

一种基于图像内容的隐写分析方法

李 韬, 平西建

解放军信息工程大学 信息工程学院, 郑州 450002

摘 要: 数字图像已成为信息隐藏的一类重要载体, 然而不同内容的图像对隐写分析呈现出不同的检测性能. 该文基于图像信源区域平稳马尔可夫特性, 用方差度量图像的区域复杂度, 运用四叉树分割方法对图像进行递归分解, 提取对数字隐写较为敏感的平坦区域, 并从这些区域提取游程长度直方图统计矩特征, 最后结合支持向量机实现隐藏信息检测. 实验结果表明, 依据图像内容分割出敏感区域进行隐写分析可有效改善检测性能.

关键词: 隐写分析; 图像内容; 四叉树; 游程长度直方图

中图分类号: TP391

文章编号: 0255-8297(2013)04-0394-08

Steganalysis Based on Image Contents

LI Tao, PING Xi-jian

Institute of Information Engineering, PLA Information Engineering University, Zhengzhou 450002, China

Abstract: Digital images are important covers in information hiding. It has been shown that steganalytic methods perform differently on images with different contents. Based on regionally stationary Markov characteristics of natural images, a new steganalysis method based on image segmentation is proposed. Quad-tree segmentation based on local Variance is used to locate areas sensitive to message embedding. Features based on run length histogram are then extracted and optimized. Finally, detection of steganography is done using support vector machine. Experimental results show that the proposed steganalysis method outperforms previous methods in detection accuracy.

Keywords: steganalysis, image content, quad-tree, run length histogram

数字图像广泛存在于互联网之中并成为人们获取信息的必要媒介, 正是它的公开性和海量性, 使之成为数字隐写的重要载体. 随着当代隐写术的发展, 隐写分析已成为信息安全领域的研究热点.

不同图像对隐写分析的结果造成的影响不同, 已有研究者在这方面进行了初步的探索. 文献 [1] 指出, 特征函数质心 (histogram characteristic function center of mass, HCFCOM) 特征在不同图像之间变化很大, 其变化范围远大于载体、载密图像对之间的差异; 为了减小不同图像对特征值的影响, 以降采样图像作为载体图像的估计, 并用降采样图像的特征值对原图像的特征进行校准, 定义了校准的 HCFCOM 特征. 文献 [2] 对 3 种不同 LSB 匹配隐写分析算法在不同图像库上进行比较研究, 结果表明 3 种算

法在各图像库上的绝对性能和相对性能均有相当大的变化. 文献 [3] 分析了载体图像固有特征和数字隐写引入特征的差异, 提出了一种基于整体性特征提取和融合的隐写分析方法. 文献 [4] 分析了隐蔽通信系统与传统通信系统的差异, 指出信息隐藏中具体的载体数据是一个具有实际信息内容含义的确定型信号, 在隐写分析的算法设计中将载体、攻击信道看作高斯白噪声的假设是不成立的或只能假设为近似成立. 文献 [5] 将图像小波子带系数建模为广义高斯分布, 用这一分布的形状参数来刻画图像复杂度, 提取了位平面内和位平面间的相关性特征, 并用实验验证了图像统计显著性、隐写分析检测性能、数字隐写嵌入率与图像复杂度之间存在着联系. 文献 [6] 认为隐写嵌入对具有相似内容的图像具有相似的影响, 进而把图像分成大

收稿日期: 2011-12-12; 修订日期: 2012-04-11

基金项目: 国家自然科学基金(No.60970142)资助

通信作者: 平西建, 教授, 博导, 研究方向: 图像处理、模式识别、信息隐藏, E-mail: pingxijian@yahoo.com.cn

小一致的块，分块分析并联合判决，该方法具有一定的启发意义。文献 [7] 把图像分为 low activity 和 high activity 区域，然后从较平坦的 low activity 区域中计算参数，估计嵌入率。

图像反映客观世界中不同景物空间结构关系所呈现的视觉信息。自然景物千差万别，获取图像的手段和技术的不同，都使图像的内容大不相同。然而，具有实际内容的图像，其内部的灰度、颜色、纹理具有某种同一性关系。在某一局部区域内，图像数据统计特性的变化较小，具有近似的平稳性。如细节较少、内容比较单一的平坦区域，像素间的相关性大；而细节较多的纹理区和灰度变化较大的边缘区，数据之间相关性则较小。因此，可以将图像信源建模为区域平稳的 Markov 信源。

