

彩色图像通用隐写分析的多类统计特征

孙文颢 刘婷婷 张新鹏 王朔中

(上海大学通信与信息工程学院, 上海 200072)

摘要 提出了用于彩色图像通用隐写分析的3类统计特征。针对隐写所导致的图像DCT系数分布、空域像素值平坦性以及颜色空间一致性的改变,从系数分布模型背离程度、空域相关性和不同颜色分量间的相关性3个方面设计能够体现原始载体和含密图像差异的特征,提取了10维特征向量。这些特征对隐写行为较为敏感、涵盖面广,且维数低、计算方便。在这些特征的基础上,可利用支持向量机分析彩色JPEG图像是否含有秘密信息。实验结果表明,基于这些特征的通用隐写分析方法可以准确、高效地检测以Jsteg、F5或MB隐写方法嵌入的秘密信息,辨别隐写方法种类,具有很好的通用性。

关键词 通用隐写分析 JPEG图像 特征向量 支持向量机

中图法分类号:TP309 文献标识码:A 文章编号:1006-8961(2008)10-1914-04

Statistical Features for Universal Steganalysis on Color Images

SUN Wen-yong, LIU Ting-ting, ZHANG Xin-peng, WANG Shuo-zhong

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

Abstract In this paper we propose three types of features which can be used to realize efficient universal steganalysis on color images. After measuring the deviation of DCT coefficient distribution, the correlation of smoothness in spatial domain, and the correlation between different components, a 10-dimensional feature vector is produced for each color image. Those features are sensitive to steganographic process and the feature vector, having a low dimension, is easy to calculate. Then, a SVM classifier is built to distinguish the stego-images from the original images. Experiments show that the universal steganalytic approach can effectively detect the presence of secret message hidden by Jsteg, F5 or MB steganography, and can identify the steganographic techniques used.

Keywords universal steganalysis, JPEG images, feature vector, support vector machine (SVM)

1 引言

隐写分析作为主要的反隐写技术,其目的是检测载体中秘密信息的存在性,一旦发现秘密信息存在,便可中止隐蔽通信并缉捕发送者。与针对特定隐写方法的隐写分析不同,通用隐写分析可用于分析多种隐写方法,通过特征训练建立分类模型,从而实现对待检测图像的有效检测,判断其中是否存在秘密信息。其关键在于统计特征的提取并根据特征

向量进行分类。隐写分析用到的特征通常是对隐写操作敏感而对其他正常操作不敏感的特征。现有的通用分析算法提取的特征主要有基于图像质量的特征^[1]、小波域统计特征^[2]以及DCT系数分布特点和图像块边界特征^[3]等。

JPEG图像是互联网中常见的图像格式,以往方法多针对灰度JPEG图像进行隐写分析。另一方面,JPEG图像中的隐写操作多在DCT域进行,而这种行为也会破坏图像的空域相关性。本文同时从变换域和空域提取对隐写行为敏感的特征,包括DCT

基金项目:国家自然科学基金项目(60502039, 60773079);国家863计划(2007AA01Z477);上海市青年科技启明星计划项目(06QA14022)

收稿日期:2008-07-11;改回日期:2008-08-05

第一作者简介:孙文颢(1984~),女,上海大学通信与信息工程学院硕士研究生。主要研究方向为信息安全,图像处理。

E-mail:sunwenyong@smg.sh.cn

系数分布模型背离程度、空域平坦相关性,并且提取体现不同颜色内容之间相关性的平坦性差异,突破了现有方法把不同颜色分量独立处理的局限,提高了分析的准确性和效率。

2 基于系数分布模型和相关性的特征

应用 3 类统计特征:系数分布模型背离度、空域平坦相关性、色度分量平坦性差异。系数分布模型背离度可以在不构造参考图像情况下比较图像 DCT 系数分布模型与实际系数分布之间的差异;空域平坦相关性可以体现隐写行为对空域平坦性的破坏程度;色度分量平坦性差异描述了不同色度分量中图像子块边缘和内部的平坦性关系。这些特征向量计算简单、维数较低,可同时保证分析准确率和运算效率。

2.1 系数分布模型背离程度

DCT 系数直方图可以用来表征 DCT 系数的分布特性。Jsteg^[4]、F5^[5]、MB^[6] 等隐写方法会不同程度地改变 DCT 系数直方图分布。统计不同频率位置上的 DCT 系数直方图,并估计 DCT 系数的分布模型,考察 DCT 系数的实际分布与分布模型之间的差异,提取两个参数来衡量系数分布背离程度,用于隐写分析。提取步骤如下:

