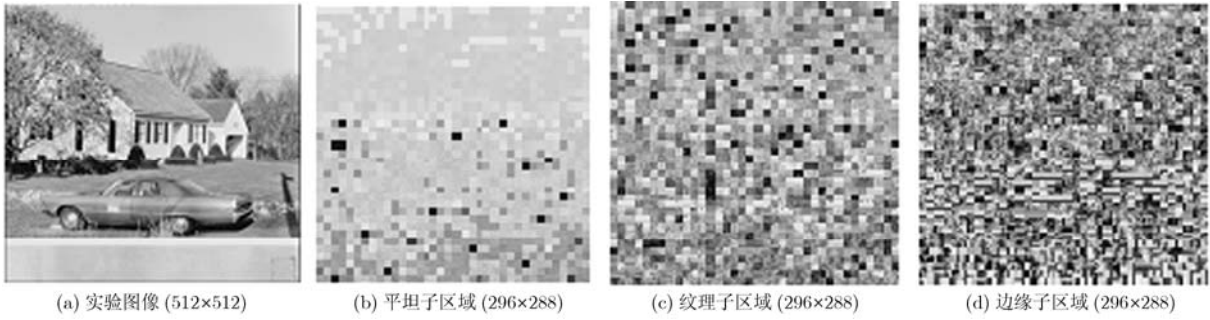


图 3 图像分割结果示意图

如图 4(b))和列扫描($f=2$, 扫描矩阵为 M_1 转置)能保持水平和垂直方向的相关性; Z 扫描($f=3$, 扫描矩阵 M_3 如图 4(c))与反 Z 扫描($f=4$, 扫描矩阵为 M_3 转置)能保持对角方向的相关性; Hilbert 扫描($f=5$, 扫描矩阵 M_5 如图 4(d))是一种随机扫描, 有利于保持基本单元内的整体相关性。

图像各子区域 $Area_k$ 内的图像基本单元经扫描转化为相应的 1 维像素序列, 再将子区域内所有像素序列连接成总像素序列 S_{kf}^p 。

4.2 序列灰度直方图频谱分析

记载体、载密像素序列灰度直方图为 h_c, h_s , 基于加性噪声模型^[2]可得式(12)。

$$h_s(x) = h_c(x) * h_p(\Delta x) \quad (12)$$

其中 $h_p(\Delta x)$ 表示嵌入率为 ρ 时嵌入隐写噪声的概率分布, 如式(3), “*” 表示卷积运算符。

对灰度直方图 $h(x) (x \in [x_{\min}, x_{\max}])$, 进行 N^h ($N^h = 2^{\lceil \log_2(x_{\max} - x_{\min} + 1) \rceil}$), $\lceil x \rceil$ 表示对 x 向上取整)点离散傅里叶变换(DFT)得到其频谱 H 为

$$H(v) = \sum_{n=0}^{N^h-1} h(n + x_{\min}) \exp\left\{-\frac{j2\pi nv}{N^h}\right\} \\ = |H(v)| \exp(\varphi(v)), 0 \leq v \leq N^h - 1 \quad (13)$$

其中 $|H(v)|$ 为频谱幅度, $\varphi(v)$ 为频谱相位, v 是数字频率, 等于数字角频率(rad)与 2π (rad) 的比值, 是一个无量纲的物理量。时域卷积等效于频域相乘, 据式(12), 式(13)可得

$$H_s(v) = H_c(v) \times H_p(v) \quad (14)$$

据傅里叶变换的性质, 噪声频谱幅度满足式(15):

$$|H_p(v)| = \left| \sum h_p(n) \exp\left\{-\frac{j2\pi nv}{N}\right\} \right| \leq \sum |h_p(n)| = 1 \quad (15)$$