本文从图像信源统计特性出发，分析隐藏信息存在性特征与图像内容之间的关系，提出一种基于图像内容特征的隐写分析方法。首先以方差作为图像复杂度的测度，用四叉树分割的方法把图像中对隐写特征敏感的区域提取出来，然后在这些区域提取游程长度直方图统计矩特征。实验结果表明：与未考虑图像内容的方法相比，本文算法的检测率有所改善。

1 基于图像分割的隐写分析方法

图像信号是一种记忆性的信源，对于一幅尺寸为 $M \times N$ 的二维静止灰度图像 I ，其第 i 行第 j 列的像素灰度值为 $I_{i,j}$ 。如果每个像素有 p 个灰度等级，则该图像共可能有 $p^{M \times N}$ 种不同灰度分布的图案。图像的种类多种多样，内容千变万化， $p^{M \times N}$ 个信源符号可能存在于一个非常大的概率空间，以图像像素灰度值 8 比特表示为例，此概率空间就有 $256^{M \times N}$ 个信源符号。由于图像的高维特性，对自然图像的统计特性进行分析是非常困难的。在实际情况中，可能出现的信源符号在 $p^{M \times N}$ 中只占一小部分。在小的空间邻域条件下，某一像素的概率密度与该邻域内的像素相关性较大，而与邻域外的像素相关性较小；邻域内像素的分布与邻域在图像中的绝对位置无关。因此，图像信源可以被看作区域平稳的 Markov 信源。

1.1 图像统计特征分析

特定目标图像内容丰富，复杂多变。为了研究秘密信息序列的嵌入在不同图像内容中的差异，本文建立了一个高精度未压缩图像库，其中所含图像使用 Nikon D40X 拍摄，大小为 $3\,872 \times 2\,592$ 。纹理是描述图像内容的一个重要特征，对自建图像库中的图像按照纹理复杂度不同划分为平坦图像和纹理图像两类，裁剪至 512×512 大小，每类图像各 500 幅。从视觉上可以看出，平坦图像内容变化缓慢，细节较少，表现

在频域就是大部分为低频信号；而自然纹理图像具有较多的细节内容，故在频域中有大量高频分量。

为了考察图像内容与隐写分析的相关性，以 LSB 匹配隐写方法为例，分别对两个图像库嵌入 1.0 bpp(bit per pixel) 和 0.5 bpp 的秘密信息，用比较有代表性的方法：Harmsen 直方图质心^[8]、局部极值特征^[9]和 Ker 邻接直方图质心特征^[10]分别在两个图像库上提取特征，其中对载体载密图像库特征的统计量取平均，统计结果如表 1 所示。可以看出，对于同一内容类型的图像，嵌入率越高，载体、载密图像之间统计量的差异越大；对于不同类型的图像，图像越平坦，载体、载密图像之间统计量的差异越大。

表 1 自建分类载体/载密图像库上不同隐写分析特征的均值

	平坦图像			自然纹理图像		
	载体	载密	载密	载体	载密	载密
	0.5 bpp		1.0 bpp	0.5 bpp		1.0 bpp
HCF	0.012	0.009	0.010	0.004	0.004	0.004
ALE	0.065	0.044	0.031	0.040	0.038	0.036
AHCF	77.361	57.806	44.776	97.184	92.658	90.420

不同的数字图像，内容不尽相同，其统计特性亦不相同，而嵌入强随机秘密信息对不同图像的统计特性的改变也将不同，因此对由不同内容构成的图像子区域的统计特性产生不同的影响。平坦图像对秘密信息的嵌入比较敏感，对这一类图像的检测也较容易；而纹理丰富的图像因为本身统计特性复杂，有类似于隐写噪声的高频成分，对其检测比较困难。文献 [9] 从不同照相机拍摄和不同内容图片的隐写分析实验中得出：有一些图像在检测中总是被误判（载体被判为载密，载密被判为载体），而这些图片的共同点就是拥有大量的纹理区域或是内容丰富而多变。那些始终能被正确检测的图片，往往是那些内容比较单一或者拍摄时失焦、抖动（造成的结果是图像像素间相关性很强）。实验表明，对纹理复杂图像的隐写分析即使在高维特征空间中也很难做到，这与本文的实验结果相符合。

1.2 四叉树图像分割

由以上分析可知，充分利用图像中对隐写分析特征敏感的区域，尽量减少不敏感甚至会对判别结果产生误导的区域在隐写分析中的比例，可以提高隐写分析的可靠性。要达到这样的目的，引入图像分割的方法。四叉树算法^[11]是 1991 年由 Fisher 提出的，广泛

应用于数据压缩,具有分块灵活性高以及压缩比率高的优点.它将图像表示成一棵四叉树,树根就是原图像本身,除叶节点外,树中每个节点均有4个子节点,分别对应于原图像(或图像块)4个象限的子块.

