文章编号:1001 - 9081(2007)09 - 2140 - 03

基于马尔可夫模型和支持向量机的 JPEG 图像隐写分析

崔 霞,童学锋,黄 聪
(同济大学 计算机科学与技术系,上海 200092)
(cx_university@126.com)

摘 要:提出了一种新的针对 JPEG 图像的通用隐写分析方法,利用马尔可夫模型,挖掘量化后的分块 DCT 系数中低频系数的相关性,提取出 360 维特征,然后采用支持向量机(SVM)分类方法进行识别。对四种公认的安全性较高的 JPEG 嵌入方法 F5、Outguess、MB1 和 MB2 进行隐写分析。在 CorelDraw 图像库的实验结果显示:该方法的检出率高、稳定性好且运算速度快。

关键词:隐写分析;JPEG 图像;马尔可夫模型;支持向量机
中图分类号: TP391.41 **文献标志码**:A

Steganalysis based on Markov model and SVM for JPEG images

CUI Xia, TONG Xue-feng, HUANG Cong
(*Department of Computer Science and Technology, Tongji University, Shanghai* 200092, *China*)

Abstract: A new and universal steganalysis scheme based on Markov model for JPEG images was proposed. 360 dimensional feature vectors were derived from the markov matrix, which was calculated directly in DCT domain and sensitive to the data embedding process. Then Support Vector Machine (SVM) was used to classify. The experimental results of 4 popular JPEG steganographic schemes (F5, Outguess, MB1 and MB2) demonstrate that the proposed scheme has the advantage in detection rate, stability, and speed.

Key words: steganalysis; JPEG image; Markov model; Support Vector Machine(SVM)

0 引言

隐写分析技术是数据隐藏的反过程,目的是为了揭示数字媒体中隐藏信息的存在。它一方面可以促进信息伪装算法安全性的提高,也可以作为评判隐写技术、数字水印或数据隐藏的质量好坏的一个指标;另一方面从有隐藏信息的载体中发现隐藏信息,可以直接用于侦察,可以防止信息伪装被不法分子滥用。目前的隐写分析技术可以分为通用隐写分析和针对特定数据隐藏方法的隐写分析。本文讨论的是通用隐写分析。

文献[1]提出了从小波直方图提取 72 维高阶统计量作为特征的隐写分析方法,取得较好的结果。文献[2]提出了从小波直方图 DFT 提取 39 维高阶统计量作为特征的隐写分析方法,有所改进。以上两种方法都是用一阶统计量(直方图)作为特征的隐写分析方法,用于 JPEG 图像时却不合适,因为 JPEG 图像量化后的 DCT 块能量都集中在低频,其他大部分都是零,可利用的特征太少。文献[3]提出了一种针对扩谱隐写的隐写分析方法,通过求取共生矩阵来挖掘相邻像素间的相关性,处理 JPEG 图像时需要转到空域中进行,而且由于空域灰度范围太大造成维数太高,作者仅保留共生矩阵对角线上及其附近的系数来降低维数,这两个过程都造成大量的信息丢失。文献[4]利用二阶统计量总结了 23 个对于数据隐写比较敏感的特征,结果有所改善,但特征偏少,运算量却很大。

本文直接利用量化后的 DCT 系数来挖掘信息,避免了转到空域处理而丢失信息,而且维数不会太高。实验表明该方法在通用性以及小嵌入量隐写分析上均取得很好的性能。

1 特征提取

本文首先采用多种方式扫描 DCT 块的中低频系数,形成相应的扫描序列,然后根据所选阈值处理这些系数,再利用马尔可夫模型计算出各序列的马尔可夫矩阵,形成 360 维特征。

图 1 图像 16053.jpg

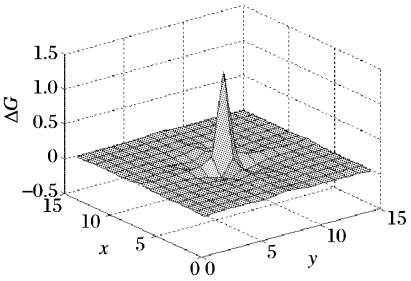

图 2 马尔可夫矩阵之差($\Delta \boldsymbol{G} = \boldsymbol{G}_{steg} - \boldsymbol{G}_{ori}$)