因此, 据式(14), 式(15)可得

$$|H_s(v)| = |H_c(v)| \times |H_p(v)| \leq |H_c(v)| \quad (16)$$

式(16)说明嵌入秘密消息之后, 图像像素序列的灰度直方图频谱的幅值会减小。

4.3 序列差分直方图频谱分析

进一步, 利用像素差分序列来刻画 LSB 匹配隐写对序列 S^p 相关性的影响。运用差分模板 m_i 对像素灰度序列 S^p 滤波得到像素差分序列 S^d , 滤波处理如式(17)所示。

$$S_i^d = S^p * m_i \quad (17)$$

其中所用差分模板包括: $m_1 = [1 \ -1]$, $m_2 = [1 \ -2 \ 1]$, $m_3 = [1 \ 0 \ -1]$ 。

设载体、载密像素序列差分直方图为 h_c^d, h_s^d , 据文献[4]和文献[11]分析可得式(18):

$$h_s^d(x^d) = h_c^d(x^d) * h_p(\Delta x) * h_p(\Delta x) \quad (18)$$

根据式(13)对序列差分直方图 $h^d(x^d)$ 进行 DFT 得到频谱 H^d , 同理, 据式(15), 式(18)可得

$$|H_s^d(v)| = |H_c^d(v)| \times |H_p(v)|^2 \leq |H_c^d(v)| \quad (19)$$

即嵌入秘密消息之后, 序列差分直方图频谱的幅值也会减小, 且减小幅度大于灰度直方图。

提取图 3 中 3 个子区域经行扫描和差分模板 m_1 处理后的差分直方图频谱如图 5 所示。

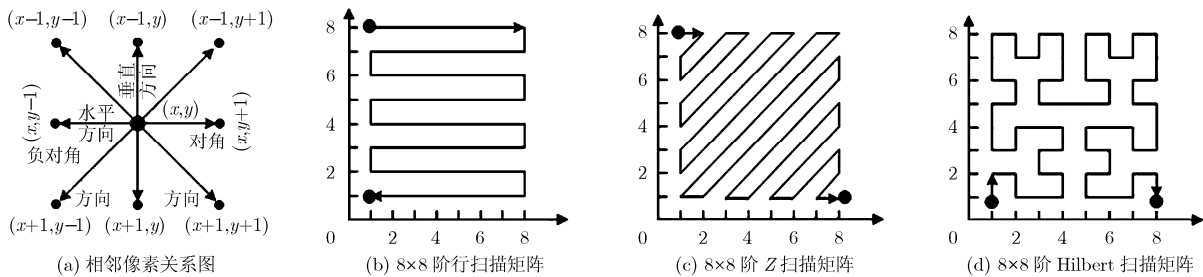


图 4 扫描矩阵示例

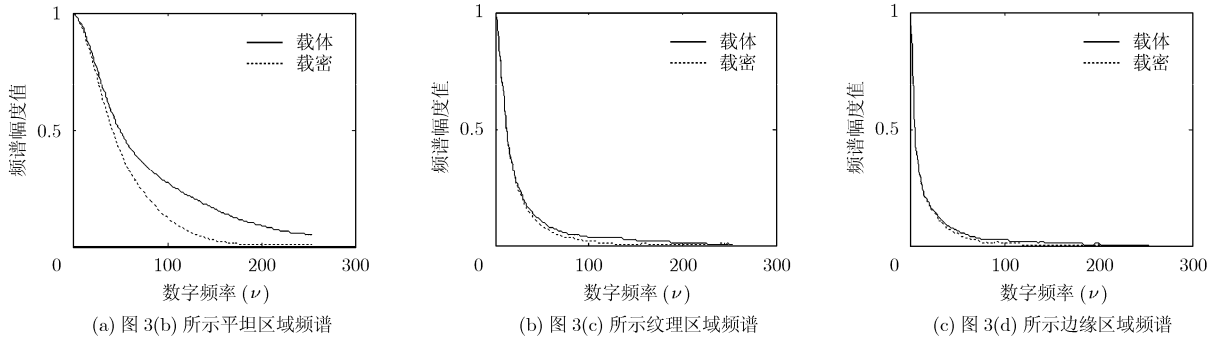


图 5 各子区域序列差分直方图频谱示意图

图 5 表明 LSB 匹配隐写使得差分直方图频谱幅度值减小, 实验结果与式(19)分析相符合。对比可得, 子区域内容复杂度越小, 频谱幅度差异越大。灰度直方图频谱具有相似性质。

