

Markov 隐写检测特征的一种新设计

张昊* 平西建

(解放军信息工程大学 郑州 450002)

摘要: 如何在保持 Markov 特征隐写检测能力的同时降低其维数是隐写分析领域中一个重要问题。该文通过将空域图像各向同性的统计假设扩展到离散余弦变换(DCT)域,给出了一种新的特征设计方法。该方法可将传统的块内特征的维数降低 36%,块间特征的维数降低 72%,且广泛地适用于不同提取源不同阶数的 Markov 特征。实验表明,该文的设计方法降低维数的同时还提高了特征的检测性能。

关键词: 隐写分析; Markov 特征; 降维; 联合图像专家组(JPEG)

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2013)08-1907-07

DOI: 10.3724/SP.J.1146.2012.01508

New Design of Markov Steganalytic Features

Zhang Hao Ping Xi-jian

(PLA Information Engineering University, Zhengzhou 450002, China)

Abstract: How to reduce Markov features' dimensionality while keeping their steganalytic ability is an important issue in the field of steganalysis. This paper generalizes the statistical hypothesis that images are isotropic from spatial domain to Discrete Cosine Transform (DCT) domain, and provides a new design method. The proposed method is suitable for Markov features of different orders with respect to different sources of extraction and it can effectively deduce the dimensionality of Markov features. Concretely, it can reduce the dimensionality of traditional intrablock features by 36% and that of interblock features by 72%. Experimental results show that the proposed method can also enhance the features' detection ability.

Key words: Steganalysis; Markov features; Dimensionality reduction; Joint Photographic Experts Group (JPEG)

1 引言

近年来,以数字图像为载体的信息隐藏技术成为信息安全领域的一个研究热点^[1]。数字隐写和隐写检测技术是其主要的研究内容。联合图像专家组(Joint Photographic Experts Group, JPEG)图像压缩比高、质量好、获取方便,是现代隐写术的主要载体。当前针对 JPEG 隐写的检测技术往往以提取有效的隐写检测特征为手段,使用分类器实现对载密图像的检测^[2,3]。

由于大多数 JPEG 隐写术以修改量化后 DCT (Discrete Cosine Transform)系数的方式,将与载体数据不相关的秘密信息比特流嵌入到图像中,从而在一定程度上破坏了相邻系数的弱相关性,这就使得体现统计相关性的 Markov 转移和联合概率成为一类重要的隐写检测特征。自 2006 年 Shi 等人^[4]首次提出基于相邻系数绝对值的差值的 324 维 Markov

特征以来,该类特征受到研究者的广泛关注。2007 年,文献[5]将文献[4]的特征并入混合特征的设计以获得一个更全面的隐写检测特征集;2008 年,文献[6]为进一步完善 Markov 特征,又在块内特征^[4]基础上增加了 162 维的块间 Markov 特征;2010 年,Davidson 等人^[7]基于所谓的偏序 Markov 模型提出了从相邻 DCT 系数中提取的 Markov 特征;2011 年,Liu 等人^[8]为了改进文献[4]和文献[6]的特征建立了分别从相邻系数和系数绝对值中提取的联合概率特征。除上述从 DCT 域提取的特征外,一些空域和小波域 Markov 特征^[9,10]也能有效检测 JPEG 隐写算法。

由于传统 Markov 特征^[4,6]维数较高,文献中普遍使用将不同方向特征矩阵合并的方式降维。该方法在空域特征的处理中取得了理想的效果^[9,11-13],其有效性依据了空域图像各向同性的统计假设。而在与空域统计特性不同的 DCT 域,文献[5,7]中直接将 4 方向特征合并的降维方法是以降低检测性能为代价的,且其有效性也缺少合理的解释。本文考虑将各向同性的统计假设扩展到 DCT 域,从而给出

2012-11-20 收到,2013-03-07 改回

国家自然科学基金(60970142)资助课题

*通信作者:张昊 hao_zhangnew@163.com

一种合理有效的新设计。由于目前 DCT 域特征的区别大多体现在特征提取源的变化上^[7,8], 而没有改变传统特征的设计框架, 因此我们仍以传统特征的改进为说明对象, 其它提取源的提取过程类似。为了方便说明, 我们首先给出新特征的提取过程, 然后通过比较新特征与传统特征的区别, 逐步归纳出具有一般性的新设计和特征提取步骤。

2 特征设计

2.1 各向同性与图像转置

不失一般性, 假设 I_0 为一幅尺寸为 $8M \times 8N$ 的空域图像, 我们将图像 I_0 分成 $M \times N$ 个不交的 8×8 分块, 并对每个分块做 2 维离散余弦变换, 则坐标为 (k, l) 的块内 (m, n) 位置上的系数为

$$\begin{aligned} & \mathbf{F}_0(8k-8+m, 8l-8+n) \\ &= \frac{C(m)}{2} \frac{C(n)}{2} \sum_{i=1}^8 \sum_{j=1}^8 I_0(8k-8+i, 8l-8+j) \\ & \quad \cdot \cos\left(\frac{(2i-1)(m-1)\pi}{16}\right) \cos\left(\frac{(2j-1)(n-1)\pi}{16}\right) \end{aligned} \quad (1)$$