(1) 将待检测图像变换到 YCbCr 空间,对 Y 分量中的 63 个非直流系数位置分别统计 DCT 系数直方图,用 $h_{i,j}(a)$ 表示 (i, j) 位置上值为 a 的 DCT 系数个数。

(2) 把除 $h_{i,j}(0)$ 以外的 $h_{i,j}(a)$ 两两组对,求每对的系数个数和

$$x_{i,j}(k) = \begin{cases} h_{i,j}(2k-1) + h_{i,j}(2k) & k > 0 \\ h_{i,j}(k) & k = 0 \\ h_{i,j}(2k+1) + h_{i,j}(2k) & k < 0 \end{cases} \quad (1)$$

(3) 用广义柯西分布函数

$$L(k) = \frac{p-1}{2s} \left(\left| \frac{k}{s} \right| + 1 \right)^{-p} \quad (2)$$

拟合 $x_{i,j}(k)$,刻画图像 DCT 系数直方图分布特性,选择参数 p 和 s 使 $L(k)$ 与 $x_{i,j}(k)$ 的差值平方和最小。根据 $L(k)$ 得到 DCT 系数分布模型为

$$h'_{i,j}(a) = \frac{p-1}{2s} \left(\left| \frac{a+0.5}{s} \right| + 1 \right)^{-p} - \frac{p-1}{2s} \left(\left| \frac{a-0.5}{s} \right| + 1 \right)^{-p} \quad (3)$$

用于表示 (i, j) 位置上 DCT 系数模型中值为 a 的 DCT 系数个数。含密图像的 $h'_{i,j}(a)$ 与实际图像直方图分布通常有较大差异。

(4) 令 N 为待检测图像 8×8 小块的总数,为减

少 DCT 系数的随机性对特征向量的影响,对满足 $x_{i,j}(k) > N/20$ 的 k ,衡量 (i, j) 位置上 $h_{i,j}(a)$ 与 $h'_{i,j}(a)$ 的差异,以 $D_{i,j}(k)$ 衡量实际系数分布与模型 $h'_{i,j}(a)$ 的背离程度, $C_{i,j}(k)$ 标记此背离程度所属类别。 $D_{i,j}(k)$ 和 $C_{i,j}(k)$ 定义如下:当 $k > 0$ 时,

$$D_{i,j}(k) = \begin{cases} h_{i,j}(2k-1) - \frac{h_{i,j}(2k) \times h'_{i,j}(2k-1)}{h'_{i,j}(2k)} \\ \left(h_{i,j}(2k-1) \geq \frac{h_{i,j}(2k) \times h'_{i,j}(2k-1)}{h'_{i,j}(2k)} \right) \\ h_{i,j}(2k) - \frac{h_{i,j}(2k-1) \times h'_{i,j}(2k)}{h'_{i,j}(2k-1)} \\ \left(h_{i,j}(2k-1) < \frac{h_{i,j}(2k) \times h'_{i,j}(2k-1)}{h'_{i,j}(2k)} \right) \end{cases}$$

$$C_{i,j}(k) = \begin{cases} 0 & h_{i,j}(2k-1) \geq \frac{h_{i,j}(2k) \times h'_{i,j}(2k-1)}{h'_{i,j}(2k)} \\ 1 & h_{i,j}(2k-1) < \frac{h_{i,j}(2k) \times h'_{i,j}(2k-1)}{h'_{i,j}(2k)} \end{cases} \quad (4)$$

当 $k < 0$ 时,

$$D_{i,j}(k) = \begin{cases} h_{i,j}(2k+1) - \frac{h_{i,j}(2k) \times h'_{i,j}(2k+1)}{h'_{i,j}(2k)} \\ \left(h_{i,j}(2k+1) \geq \frac{h_{i,j}(2k) \times h'_{i,j}(2k+1)}{h'_{i,j}(2k)} \right) \\ h_{i,j}(2k) - \frac{h_{i,j}(2k+1) \times h'_{i,j}(2k)}{h'_{i,j}(2k+1)} \\ \left(h_{i,j}(2k+1) < \frac{h_{i,j}(2k) \times h'_{i,j}(2k+1)}{h'_{i,j}(2k)} \right) \end{cases}$$

$$C_{i,j}(k) = \begin{cases} 0 & h_{i,j}(2k+1) \geq \frac{h_{i,j}(2k) \times h'_{i,j}(2k+1)}{h'_{i,j}(2k)} \\ 1 & h_{i,j}(2k+1) < \frac{h_{i,j}(2k) \times h'_{i,j}(2k+1)}{h'_{i,j}(2k)} \end{cases} \quad (5)$$