用 $\boldsymbol{G}_{ori}$ 表示原图的马尔可夫矩阵,$\boldsymbol{G}_{steg}$ 表示嵌入数据后该图的马尔可夫矩阵,对于 CorelDraw 库中编号为 16053 的图像(如图 1),利用 *F5* 进行数据嵌入前后马尔可夫矩阵发生的变化 $\Delta \boldsymbol{G} = \boldsymbol{G}_{steg} - \boldsymbol{G}_{ori}$(如图 2)。可以看出本文方法所提取的特征能够充分反映图像在嵌入数据前后的变化。

收稿日期:2007 - 03 - 26;**修回日期**:2007 - 06 - 29。 **基金项目**:国家自然科学基金资助项目(90304017)。

作者简介:崔霞(1982 -),女,江苏扬州人,硕士研究生,主要研究方向:图像处理、模式识别; 童学锋(1963 -),男,湖北蕲州人,教授,主要研究方向:图像处理、模式识别; 黄聪(1981 -),男,广西藤县人,硕士研究生,主要研究方向:图像处理、模式识别。

1.1　提取特征步骤

1)直接读取 JPEG 图像的 DCT 系数(目前只读取 YUV 分解的亮度 Y);

2)对每个 8×8 DCT 分块的中低频部分(包括 DC 系数)按照图 3 所示的顺序展开成 3 个一维向量 $\boldsymbol{V}_i$(i 表示扫描顺序的种类,i = 1,2,3)。为更加直观,仅显示了 8×8 系数矩阵的中低频部分,图中数字表示对其进行扫描的顺序。

$$\begin{bmatrix} 0 & 1 & 5 & 6 & 14 & 15 \\ 2 & 4 & 7 & 13 & 16 & \\ 3 & 8 & 12 & 17 & & \\ 9 & 11 & 18 & & & \\ 10 & 19 & & & & \\ 20 & & & & & \end{bmatrix}$$

(a) ZigZag 扫描

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 10 & 9 & 8 & 7 & 6 & \\ 11 & 12 & 13 & 14 & & \\ 17 & 16 & 15 & & & \\ 18 & 19 & & & & \\ 20 & & & & & \end{bmatrix}$$

(b) 局部水平扫描

$$\begin{bmatrix} 0 & 10 & 11 & 17 & 18 & 20 \\ 1 & 9 & 12 & 16 & 19 & \\ 2 & 8 & 13 & 15 & & \\ 3 & 7 & 14 & & & \\ 4 & 6 & & & & \\ 5 & & & & & \end{bmatrix}$$

(c) 局部垂直扫描

图 3　扫描 DCT 块的三种顺序

3)将 V_i 中的系数值钳位于[$-T$,T]内,即将小于 $-T$ 的系数全部改为 $-T$,大于 T 的系数全部改为 T。本文取 T = 7,就是说钳位于[-7, 7],共计有效整数 15 个,即[-7, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7]。因而形成 15×15 的马尔可夫矩阵;

4)利用马尔可夫模型,统计各 DCT 块一维向量 $\boldsymbol{V}_i$ 的马尔可夫矩阵 $\boldsymbol{G}_{ik}$;

5)计算全局平均马尔可夫矩阵 $\boldsymbol{G}_i = \frac{1}{n}\left(\sum_{k=1}^{n} \boldsymbol{G}_{ik}\right)$,$n$ 为一幅图中 DCT 块总数,i = 1,2,3;

6)由于 15×15 的马尔可夫矩阵 $\boldsymbol{G}_i$ 是对称矩阵,所以只需选择上三角元素作为特征,即对于每个 $\boldsymbol{G}_i$ 可提取 120 维((15×15 + 15)/2 = 120)特征,所以最终形成 360 维特征。

1.2　DCT 域分析

针对 JPEG 图像的数据隐写技术,直接在 DCT 域进行隐写分析优于空域。F5[5]、Outguess[6]、MB[7] 等都是直接在 DCT 域进行数据隐写的。如果将其转到空域进行隐写分析,在转换过程中,DCT 域比较显著的变化由于被分散到空域一定范围的系数上而变得非常微弱,并且取整过程使得精度丢失,以致嵌入信息的丢失,从而大大降低了检出率。从本文的实验可以看出,直接在 DCT 域进行分析的方法(文献[4]和本文)比经过空域转换的方法[2,3,8] 的检出率要提高 15% ~ 30%。

1.3　三种方式扫描

如图 3 所示,图(a)是 ZigZag 顺序,它描述的是 DCT 域频率由小到大的顺序,对于原始图像,按照此顺序展开的系数相邻之间具有较高的连续性和相关性。而嵌入数据后,这种连续性和相关性会遭到一定程度的破坏。本文增加了(b)和(c)两种扫描顺序,主要是为了尽可能多地提取系数的空间相关性。我们仅采用(a)方式扫描来提取特征,对 F5 做了实验,检出率如表 1。将此结果和表 4 比较,可以看出本文方法更加有效。