四叉树图像分割的方法运用递归分解的原则,将原始图像分成4个相同大小的方块,判断每个方块是否满足某一判别准则,如果满足就不再继续分解,如果不满足就再细分成4个方块.这个迭代重复的过程直到所有块都满足这一判别准则才停止.最后,图像四叉树分割的结果可能包括多种不同尺寸的方块.方差可以体现块内像素变化的剧烈程度,也可以一定程度上反映图像块的复杂程度,设定大小为 $M \times M$ 图像 I 中的像素值为 $p(i, j)$, 首先计算图像整体方差

$$V(I) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M \left[p(i, j) - \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M p(i, j) \right]^2 \quad (1)$$

然后,计算子块的方差 v , 本文所设定的判别准则如下:

$$v \geq \lambda V(I) \quad (2)$$

图1是标准图像Lena取不同 λ 四叉树分割后的示意图.可以看出,此分割算法可以很好地体现图像的纹理特性,较大的图像块反映了平坦的区域,而较小的块反映了纹理复杂的区域或边缘区域. λ 取值直接决定了分割的精细程度.试验表明,当 λ 取值较小时,更多的子图像块被分解,对图像的纹理更加敏感,但是容易过分解,对纹理较复杂的图片分割效果不好;反之,较大的 λ 值会造成分解不足,不能准确刻画边缘信息.结合本文的运算需求,在既能刻画图像纹理特性又能提取一定数量平坦区域的原则下,本文把 λ 设定为5,这样的阈值可以达到不错的分割效果,并且对纹理较复杂的图像也能进行有效的分割.运用四叉树分割有如下几点优点:1) 分割算法简单,运算复杂度低,计算时间短;2) 分块比较整齐,均为 $2^N \times 2^N$ 大小的方块,易于提取并作下一步运算;3) 分割后子图像块的大小在一定程度上就可以反映图像的纹理复杂度,于是可以依据块的大小来对不同内容图像块进行分类处理.

1.3 游程长度直方图

由1.2节可知,图像被分解为尺寸不同、纹理复杂度不同的子图像块.不同类图像块在空间上没有完全连通,基于像素间相关性的特征很难被使用,于是引入游程长度直方图的概念,可以在空间位置上不连续的区域之间提取特征.

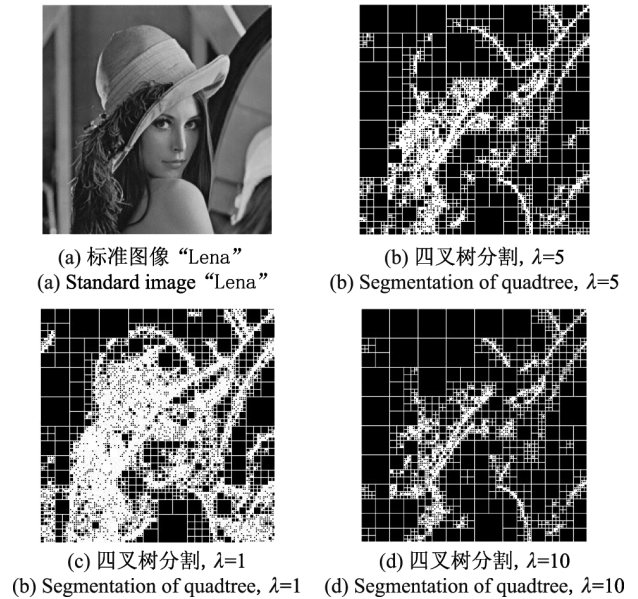


图1 图像四叉树分割

Figure 1 Example of quadtree segmentation

游程的概念是在20世纪50年代提出的.在数字图像中,沿某一确定方向 θ 具有相同灰度的一批邻接像素之集为一游程,集合中像素点的数目为游程长度.如果给定图像 I ,图像的灰度等级游程矩阵可以定义为

$$\begin{aligned} r(g, l|\theta) &= \#\{(x, y) | p(x, y) = p(x+s, y+t) = g; \\ &p(x+u, y+v) \neq g; \\ &0 \leq s < u \text{ \& } 0 \leq t < v; u = l \cos(\theta) \text{ \& } v = l \sin(\theta); \\ &0 \leq g \leq G-1 \text{ \& } 1 \leq l \leq L \text{ \& } 0^\circ \leq \theta \leq 180^\circ\} \end{aligned} \quad (3)$$