记 $D_{i,j}(k)$ 和 $C_{i,j}(k)$ 的均值为 F_1 、 F_2 ,用来衡量 DCT 系数的实际分布与模型的背离度。 F_1 表明了图像实际系数分布与模型的差异度,原始载体的 F_1 应接近于 0,而含密图像的 F_1 应较大; F_2 则描述了差异的分类情况,原始载体的 F_2 应接近于 0.5,而含密图像的 F_2 值远大于 0.5 或接近 0。 F_1 和 F_2 即用于隐写分析的前二个特征。

2.2 空域平坦相关性

隐写会削弱图像子块边界的相关性和子块内部的相关性,使子块边界或内部变得不平滑。按照 JPEG 图像的 8×8 分块规则,考察图像子块边缘平坦程度与子块内部平坦程度的差异,以此作为衡量

子块边缘相关性的特征向量,并用奇异值分析的方法衡量图像子块的平坦情况,以此作为衡量子块内部相关性的特征向量。提取步骤如下:

(1)将 JPEG 图像从 DCT 系数转换为像素值,设图像大小为 $H \times L$,用 $P_s(i,j)$ ($s \in \{Y,Cb,Cr\}$),

$$D_s(u,v) = \frac{\sum_{i=1}^8 \left\{ \left| [P_s(8u,8v+i) - P_s(8u+1,8v+i)] \right| + \left| [P_s(8u+i,8v) - P_s(8u+i,8v+1)] \right| \right\}}{\sum_{i=1}^8 \left\{ \left| [P_s(8u+4,8v+i) - P_s(8u+5,8v+i)] \right| + \left| [P_s(8u+i,8v+4) - P_s(8u+i,8v+5)] \right| \right\}}$$

$$1 < u < \frac{H}{8}, 1 < v < \frac{L}{8}, s \in \{Y,Cb,Cr\} \quad (6)$$

根据式(6),统计亮度分量 Y 中 $D_Y(u,v)$ 的均值、方差,并统计色度分量 Cb 和 Cr 中 $D_{Cb}(u,v)$ 、 $D_{Cr}(u,v)$ 的均值,得 3 个统计特征

$$F_3 = \frac{\sum_{u=2}^{H/8-1} \sum_{v=2}^{L/8-1} D_Y(u,v)}{\left(\frac{H}{8} - 2\right) \times \left(\frac{L}{8} - 2\right)} \quad (7)$$

$$F_4 = \frac{\sum_{u=2}^{H/8-1} \sum_{v=2}^{L/8-1} [D_Y(u,v) - F_3]^2}{\left(\frac{H}{8} - 2\right) \times \left(\frac{L}{8} - 2\right)} \quad (8)$$

$$F_5 = \frac{\sum_{u=2}^{H/8-1} \sum_{v=2}^{L/8-1} [D_{Cb}(u,v) + D_{Cr}(u,v)]}{2 \times \left(\frac{H}{8} - 2\right) \times \left(\frac{L}{8} - 2\right)} \quad (9)$$

F_3 、 F_4 从亮度, F_5 从色度上衡量了子块边缘平坦程度与内部平坦程度的差异。隐写破坏了空域平坦性,增大了这种差异,使 F_3 、 F_4 、 F_5 的值较大;而原始载体的 F_3 、 F_4 、 F_5 的值应较小。

(3)用 $B_s(u,v)$ ($1 \leq u \leq H/8, 1 \leq v \leq L/8$) 表示图像在 (u,v) 位置的子块,对图像每个子块 $B_s(u,v)$ 进行奇异值分解,用 $SVD_{s,i}(u,v)$ ($i = 1, 2, \dots, 8$) 表示图像 s 分量上 (u,v) 位置子块第 i 位奇异值,并用不同位置上的奇异值的比值,来衡量图像子块内部的平滑情况。

$$C_{s,i,j}(u,v) = \frac{SVD_{s,i}(u,v)}{SVD_{s,j}(u,v)} \quad (10)$$

取 Y 分量中 $C_{Y,5,1}$ 、 $C_{Y,8,1}$ 的均值,得第 6、7 维特征

$$F_6 = \frac{\sum_{u=1}^{H/8} \sum_{v=1}^{L/8} C_{Y,5,1}(u,v)}{\frac{H}{8} \times \frac{L}{8}} \quad (11)$$

$$F_7 = \frac{\sum_{u=1}^{H/8} \sum_{v=1}^{L/8} C_{Y,8,1}(u,v)}{\frac{H}{8} \times \frac{L}{8}} \quad (12)$$