4.4 特征提取

进一步分析图 5 可得, 平坦子区域频谱幅度的差异主要分布在中高频区域, 边缘子区域的频谱幅度差异主要分布在中频或中低频区域, 纹理子区域频谱介于两者之间。据此设计具有中低频和中高频带通的两组滤波器捕捉频谱幅度的最大差异, 滤波器系统函数频谱如式(20), 式(21)所示。

$$F_u^1(v) = \frac{2\tau}{1 + \tau^2} = \frac{4v(u+1)N}{N^2 + 4v^2(u+1)^2},$$

$$\tau = \frac{v(u+1)}{N/2}, v = 0, \dots, \frac{N}{2} - 1 \quad (20)$$

$$F_u^2(v) = \sin\left(\frac{\pi(u+1)v}{N}\right), v = 0, \dots, \frac{N}{u+1} - 1 \quad (21)$$

其中 $F_u^1(v)$ 具有高通滤波性质, $F_u^2(v)$ 具有中低通滤波性质; $u(1 \leq u \leq U)$ 的变化可以控制滤波器的滤波性能(如图 6 所示)。实验可得载体、载密序列直方图频谱的最大差异位置约位于 $N/4 \sim N/10$ 之间, 因此设定 $U = 4$ 。 N 为直方图频谱的频率点数, 频谱具有对称性, 因此考虑半频谱 ($v = 0, 1, \dots, N/2 - 1$)。

将滤波器系统函数频谱 $\hat{F}(v)$ 与直方图频谱

$\hat{H}(v)$ 相乘得滤波后频谱 $\widehat{HF}(v)$, 如图 6 所示。

计算滤波后所得频谱 $\widehat{HF}(v)$ 内所有频率幅度值之和的归一化值作为特征, 如式(22):

$$\widehat{\text{Sum}} = \frac{\sum_{v=v_{\min}}^{v_{\max}} |\widehat{HF}(v)|}{\sum_{v=v_{\min}}^{v_{\max}} |\hat{H}(v)|} = \frac{\sum_{v=v_{\min}}^{v_{\max}} |\hat{H}(v)| |\hat{F}(v)|}{\sum_{v=v_{\min}}^{v_{\max}} |\hat{H}(v)|} \quad (22)$$

其中 v_{\min} , v_{\max} 分别表示 $\widehat{HF}(v)$ 的起始和截止频率, 归一化处理能够减小特征取值的动态范围。

实验中利用以上特征计算方法, 对 K 类区域分别用 $F_u^1(v)$ 和 $F_u^2(v)$ 提取前 $\lfloor K/2 \rfloor$ 类子区域和后 $K - \lfloor K/2 \rfloor$ 类子区域特征。灰度序列 S^p 其与像素空间位置无关, 因此任取一扫描序列依据带通滤波器数 u 提取 4 维灰度直方图特征 Feature^1 。对差分序列 S_{mf}^d , 依据带通滤波器数 u , 扫描方式 f 及滤波模板 m 的不同构成 $60(4 \times 5 \times 3 = 60)$ 维特征 Feature^2 , 如式(23):

$$\text{Feature}^i = \begin{cases} \widehat{\text{Sum}}^{iu}, & i = 1 \\ \widehat{\text{Sum}}_{mf}^{iu}, & i = 2 \end{cases},$$

$$u = 1, 2, 3, 4, f = 1, 2, \dots, 5, m = 1, 2, 3 \quad (23)$$

据式(23)将两类特征合并得到 64 维特征向量:

$$\text{Feature} = [\text{Feature}^1, \text{Feature}^2].$$

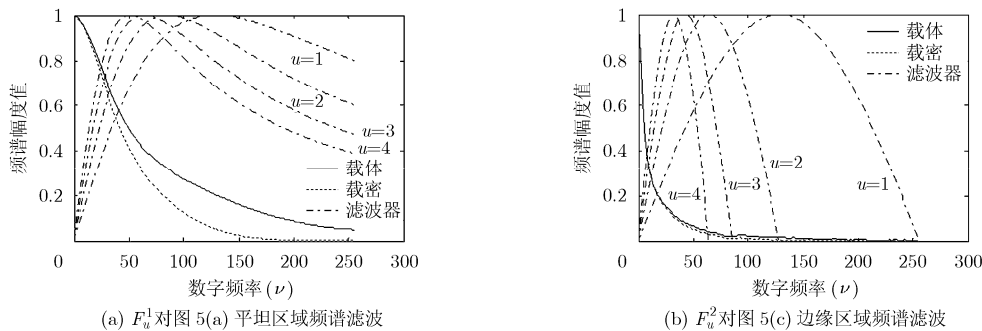


图 6 直方图频谱带通滤波示意图

5 分类测试与联合判决

贝叶斯(Bayes)决策理论对求解问题建立概率模型, 所得后验概率能够保留更多信息用于各子区域检测结果融合, 因此下面基于 Bayes 分类器进行分类测试和联合判决。

训练阶段主要是确定 K 类子区各自权重(如图 2)。由训练可得第 k 类子区域的 ROC 曲线, 记 ROC 下的面积为 AUC_k , 据其计算得到第 k 类子区域检测可靠性 η_k 值, 再据 η_k 可得各子区域类别的权重 w_k 为

$$w_k = \frac{\eta_k}{\sum_{k=1}^K \eta_k} = \frac{2AUC_k - 1}{\sum_{k=1}^K (2AUC_k - 1)}, \quad k = 1, 2, \dots, K \quad (24)$$

其中 $\eta_k = 2AUC_k - 1$ ($0 \leq \eta_k \leq 1$), 子区域检测性能越好, 则 η_k 越大, 分配的权重 w_k 也越大。

在测试阶段(如图 2), 设待测图像被分割成 K 个图像子区域, K 个子区域共 L 类, 每一类记为 Y_l ($1 \leq l \leq L$), 第 k 个图像子区域类别为 L_k ($1 \leq L_k \leq K, L \leq K$), 经 Bayes 分类器判别后得到其属于载体、载密的后验概率为 $P_{L_k r}$, 显然 $P_{L_k 1} + P_{L_k 2} = 1$ 。现采用“先类内平均, 后类间加权”的融合规则对 K 个后验概率 $P_{L_k r}$ 融合。

据上述参数可得“类内平均”的计算如式(25):

$$P_{Y_l r} = \begin{cases} \frac{1}{NY_l} \sum P_{L_k r}, & L_k = Y_l (1 \leq k \leq K) \\ 0, & \text{其它} \end{cases}, \quad Y_1 \leq Y_l \leq Y_L \quad (25)$$

其中 NY_l 表示 K 个子区域中属于类别 Y_l 的数目, $r \in \{1, 2\}$ 分别表示载体、载密图像类别。

将不同类别子区域的后验概率 $P_{Y_l r}$ 进行“类间加权”得到融合的后验概率 P_r 计算如式(26):

$$P_r = \frac{1}{L} \sum_{Y_l=Y_1}^{Y_L} [w_{Y_l} (P_{Y_l r} - 0.5) + 0.5] \quad (26)$$

据式(25), 式(26)可得 $P_1 + P_2 = 1$ 。因此可得 K 个子区域的后验概率 $P_{L_k r}$ 融合后所得 P_r ($r=1, 2$) 满足概率要求。据所得后验概率 P_1 和 P_2 , 采用最大后验概率准则对整幅图像进行最终判决。

6 实验结果与分析

6.1 实验设置

为了较全面地评价算法的检测性能, 分别在以下 4 个未压缩图像库上进行实验: (1)CAMERA 图像库^[18]: 包含 3164 幅大小为 512×512 的彩色图像, 随机选择 2000 幅图像; (2)UCID 图像库^[19]: 包含 1338 幅大小为 512×384 或 384×512 的彩色图像, 随机选择 1000 幅图像; (3)NRCS 图像库^[20]: 包含

3185 幅分辨率为 1500×2100 或 2100×1500 的彩色图像, 随机选择 2000 幅图像; (4)Digital Image101 (DI101)自建图像库: 包含 3000 幅分辨率为 3872×2592 的 RAW 格式彩色图像, 随机选择 2000 幅图像并中心裁剪为 512×512 大小。