式中, $p(x, y)$ 代表了在 (x, y) 位置上的像素值;“#”代表了满足该条件的像素集合,即集合的势; G 和 L 分别为最大的灰度等级和最大的游程长度.因为游程矩阵数据量大,用它里面的元素当成特征计算太过复杂,所以把二维的矩阵转化成一维的游程长度直方图

$$H(l) = \sum_{g=1}^G r(g, l|\theta), \quad 1 \leq l \leq L \quad (4)$$

使用游程长度直方图作为隐写分析特征的主要原因游程长度的改变反映了秘密嵌入的过程.未嵌密的自然图像像素之间有较强的相关性,长游程出现的频次较高.隐写发生后,像素的灰度值会随机增加1或者减少1,于是长游程就变成若干个短游程,而若干短游程变成一个长游程的概率远小于前者.图2为隐写前后游程长度直方图的示意图,纵轴是 $\lg(H(l) + 1)$, 横轴是游程长度 l .从图2中可以看出,图像游程长度直方图嵌密后会缩短,长游程出现的频

次减小,这与分析结果相符.而从不同内容的图像的游程长度直方图可以看出,平坦图像载体载密区分度

比较高,而相比之下,纹理复杂的载体载密图像则很难区分出来,这与1.1节的分析相吻合.

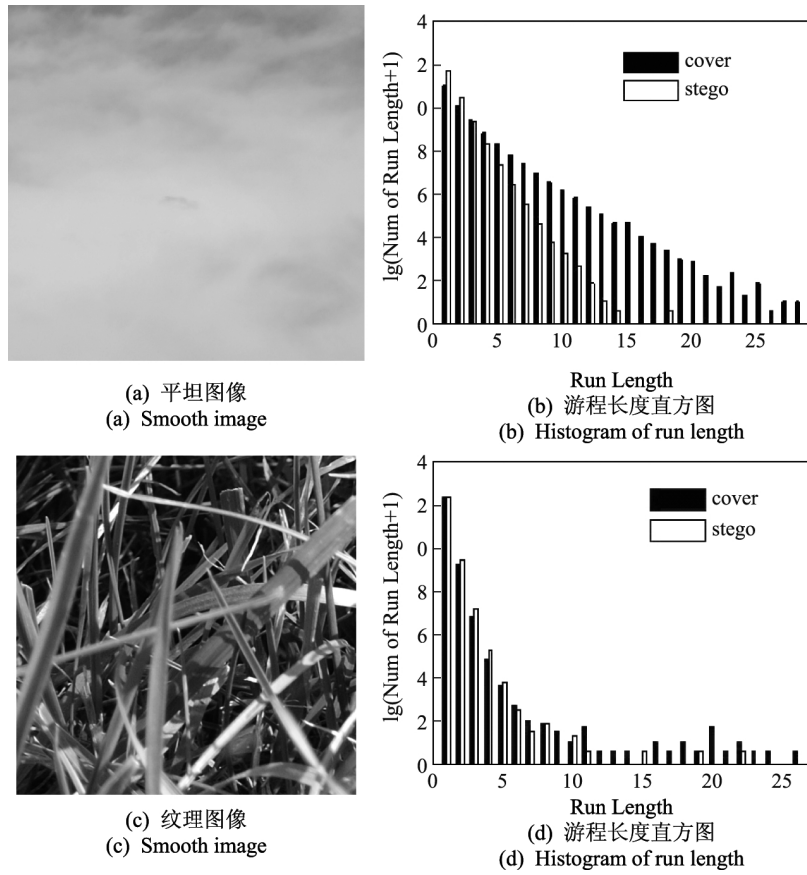


图2 平坦、纹理图像及其游程长度直方图

Figure 2 Examples of RLH of different kinds of images

1.4 特征提取

从图2中可以看出:秘密信息的嵌入使游程长度直方图向内收缩,载体载密直方图的差别主要体现在长游程的数目上,在较短游程数目上无法区别载体载密图像.直方图统计矩反映了纵轴上值在横轴上的累加.文献[12-13]引入的多阶统计矩特征也被证实对秘密信息的嵌入十分敏感.文献[14]提取了游程长度直方图质心特征

$$C(H(l)) = \frac{\sum_{l=1}^N lH(l)}{\sum_{l=1}^N H(l)} \quad (5)$$

利用二次嵌入方法进一步提高特征的可分性.文献[15]提取了游程长度直方图、差分图像游程长度直方图及压缩了原图像灰度取值范围的游程长度直方图

的频域特征矩 M_n

$$M_n = \frac{\sum_{k=0}^{\frac{K}{2}-1} |\Phi(k)|k^n}{\sum_{k=0}^{\frac{K}{2}-1} |\Phi(k)|}, \quad n \geq 1 \quad (6)$$