表 1　仅用(a)方式扫描的实验结果

嵌入量	检出率/%
1kB	81
2kB	92
4kB	98

1.4　保留 DC 系数

由于目前的主流 JPEG 隐写方法仅对非 0 的 AC 系数进行嵌入,所以以前的分析方法都把 DC 系数省略了。但是我们考虑到保留 DC 系数后,就可以保留其与第一个 AC 系数的关系,这样就能最大限度地提取图像的相关性。我们在提取特征时去除 DC 系数,其他过程保持不变,对 F5 做了实验,检出率如表 2。将此结果和表 4 比较,可以看出本文方法更有效。

表 2　不加 DC 系数的实验结果

嵌入量	检出率/%
1kB	82
2kB	92
4kB	98

1.5　设定阈值

DCT 的 AC 系数基本集中于 0 附近,近似于拉普拉斯分布。所以,大部分 DCT 系数都很小。只要选择适当的阈值 T,就既能保证丢失极少的信息,又能极大地降低计算复杂性。本文选取 CorelDraw 库中的 16053.jpg 图像,选取 T = {5,6,7,8,9},统计处于[$-T$,T]范围内的 AC 系数占系数总数的百分比,如表 3。T = 8 或 9 时,虽然百分比要高一点,但并不明显,而且相应的特征维数会明显增大,使得计算复杂,运行速度大大降低,所以本文折中选择 T = 7。

表 3　统计在[$-T$,T]范围内的 AC 系数百分比(%)

[$-T$,T]	百分比	[$-T$,T]	百分比
[-5,5]	97.3	[-8,8]	98.9
[-6,6]	98.1	[-9,9]	99.2
[-7,7]	98.6		

2　支持向量机

支持向量机(SVM)的主要思想是建立一个超平面作为决策曲面,使得正例和反例之间的隔离边缘被最大化。

本文采用 OSU SVM Tool Box[9]。在对样本进行训练和测试前,需要对样本进行[-1,1]的拉伸(有 scale 函数)。选用 RBF 内积核,涉及到参数 C 和 γ 的设置,本文采用循环测试的方法($C = 2^{-5}, 2^{-3}, \cdots, 2^{15}$;$\gamma = 2^{-15}, 2^{-13}, \cdots, 2^{3}$)来选取最优参数。经过反复测试,比较好的一对参数是 $C = 2^9 = 512$,$\gamma = 2^{-1} = 0.5$。

3　实验

实验采用 CorelDraw 图像库,该库共由 1096 张图像,大小

为768×512。为了消除双重压缩造成的影响,我们对BMP图像以质量75%进行JPEG[10]压缩作为原图(cover image);对BMP图像用F5、Outguess、MB1、MB2以质量75%分别进行嵌入作为嵌入图(stego image),嵌入量为1 kB,2 kB,4 kB,分别相当于0.021 bpp,0.041 bpp和0.083 bpp,嵌入程序可以从参考网址[11][12][13]获得。随机提取896对图像用于训练,剩下200对图像用于测试(对于Outguess,由于存在嵌入失败的问题,嵌入量为1 kB,2 kB,4 kB时,嵌入成功的数量分别是1071张,1001张和630张,随机提取200张原图和200张嵌入图作为测试,其余的作为训练),程序运行10次得到平均检出率。

3.1 检出率的比较

本文与文献[2,4]的方法进行对比,所有参与比较的方法都仅针对亮度Y。实验数据见表4,可以看到本文提出的新方法最优。

表4 几种分析方法的检出率比较(%)

嵌入方法	嵌入量	隐写分析方法		
		文献[2]方法	文献[4]方法	本文方法
F5	1kB	60	74	84
	2kB	65	87	94
	4kB	81	96	99
Outguess	1kB	51	89	98
	2kB	64	97	100
	4kB	76	98	100
MB1	1kB	59	66	93
	2kB	70	86	97
	4kB	82	90	99
MB2	1kB	60	62	94
	2kB	69	76	98
	4kB	84	84	99