实验中, 将 4 个载体图像库图像转化为灰度图像, 并用 LSB 匹配隐写构成载密图像库(嵌入率^[11]分别为 10%, 25%, 50%)。对于每个图像库, 随机选取的 1/5 图像作为训练样本, 剩余 4/5 作为测试样本。选用融合后的后验概率 P_2 与设定的阈值相比较, 得到 ROC 曲线下的面积 AUC 作为性能评价指标, AUC 指标在 0~1 之间取值, 其值越接近于 1, 检测性能越好。为了减小训练, 测试样本随机性对性能的影响, 取 20 次实验的平均值作为检测性能的最终评价指标。

6.2 分类检测性能分析

分块大小(B)和子区域类别数目(K)是两个重要参数。在 UCID, NRCS 两个图像库上进行多组实验, 实验中 B 主要取 2^n , K 主要取 $2n$ ($n \in N^+$), 部分实验结果如表 2 所示。

比较表 2 中各组数据可得两点结论: (1) $K=1$ 时表示将载体图像中的所有分块单元视为同一类别, 不对图像进行分割处理, 采用式(20)对其提取特征, 其检测性能(斜体字所示)低于 K 取其它值时的性能(黑体字所示为实验中分类检测的最好性能), 由此表明对图像各子区域分类处理有利于提高检测性能。(2)分类对 NRCS 库检测性能的提高大于 UCID 库。分析原因是: UCID 库中原始图像大小为 512×384 , 由于图像大小的限制使得分类检测的优势不明显。而 NRCS 库中图像大小为 1500×2100 , 当分块大小为 $B=32, 64, K=8, 10$ 时体现出较好的分类检测性能, 因为分块大小 64×64 左右时块内像素统计特性较为一致, 有利于隐写特征提取; 分类数目 $K=8$ 左右时能够保证一定的类别数进行类间加权融合。通常情况下, 图像尺寸越大, 其图像内部各子区域统计特性更加明显, 更有利于分类检测。

6.3 不同算法检测性能比较

本文算法(记为 S0)与 9 种典型的隐写分析算法(记为 S1~S9)的检测结果如表 3 所示。

表 3 给出了 10 种算法在 3 个图像库上的实验结果, 图 7 给出了本文算法与部分算法的 ROC 曲线。分析表 3 数据可得:

(1)横向比较, 不同图像库, 不同嵌入率条件下, 本文算法都获得了较好的实验结果, 说明本文算法受图像内容影响较小, 所提特征能够有效捕获 LSB 匹配隐写在图像中所产生的变化, 在嵌入率较低时

表 2 不同分块大小(B)和类别数目(K)情况下 AUC 值(嵌入率 50%)

参数取值	UCID				NRCS			
	B=8	B=16	B=32	B=64	B=16	B=32	B=64	B=128
K=1	0.930	0.941	0.936	0.935	0.843	0.865	0.877	0.863
K=4	0.936	0.949	0.953	0.951	0.924	0.928	0.931	0.922
K=8	0.942	0.955	0.951	0.950	0.928	0.945	0.943	0.934
K=10	0.938	0.947	0.950	0.941	0.931	0.948	0.951	0.937

表 3 不同算法在 3 个图像库上的 AUC 指标(fea 表示特征, F 表示采用 Fisher 线性分类器)

算法	嵌入率		UCID			CAMERA			DI101			
			50%	25%	10%	50%	25%	10%	50%	25%	10%	
针对性检测	S1	文献[3]	2 fea(F)	0.646	0.576	0.523	0.731	0.614	0.551	0.698	0.603	0.533
	S2	文献[4]	1 fea(F)	0.665	0.559	0.519	0.679	0.576	0.527	0.675	0.568	0.520
	S3	文献[15]	1 fea(F)	0.677	0.591	0.538	0.728	0.601	0.542	0.680	0.595	0.541
	S4	文献[11]	1 fea(F)	0.642	0.554	0.511	0.723	0.59	0.516	0.691	0.575	0.515
	S5	文献[7]	50 feas(F)	0.837	0.698	0.569	0.957	0.845	0.721	0.918	0.810	0.690
	S6	文献[10]	24 feas(F)	0.875	0.818	0.677	0.988	0.964	0.870	0.935	0.861	0.750
	S7	文献[9]	70 feas(F)	0.912	0.819	0.725	0.997	0.980	0.914	0.952	0.913	0.802
盲检测	S8	文献[14]	27 feas(F)	0.687	0.599	0.523	0.902	0.838	0.719	0.867	0.782	0.622
	S9	文献[13]	104feas(F)	0.894	0.781	0.642	0.963	0.902	0.788	0.932	0.842	0.687
本文	S0	B=16 K=8	64 feas	0.955	0.851	0.768	0.998	0.986	0.922	0.989	0.962	0.858