式中, $\Phi(k)$ 为 $H(l)$ 的离散傅里叶变换 (discrete Fourier transform, DFT), K 为 DFT 的长度, k 为傅氏变换的第 k 次频率. 本文提取游程长度直方图的前4阶空域统计矩作为隐写分析特征, 定义直方图空域 n 阶统计矩 m_n 为

$$m_n = \frac{\sum_{l=1}^N l^n H(l)}{\sum_{l=1}^N H(l)}, \quad n \geq 1 \quad (7)$$

式中, n 为矩的阶数, N 为直方图水平轴变量的数目, l 为直方图水平轴变量. 式(7)在 $n=1$ 时等价于

式(5),之所以采用前四阶矩是因为在实验过程中发现一味地增多特征矩的阶数并不会提高检测可靠性,当 $n=4$ 时可以达到最好的检测效果。

隐写分析实际上是一个模式识别问题,即按图像是否嵌入隐藏信息,将图像分为载体和载密两类.对于模式识别来说,好的特征应当是:类内聚合,类间分散.同样,对于隐写分析来说,好的特征应当是:1)对图像内容的不敏感,从各种载体图像中提取的特征相对聚合,且使用各种隐写方法获得的隐写图像中提取的特征也相对聚合;2)特征对于信息隐藏过程敏感,隐写前后特征相对分散.而实际情况是,图像内容千差万别,不同内容图像之间的差异大于载体载密图像之间的差异.为了达到更好的分类效果,在秘密信息均匀嵌入这一前提下,利用图像分割的手段把对隐写分析敏感的平坦区域提取出来.由于这些区域对秘密信息的嵌入较为敏感,并且对这些区域的隐写分析有较高的检测率,利用这些区域进行分析可以有效减少图像内容造成的影响.图像经过四叉树分解后形成不同尺度的子图像块,而块的大小在一定程度上反映了图像复杂度,此处设定 8×8 大小以下(包括 8×8)的块为纹理复杂区域, 8×8 以上的块为平坦区域.对每个平坦子图像块,计算游程长度直方图,最终的游程长度直方图为所有平坦图像块直方图之和,式(8)给出分割后游程长度直方图的表达式

$$H'(l) = \sum_{N=3}^9 H_N(l|B=2^N \times 2^N), \quad 1 < l < L \quad (8)$$

式中, B 表示最小图像块的大小.由于分割后的图像相对于原图丢失了一部分信息,各个游程长度对应的数值会有所下降,统计矩特征值整体上小于对原始图像游程长度直方图的统计矩特征值,但特征的区分度明显提高.定义某一隐写分析特征 f 的区分度函数 $D(f)$ 为

$$D(f) = \frac{f_c - f_s}{f_c} \quad (9)$$

式中, f_c 和 f_s 分别为嵌入前后的特征值.区分度函数值越大,说明隐写造成特征的改变比例越大,特征的可分性也越大.图3给出CAMERA图像库[16]中500幅图像游程长度直方图一阶矩的区分度函数分布,虚线为对整体图像提取特征的区分度函数分布,实线为对分割后敏感区域提取特征的区分度函数分布, D_1 是 $H'(l)$ 的特征值区分度函数分布, D_2 是 $H(l)$ 的区分度函数分布.由图4可以看出,利用本文算法对图像分割后再提取特征,特征的区分度更大。

最终提取图像四个方向上($0^\circ, 45^\circ, 90^\circ, 135^\circ$)的游程长度直方图 $H'(l)$ 的前四阶区域统计特征矩,共 $4 \times 4 = 16$ 维特征来进行分类器的训练和结果的判定。

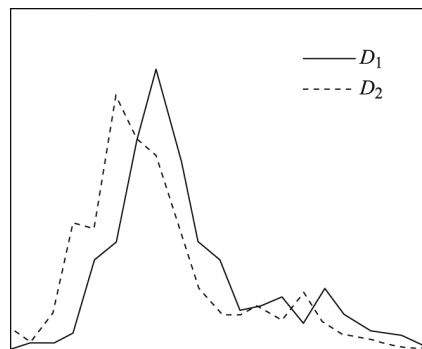


图3 $H(l)$ 和 $H'(l)$ 一阶矩的区分度函数分布

Figure 3 Differentiation of the first moment of $H(l)$ and $H'(l)$

1.5 分类器

支持向量机 (support vector machine, SVM) 是根据统计学习理论的结构风险最小化原则而提出的一种机器学习方法,不过分依赖于输入样本的数量,能够适应小样本集学习,具有很强的泛化能力,是现在广泛运用的一种分类器. LIBSVM^[17]是一个集成了分类、回归和分布估计的软件,提供了多种常用的核函数,可以有效解决多类问题、交叉验证选择参数等.对于本文的支持向量分类,选取径向基函数 (radial basis function, RBF) 作为核函数,用交叉验证选择最佳参数 C 与 g ,采用最佳参数 C 与 g 对整个训练集进行训练获取支持向量机模型,利用获取的模型进行分类判决。