$1 \leq i \leq H, 1 \leq j \leq L$ 来表示在 s 分量 (i,j) 位置上的像素值。

(2)为便于计算,忽略图像最外围 8×8 子块,统计其他 8×8 子块边缘相邻像素差值与该子块内部 4×4 子块边缘相邻像素差值,并取两者的比值,

取 Cb、Cr 分量中 $C_{Cb,5,1}$ 与 $C_{Cr,5,1}$ 两者和的均值,得第 8 维特征

$$F_8 = \frac{\sum_{u=1}^{H/8} \sum_{v=1}^{L/8} [C_{Cb,5,1}(u,v) + C_{Cr,5,1}(u,v)]}{2 \times \frac{H}{8} \times \frac{L}{8}} \quad (13)$$

F_6 、 F_7 从亮度, F_8 从色度上衡量了图像子块内部的平滑情况。原始载体的平坦性较好, F_6 、 F_7 、 F_8 值应较小;而含密图像的 F_6 、 F_7 、 F_8 值应较大。

2.3 色度分量平坦性差异

对于彩色图像,隐写会在某个颜色空间的各分量上造成不同程度的影响,破坏图像原有的平坦一致性特点。彩色图像在 YCbCr 空间中,Cb 和 Cr 分量有较好的平坦一致性,故考虑衡量这两个色度分量之间的图像子块的相关性。统计 Cb 和 Cr 分量里图像子块边缘的平坦情况,并按平滑程度顺序编号,比较不同色度分量里编号值之间的距离,以此作为不同色度分子块边缘的特征向量,若距离较远,图像就可能被隐写。用奇异值分析的方法衡量图像子块的平坦情况,并在各色度分量对该值排序,比较不同色度分量里排序值之间的距离,以此作为不同色度分子块内部的特征向量。提取步骤如下:

(1)统计色度分量 Cb 和 Cr 的相关性,对图像用式(6)衡量 Cb 和 Cr 分子块边缘平滑情况,得色度分子块边缘平滑特性矩阵 D_{Cb} 和 D_{Cr} 。分别将 D_{Cb} 和 D_{Cr} 中的元素 $D_s(u,v)$ 按照取值从小到大排序,得到相应的序号矩阵 PD_{Cb} 和 PD_{Cr} ,计算两个序号矩阵的距离,得第 9 维特征

$$F_9 = \frac{\sum_{u=2}^{H/8-1} \sum_{v=2}^{L/8-1} |PD_{Cb}(u,v) - PD_{Cr}(u,v)|}{\left(\frac{H}{8} - 1\right) \times \left(\frac{L}{8} - 1\right)} \quad (14)$$

(2)用式(10)衡量子块内部平滑情况,得色度分子块内部平滑特性矩阵 $C_{Cb,5,1}$ 和 $C_{Cr,5,1}$ 。分别

将 $C_{Cb,s,1}$ 和 $C_{Cr,s,1}$ 中的元素 $C_{s,s,1}(u,v)$ 按照取值从小到大排序,得到相应的序号矩阵 PC_{Cb} 和 PC_{Cr} ,计算两个序号矩阵的距离,得第 10 维特征

$$F_{10} = \frac{\sum_{u=1}^{H/8} \sum_{v=1}^{L/8} |PC_{Cb}(u,v) - PC_{Cr}(u,v)|}{\frac{H}{8} \times \frac{L}{8}} \quad (15)$$

F_9 、 F_{10} 分别从边缘与内部平坦性差别及图像内部平坦性衡量了两个色度分量间的差异,原始载体的 F_9 、 F_{10} 应较小;含密图像的 F_9 、 F_{10} 应较大。

3 实验结果

对以上提取的 10 维特征向量,采用支持向量机^[7]分类方法进行识别。

训练样本集由 1 000 幅不同嵌入率的 Jsteg 含密图像和 1 000 幅原始载体构成,提取特征向量后利用支持向量机软件包 LIBSVM^[8]进行训练,得到两类分类器。然后由 1 000 幅不同嵌入率的 Jsteg、F5、MB 含密图像和原始载体构成测试样本集,将训练得到的分类器作为工具分析测试样本,结果如表 1 所示,可以比较准确地判断出图像是否经过 Jsteg、F5、MB 隐写。