3.2 稳定性实验

为了验证检出率的稳定性,我们对每次实验都重复运行10次,并记录所得结果如图4。

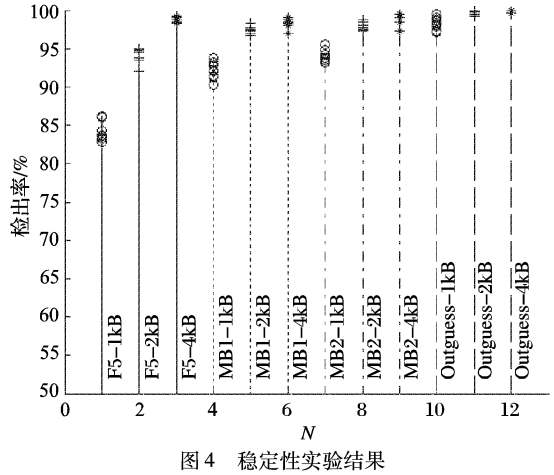

图4 稳定性实验结果

图4中:N表示不同的嵌入方法;"o"、"+"或"*"表示一次随机挑选训练测试样本的实验结果。由图4可见,对于不同训练测试样本的组合,检出率是相当稳定的。

3.3 运行速度比较

由于所用分类器相同,参数也可以预先训练出来,所以本文仅对不同隐写分析方法的特征提取进行时间上的测量。实验电脑配置为:Celeron CPU 1.7 GHz,256 MB内存,Matlab 7.1版。随机挑选CorelDraw图像库中100张原图进行测试,结果如表5所示,可以看出本文提出的方法虽然维数较高,但所耗时间最少。

表5 不同隐写分析方法的运行速度

隐写分析方法	特征维数	百张总耗时/s	平均每张耗时/s
文献[2]方法	39	406.46	4.06
文献[4]方法	23	1159.20	11.59
本文方法	360	313.15	3.13

4 结语

1)利用二阶统计量作为特征的隐写分析方法,直接在分块DCT系数域提取特征,避免了转到空域处理所造成的隐写信息的丢失;

2)分别采用三种方式扫描DCT块提取特征,保留DC系数,充分利用了各系数不同方向的空间相关性,特征维数适当增加,包含了更多有效信息;

3)选择适当的阈值T,保留特定范围内的系数,大大降低计算复杂性;

4)本文对F5、Outguess、MB1和MB2四种JPEG嵌入方法进行隐写分析,在CorelDraw图像库上做实验,稳定性好,检出率和运行速度上都明显优于现有算法,具有很大的实际应用潜力。

参考文献:

[1] FARID H. Detecting hidden messages using higher-order statistical models[C]// Proceedings of the IEEE International Conference on Image Processing 02. New York: IEEE Press, 2002, II: 905-908.

[2] XUAN G R, SHI Y Q, GAO J J, et al, Steganalysis Based on multiple features formed by statistical moments of wavelet characteristic functions [C]// Information Hiding Workshop (IH2005). Berlin: Springer-Verlag, 2005, 3727: 262-277.

[3] SULLIVAN K, MADHOW U, CHANDRASEKARAN S. *et al*. Steganalysis of Spread Spectrum Data Hiding Exploiting Cover Memory [C]// SPIE2005. [S.l.]: San Jose, 2005.

[4] FRIDRICH J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes [C]// The 6th Information Hiding Workshop. Toronto: [s.n.], 2004.

[5] WESTFELD A. F5 a steganographic algorithm: High capacity despite better steganalysis [C]// 4th International Workshop on Information Hiding. Berlin: Springer-Verlag, 2001.

[6] PROVOS N. Defending against statistical steganalysis[C]// 10th USENIX Security Symposium. Washington: [s.n.], 2001.

[7] SALLEE P. Model-based methods for steganography and steganalysis [J]. International Journal of Image and Graphics, 2005, 5(1): 167-190.

[8] LYU S, FARID H. Detecting hidden messages using higher-order statistics and support vector machines[C]// 5th International Workshop on Information Hiding IH 2002. Noordwijkerhout: [s.n.], 2002.

[9] OSU SVM Toolbox for Matlab [EB/OL]. [2006-06-28]. http://sourceforge.net/projects/svm/.

[10] UG library for JPEG image compression [EB/OL]. [2006-06-25]. http://www.ijg.org/.

[11] Steganography software Outguess [EB/OL]. [2006-07-06]. http://www.outguess.org/.

[12] Steganography software F5 [EB/OL]. [2006-07-06]. http://wwwrn.inf.tu-dresden.de/~westfeld/F5.html.

[13] Steganography software Model-Based method [EB/OL]. [2006-07-06]. http://redwood.ucdavis.edu/phil/papers/iwdw03.htm.