2 实验结果分析

为了验证本文方法对不同隐写算法的检测性能,设置图像库如下:两个未压缩图像库 IS-1: CAMERA 库和 IS-2: BOWS2 库^[18]上, CAMERA 库包含 3164 幅图像,内容包括自然景观、建筑物和人物等,由不同型号数码相机获取,经过裁剪到分辨率为 512×512 ; BOWS2 库包含 10700 幅图像,经过缩放到 512×512 大小,从中随机选取 2000 幅图像作为本文试验图像库.对 IS-2 库进行质量因子为 75 的 JPEG 压缩,得到 IS-3.对 IS-1 和 IS-2 图像库分别采用 LSB replacing, LSB matching, stochastic modulation(SM)和一种边缘自适应嵌入算法(EAI)^[19],按 1.0 bpp、0.5 bpp、0.25 bpp 的嵌入率嵌入秘密信息.对 IS-3 图像库采用 Jsteg、JPHide、F5 嵌入算法分别嵌入 0.5 bpnc、0.2 bpnc 秘密信息.随机选取实验图像库的 30% 作为训练样本,其余 70% 作为测试。

RLHS 代表对整幅图像提取与本文相同游程直方图的特征,RLHS4+, RLHS8+, RLHS16+ 分别代表了提取最小图像块大小为 4×4 , 8×8 , 16×16 的游程

直方图特征. 与本文方法作对比的是同样使用游程直方图特征的文献 [14-15] 的方法, 另外与两种经典盲检测方法 OPT156^[12]和 CF78^[13]进行比较, 特征维数分别为 156 和 78.

详细的检测结果如表 2 和 3 所示, 通过实验可以得到以下结论:

- 1) 本文算法相比于未进行图像分割而直接在整体图像上提取特征的方法检测率更高;
- 2) 本文算法对非自适应隐写算法有一定的通用性, 对空域、频域不同嵌入算法较之前的几种隐写分析算法实现了检测率的提升;
- 3) 当提取最小图像块大小为 8×8 , 分割参数 λ 为 5 时, 检测率最高. 这是因为提取更小的图像块会引入更多纹理区域, 降低特征的区分度; 而仅提取较

大的图像块会降低数据量, 使判别结果不准确. 参数 λ 越小, 分割得越精细, 对于纹理复杂的图片, 能提取的平坦区域就相对较小; 较大的 λ 值会造成分解不足, 不能准确提取平坦区域. 以 CAMERA 图像库为例, 分割出的平坦区域占原图像的比例小于 20% 的有 146 幅, 而当提取最小图像块大小为 4×4 , 16×16 时, 这个数字分别为 28 和 642, 分割后可用于统计的图像尺寸也会对检测结果产生影响.

4) 对 JPEG 类图像隐写算法检测效果并不理想, 因为 JPEG 类隐写算法对数据的改变发生在频域. 文献 [20] 提出: 数字隐写发生在哪个域, 就在哪个域提取特征进行隐写分析, 这样较为直接有效. 本文提取的空域游程直方图特征对频域隐写检测能力有限.