表 1 不同嵌入率下 Jsteg、F5、MB 隐写分析结果(正确检测概率)
Tab.1 Steganalysis result against Jsteg F5/MB with various embedding rates(correct probability)

	单位:%					
	嵌入量					
	0	60	70	80	90	100
Jsteg	95.1	93.3	94.6	94.8	95.8	96.4
F5	92.5	88.6	89.9	91.6	92.3	93.4
MB	88.2	85.7	87.1	91.4	92.5	92.6

对 1 000 幅 JPEG 图像以不同嵌入率分别进行 Jsteg、F5 和 MB 隐写,用这 3 000 幅含密图像和 1 000 幅原始载体构成训练样本集,提取特征向量,训练 LIBSVM 得到多类分类器,用其作为工具分析由不同嵌入率的 Jsteg、F5 和 MB 含密图像各 1 000 幅构成测试样本,结果如表 2 所示。可见使用本文方法所提取的特征可以比较准确地判断出含密图像所使用的隐写方法。

4 结论

从 DCT 系数分布背离程度、空域相关性、不同颜色内容之间相关性三方面提出应用于彩色 JPEG 图像通用隐写分析的统计特征,从中提取 10 维特征向量,

用其训练的分类器能有效地检测 Jsteg、F5 和 MB 隐写,而且这些特征的选取大大减少了特征向量的维数,较之以往方法降低了运算复杂度,改善了运算效率。

表 2 不同嵌入率下 Jsteg、F5、MB 方法的通用隐写分析结果
Tab.2 Results of universal steganalysis against Jsteg, F5 and MB

隐写方法	嵌入率	载体图像	单位:%		
			Jsteg	F5	MB
Jsteg	100	2.7	91.9	1.7	3.7
	90	2.9	90.7	2	4.4
	80	3.0	90.4	2.1	4.5
	70	4	89.1	2.7	4.2
F5	100	6.4	0.2	92.3	1.1
	90	7.8	0.2	91.2	0.8
	80	6.8	0.9	90.9	1.4
	70	9.9	0.8	86.8	2.5
MB	100	6.3	3.3	3.3	87.1
	90	4.3	5.7	3.1	86.9
	80	5.7	7.1	3.1	84.1
	70	4.6	9.5	3.3	82.6
原始载体	0	88.9	1.3	6.5	3.3

本文是针对现有主流隐写方法进行的通用隐写分析,随着隐写技术的进一步发展,隐写分析将面临新的挑战。如何找到更具普遍意义的能够描述载体图像与含密图像差异的特征,提高通用隐写分析的普适性将成为隐写分析努力的方向。

参考文献(References)

- 1 Avciabas I, Memon N, Sankur B. Steganalysis using image quality metrics [J]. IEEE Transactions on Image Processing, 2003, 12(2): 221 ~ 229.
- 2 Lyu S, Farid H. Steganalysis using higher-order image statistics [J]. IEEE Transactions on Information Forensics and Security, 2006, 1(1): 111 ~ 119.
- 3 Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes [A]. In: Proceedings of 6th Information Hiding Workshop [C], Toronto, Canada, 2004, 3200: 67 ~ 81.
- 4 Hsu C T, Wu J L. Hidden digital watermarks in images [J]. IEEE Transactions on Image Processing, 1999, 8(1): 58 ~ 68.
- 5 Westfeld A. F5-A steganographic algorithm [A]. In: Proceedings of 4th International Workshop on Information Hiding [C], Pittsburgh, Pennsylvania, 2001, 2137: 289 ~ 302.
- 6 Sallee P. Model-based steganography [A]. In: Proceedings of International Workshop on Digital Watermarking [C], Seoul, Korea, 2004, 2939: 154 ~ 167.
- 7 Miche Y, Roue B, Lendasse A, et al. A feature selection methodology for steganalysis [A]. In: Multimedia Content Representation, Classification and Security, International Workshop [C], Istanbul, Turkey, 2006: 4105: 49 ~ 56.
- 8 Chang C-C, Lin C-J. LIBSVM—A Library for Support Vector Machines [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/cgi-bin/matlab.cgi?> + <http://www.csie.ntu.edu.tw/~cjlin/libsvm/matlab+zip,2008-04-02/2008-06-07>.