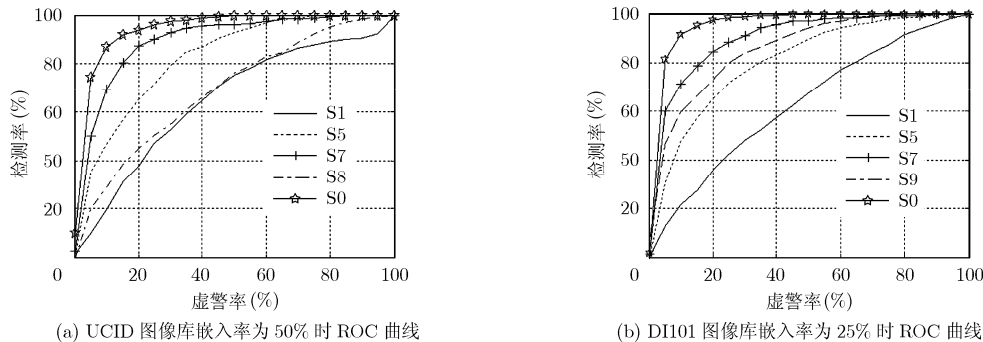


图 7 不同嵌入率下两图像库上部分算法的 ROC 曲线

(以 10%为例)也体现出较好的检测性能。

(2)纵向比较,同一图像库上各种算法检测性能均不相同,本文算法(黑体字)检测性能优于 9 种典型方法最好检测结果(S7 的黑体字)。然而,由于 3 个图像库中数据统计特性互异,对其检测性能又体现出不同特点。在 CAMERA 库上,本文算法与算法 S7 的检测性能相当,均体现出较高的检测率;在 UCID 库和自建图像库 DI101 上,本文算法在特征数目相对较少的情况下,通过对图像进行分区域处理并借助于多区域联合判决明显提高了检测性能(如图 7(a), 7(b)所示)。

7 结束语

本文以图像信源的统计特性为基础,一方面引

入化整体为区域的思想,通过图像分割将图像内容统计特性相一致的像素进行集中并分类处理,有效降低了图像内容对隐写分析的影响;另一方面借鉴现有隐写分析方法的研究成果,充分挖掘子区域内的敏感特征。实验结果表明,本文方法能有效检测未压缩图像的 LSB 匹配隐写,并对图像空域非自适应隐写方法具有一定的通用性。此外,针对新出现的空域自适应隐写方法和抗盲检测隐写方法^[21]的检测性能,有待于实验的进一步检验。因此,下一步研究方向是深入研究图像内容与隐写分析之间的关系,并将本文的思想和方法进行延拓,以进一步提高算法的检测性能和通用性。