表 2 最小子图像块尺寸和分割参数 λ 对算法性能的影响

Table 2 Algorithm effect performed by minimum of image block size and segmentation parameter λ											%
	IS-1				IS-2			IS-3			
	E-R/bpp	LSBR	LSBM	SM	LSBR	LSBM	SM	E-R/bpnc	Jsteg	JPHide	F5
RLHS, $\lambda = 5$	1.0	95.15	91.92	97.65	96.62	93.39	99.12				
	0.5	86.27	78.90	89.43	87.74	80.37	80.37	0.5	77.47	78.46	72.27
	0.25	78.26	65.71	80.93	79.73	67.18	82.40	0.2	68.27	69.93	60.96
RLHS4+, $\lambda = 5$	1.0	98.75	95.41	99.45	100	96.88	99.99				
	0.5	90.05	82.38	93.55	91.52	83.85	95.02	0.5	81.83	82.66	75.25
	0.25	82.87	71.55	85.27	84.34	73.02	86.74	0.2	72.54	73.56	63.06
RLHS8+, $\lambda = 5$	1.0	98.87	95.86	99.47	100	97.33	100				
	0.5	90.65	82.43	93.67	92.12	83.90	95.14	0.5	81.88	82.98	75.49
	0.25	83.34	71.69	85.37	84.81	73.16	86.84	0.2	72.58	73.77	63.27
RLHS16+	1.0	97.75	94.22	98.28	99.22	95.69	99.75				
	0.5	89.33	80.05	91.39	90.80	81.52	92.86	0.5	81.51	82.51	74.38
	0.25	78.50	69.34	80.82	79.97	70.81	82.29	0.2	71.72	71.73	62.34
RLHS8+, $\lambda = 1$	1.0	98.21	95.37	99.01	99.59	96.83	99.75	0.5	81.51	82.77	75.11
	0.5	89.98	81.95	93.30	91.74	83.59	94.76	0.2	72.33	73.34	62.83
RLHS8+, $\lambda = 4$	1.0	98.43	95.55	99.18	100	97.07	100	0.5	81.83	82.80	75.06
	0.5	90.22	82.07	93.39	91.98	83.78	95.01	0.2	72.49	73.69	63.01
RLHS8+, $\lambda = 6$	1.0	98.44	95.58	99.42	100	97.27	100	0.5	81.86	82.81	75.10
	0.5	90.22	82.11	93.55	92.10	83.85	94.89	0.2	72.55	73.72	63.23
RLHS8+, $\lambda = 10$	1.0	97.34	94.30	98.20	98.64	95.96	98.20	0.5	80.07	81.79	73.28
	0.5	88.45	81.84	92.43	90.87	82.27	93.04	0.2	71.89	72.05	61.91

表 3 不同算法的检测性能比较

Table 3 Comparison of detection performance among different algorithms							%
I-S	S-A	E-R	CF78	OPT156	Method in [14]	Method in [15]	Proposed
IS-1	LSBR	1.0 bpp	97.93	96.58	97.20	94.37	98.87
		0.5 bpp	90.55	89.73	89.36	84.25	90.65
		0.25 bpp	82.29	82.10	82.93	74.29	83.34
		1.0 bpp	94.56	93.44	93.42	89.95	95.86
		0.5 bpp	81.94	81.51	80.79	78.73	82.43
		0.25 bpp	70.21	71.01	69.68	65.59	71.69
	LSBM	1.0 bpp	99.91	99.73	98.90	94.37	99.47
		0.5 bpp	94.26	93.93	92.28	84.25	93.67
		0.25 bpp	85.47	85.28	84.99	74.29	85.37
		1.0 bpp	99.40	98.05	98.67	95.84	100
		0.5 bpp	92.02	91.20	90.83	85.72	92.12
		0.25 bpp	83.76	83.57	84.40	75.76	84.81
IS-2	LSBR	1.0 bpp	96.03	94.91	94.89	91.42	97.33
		0.5 bpp	83.41	82.98	82.26	80.20	83.90
		0.25 bpp	71.68	72.48	71.15	67.06	73.16
		1.0 bpp	100	99.96	99.50	95.84	100%
		0.5 bpp	95.73	95.40	93.75	85.72	95.14
		0.25 bpp	86.94	86.75	86.46	75.76	86.84
	Jsteg	0.5 bpnc	82.41	81.63	81.66	74.93	81.88
		0.2 bpnc	73.28	72.38	72.33	66.28	72.58
		0.5 bpnc	84.36	82.13	81.67	76.73	82.98
		0.2 bpnc	73.70	73.18	72.90	68.99	73.77
		0.5 bpnc	75.61	74.63	74.69	73.71	75.49
		0.2 bpnc	63.35	62.58	62.37	61.91	63.27

3 结 语

本文结合图像内容特征分析,研究了嵌入信息前后隐写分析特征的改变与图像内容之间的关系,提出了基于图像内容的隐写分析方法.该方法充分考虑了图像内容对隐写分析的影响,应用二叉树分割的方法,提取出对隐写分析特征较敏感的区域,增加了特征的区分度.实验表明本方法的检测性能相比未考虑图像内容的对应隐写分析算法有所提高.下一步的工作可以从以下方面开展:1)在特征选取和优化上需要进一步研究;2)本文未对纹理复杂区域作处理,因此对自适应的隐写算法检测效果一般,如何实现不同区域分类判别和结果加权融合是下一步研究的重点.

参考文献:

[1] KER A. Steganalysis of LSB matching in grayscale images [J]. IEEE Signal Processing Letters, 2005, 12(6): 441-444.

[2] CANCELLI G, DOËRR G, BARNI M J, COX I J. A comparative study of ± 1 steganalyzers [C]//IEEE Workshop on Multimedia Signal Processing, 2008: 791-796.

[3] 郭艳卿,何德全,尤新刚,孔祥维,王波.基于整体性的隐密分析特征提取和融合方法研究[J].电子学报,2006,34(12A):2443-2446.

GUO Yanqing, HE Dequan, YOU Xingang, KONG Xi-angwei, WANG Bo. Research on holism-based feature extraction and fusion for steganalysis [J]. Acta Electronica Sinica, 2006, 34 (12A): 2443-2446.

- [4] 钮心忻, 杨义先. 信息隐写与隐写分析研究框架探讨[J]. 电子学报, 2006, 34(12A): 2421-2424.
NIU Xinxin, YANG Yixian. Study on the frame of information steganography and steganalysis [J]. Acta Electronica Sinica, 2006, 34(12A): 2421-2424.
- [5] LIU Q, SUNG A H, RIBEIRO B, WEI M, CHEN Z, XU J. Image complexity and feature mining for steganalysis of least significant bit matching steganography [J]. Information Sciences, 2008, 178(1): 21-36.
- [6] CHO S, CHA B H, WANG J W, JAY K. Block-based image steganalysis: algorithm and performance evaluation [C]//IEEE International Symposium on Circuits and Systems, 2010: 1679-1682.
- [7] WONG P W, CHEN Hong, TANG Zhongjue. On steganalysis of plus-minus one embedding of continuous tone images [C]//Proceedings of SPIE-IS&T Electronic Imaging, 2005, 5681: 643-652.
- [8] HARMSSEN J, PEARLMAN W A. Steganalysis of additive noise modelable information hiding [C]//SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, 2003: 131-142.
- [9] ZHANG J, COX I J, DOËRR G. Steganalysis for LSB matching in images with high-frequency noise [C]//IEEE Workshop on Multimedia Signal Processing, 2007: 385-388.
- [10] FRIDRICH J, KODOVSKÝ J, HOLUB V, GOLJAN M. Steganalysis of content-adaptive steganography in spatial domain [C]//Information Hiding 13th International Conference, 2011, 6958: 102-117.
- [11] JACOBS E W, FISHER Y, BOSS R D. Image compression: a study of the iterated transformation method [J]. Signal Process, 1992, 29.
- [12] WANG Y, PIERRE M. Optimized feature extraction for learning-based image steganalysis [J]. IEEE Transaction on Information Forensics and Security, 2007, 2(1): 31-45.
- [13] 高建炯, 宣国荣, 施云庆, 柴佩琪, 杨志强. 基于直方图频域统计矩的图像隐写分析 [J]. 计算机应用, 2005, 25(10): 2371-2373.
GAO Jianjiong, XUAN Guorong, SHI Yunqing, CAI Peiqi, YANG Zhiqiang. Image steganalysis based on statistical moments in the frequency domain of the histogram [J]. Computer Applications, 2005, 25(10): 2371-2373. (in Chinese)
- [14] DONG Jing, TAN Tieniu. Blind image steganalysis based on run-length histogram analysis [C]//IEEE International Conference on Image Processing, 2008: 2064-2067.
- [15] YU X Y, BABAGUHCI N. Run length based steganalysis for LSB matching steganography [C]//International Conference on Multimedia and Expo, 2008: 353-356.
- [16] DOËRR G. Image database for steganalysis studies [DB/OL]. Available: <http://www.cs.ucl.ac.uk/staff/I.Cox/Content/Downloads.html>.
- [17] CHANG C C, LIN C J. LIBSVM: a library for support vector machines [OL]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [18] BAS P, FURON T. BOWS-2 Jul. 2007 [DB/OL]. Available: <http://bows2.gipsa-lab.inpg.fr>
- [19] LUO Weiqi, HUANG Fangjun, HUANG Jiwu. Edge adaptive image steganography based on LSB matching revisited [J]. IEEE Transactions Information Forensics and Security, 2010, 5(2): 201-214.
- [20] FRIDRICH J. Feature-based steganalysis for jpeg images and its implications for future design of steganographic schemes [C]//Proceedings of 6th Information Hiding Workshop. 2004. Springer LNCS, 3200: 67-81.

(编辑: 管玉娟)