参考文献

[1] Li B, He J, Huang J, et al. A survey on image steganography

- and steganalysis [J]. *Journal of Information Hiding and Multimedia Signal Processing*, 2011, 2(2): 142-172.
- [2] Harmsen J and Pearlman W A. Steganalysis of additive noise modelable information hiding[C]. Proceedings of the SPIE Security, Steganography, and Watermarking, San Jose, 2003: 131-142.
- [3] Ker A D. Steganalysis of LSB matching in grayscale images[J]. *IEEE Signal Processing Letters*, 2005, 12(6): 441-444.
- [4] Li X, Zeng T, and Yang B. A further study on steganalysis of LSB matching by calibration [C]. Proceedings of IEEE on Image Processing, Piscataway, NJ, 2008: 2072-2075.
- [5] Zhang Jun, Cox I J, and Doërr G. Steganalysis for LSB matching in images with high-frequency noise[C]. Proceeding of the IEEE Workshop on Multimedia Signal Processing, Piscataway, NJ, 2007: 385-388.
- [6] Cancelli G, Doërr G, Cox I J, *et al.* Detection of ± 1 LSB steganography based on the amplitude of histogram local extrema [C]. Proceedings of IEEE on Image Processing, Piscataway, NJ, 2008: 1288-1291.
- [7] Gao Y, Li X, Yang B, *et al.* Detecting LSB matching by characterizing the amplitude of histogram [C]. Proceedings of IEEE on Acoustics Speech, Signal Processing, Piscataway, NJ, 2009: 1505-1508.
- [8] Kekre H B, Athawale A A, and Patki S A. Steganalysis of LSB embedded images using gray level co-occurrence matrix [J]. *International Journal of Image Processing*, 2011, 5(1): 36-45.
- [9] Cai K, Li X, and Zeng T. Reliable histogram features for detecting LSB matching [C]. Proceedings of IEEE International Conference on Image Processing, Hong Kong, 2010: 1761-1764.
- [10] Zheng Ergong, Ping Xijian, Zhang Tao, *et al.* Steganalysis of LSB matching based on local variance histogram [C]. Proceedings of IEEE on Image Processing, Hong Kong, 2010: 1005-1008.
- [11] Zhang Tao, Li Wenxiang, Zhang Yan, *et al.* Steganalysis of LSB matching based on statistical modeling of pixel difference distributions [J]. *Information Sciences*, 2010, 180(23): 4685-4694.
- [12] Pevný T, Bas P, and Fridrich J. Steganalysis by subtractive pixel adjacency matrix [J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(2): 215-224.
- [13] Wang Y and Moulin P. Optimized feature extraction for learning-based image steganalysis [J]. *IEEE Transactions on Information Forensics and Security*, 2007, 2(1): 31-45.
- [14] Goljan M, Fridrich J, and Holotyak T. New blind steganalysis and its implications [C]. Proceedings of SPIE Security, Steganography, and Watermarking, Wash, 2006, 6072: 1-13.
- [15] Huang F and Li B. Attack LSB matching steganography by counting alteration rate of the number of neighbourhood gray levels [C]. Proceedings of IEEE on Image Processing, Piscataway, NJ, 2007: 401-404.
- [16] 廖鑫, 温巧燕. 基于拉普拉斯算子统计量的LSB替换隐写分析方法[J]. 电子与信息学报, 2009, 31(5): 1054-1058.
- Liao Xin and Wen Qiao-yan. Detection methods for LSB embedding based on laplacian statistics [J]. *Journal of Electronics & Information Technology*, 2009, 31(5): 1054-1058.
- [17] Hossein Amirkhani and Mohammad Rahmati. New frame work for using image contents in blind steganalysis systems [J]. *Journal of Electronic Imaging*, 2011, 20(1): 013016-1-013016-14.
- [18] Doërr G. Database for steganalysis studies. <http://www.cs.ucl.ac.uk/staff/I.Cox/Content/Downloads.html>. 2010.
- [19] UCID. http://www.staff.lboro.ac.uk/_cogs/datasets/UCID/ucid.html. 2010.
- [20] USDA NRCS Photo Gallery. <http://photogallery.nrcs.usda.gov>. 2010.
- [21] 刘九芬, 韩涛, 张卫明, 等. 一种基于湿纸编码和图匹配理论的直方图保持隐写算法[J]. 电子与信息学报, 2011, 33(3): 592-596.
- Liu Jiu-fen, Han Tao, Zhang Wei-ming, *et al.* A histogram-preserving steganography based on wet paper coding and graphic matching theory[J]. *Journal of Electronics & Information Technology*, 2011, 33(3): 592-596.
- 熊 钢: 男, 1986年生, 硕士生, 研究方向为图像处理、信息隐藏。
- 平西建: 男, 1953年生, 教授, 博士生导师, 研究方向为图像分析与处理、模式识别、信息隐藏。
- 张 涛: 男, 1977年生, 副教授, 硕士生导师, 研究方向为图像处理、信息隐藏。