其中 $1 \leq k \leq M, 1 \leq l \leq N, 1 \leq m, n \leq 8$, 当 $m=1$ 时, $C(m) = \sqrt{2}/2$, 当 $m \geq 2$ 时, $C(m) = 1$ 。如此得到的 \mathbf{F}_0 是一幅尺寸为 $8M \times 8N$ 的 DCT 系数图像。根据各向同性的假设, 图像 I_0 及其转置 I_1 具有相同的分布。因此由 I_1 得到的 DCT 系数图像 \mathbf{F}_1 也与 \mathbf{F}_0 具有相同的分布。我们使用相同的质量因子和 JPEG 量化表分别对 \mathbf{F}_0 和 \mathbf{F}_1 量化, 并对每个量化后的 DCT 系数取绝对值, 得到的图像记作 \mathbf{G}_0 和 \mathbf{G}_1 。由于经历了同样的变换过程, \mathbf{G}_0 和 \mathbf{G}_1 也具有相同的分布。不过, 在唯载密的隐写分析中, 我们并不能得到 \mathbf{F}_0 而只能获得 \mathbf{G}_0 , 从而也得不到 \mathbf{F}_1 和 \mathbf{G}_1 。因此我们使用 \mathbf{G}_0 的转置图像代替 \mathbf{G}_1 (下文出现的 \mathbf{G}_1 均为 \mathbf{G}_0 的转置图像)。这即是说, 我们使用的 \mathbf{G}_1 并非真正由 \mathbf{F}_1 得到的图像, 受 JPEG 量化过程对分布的影响, 我们只能近似地认为 \mathbf{G}_0 和 \mathbf{G}_1 同分布。但是, 将 \mathbf{G}_1 应用到隐写检测特征的设计中仍然有意义, 其有效性可通过 4.3 节隐写检测的实验得到验证。

2.2 新特征的设计

传统的 Markov 特征包括了水平、垂直、对角、负对角方向的块内转移概率矩阵^[4]和水平、垂直方向的块间转移概率矩阵^[6]。鉴于联合和转移概率对 Markov 链提供了相同的信息, 这里我们选用联合概率。首先考虑块内特征。如同文献[4], 计算水平 (h)、垂直 (v)、对角 (d)、负对角 (m) 4 个方向的差值图像。

$$\begin{aligned} \mathbf{G}_{\alpha,h}(s, t) &= \mathbf{G}_\alpha(s, t) - \mathbf{G}_\alpha(s, t+1), \quad 1 \leq s \leq 8M, \\ & \quad 1 \leq t \leq 8N-1 \end{aligned} \quad (2)$$

$$\begin{aligned} \mathbf{G}_{\alpha,v}(s, t) &= \mathbf{G}_\alpha(s, t) - \mathbf{G}_\alpha(s+1, t), \quad 1 \leq s \leq 8M-1, \\ & \quad 1 \leq t \leq 8N \end{aligned} \quad (3)$$

$$\begin{aligned} \mathbf{G}_{\alpha,d}(s, t) &= \mathbf{G}_\alpha(s, t) - \mathbf{G}_\alpha(s+1, t+1), \\ & \quad 1 \leq s \leq 8M-1, \quad 1 \leq t \leq 8N-1 \end{aligned} \quad (4)$$

$$\begin{aligned} \mathbf{G}_{\alpha,m}(s, t) &= \mathbf{G}_\alpha(s, t+1) - \mathbf{G}_\alpha(s+1, t), \\ & \quad 1 \leq s \leq 8M-1, \quad 1 \leq t \leq 8N-1 \end{aligned} \quad (5)$$

给定阈值 T , 令所有大于 T 的 $\mathbf{G}_{\alpha,\beta}(s, t)$ 取值 T , 令所有小于 $-T$ 的 $\mathbf{G}_{\alpha,\beta}(s, t)$ 取值 $-T$, 其中 $\alpha \in \{0, 1\}$, $\beta \in \{h, v, d, m\}$ 。计算 4 个方向的特征矩阵为

$$\begin{aligned} \mathbf{A}_{h,\alpha}(u, v) &= \frac{1}{8M(8N-2)} \sum_{s=1}^{8M} \sum_{t=1}^{8N-2} \delta(\mathbf{G}_{\alpha,h}(s, t) = u, \\ & \quad \mathbf{G}_{\alpha,h}(s, t+1) = v) \end{aligned} \quad (6)$$

$$\begin{aligned} \mathbf{A}_{v,\alpha}(u, v) &= \frac{1}{(8M-2)8N} \sum_{s=1}^{8M-2} \sum_{t=1}^{8N} \delta(\mathbf{G}_{\alpha,v}(s+1, t) = u, \\ & \quad \mathbf{G}_{\alpha,v}(s, t) = v) \end{aligned} \quad (7)$$

$$\begin{aligned} \mathbf{A}_{d,\alpha}(u, v) &= \frac{1}{(8M-2)(8N-2)} \sum_{s=1}^{8M-2} \sum_{t=1}^{8N-2} \delta(\mathbf{G}_{\alpha,d}(s, t) = u, \\ & \quad \mathbf{G}_{\alpha,d}(s+1, t+1) = v) \end{aligned} \quad (8)$$

$$\begin{aligned} \mathbf{A}_{m,\alpha}(u, v) &= \frac{1}{(8M-2)(8N-2)} \\ & \quad \cdot \sum_{s=1}^{8M-2} \sum_{t=1}^{8N-2} \delta(\mathbf{G}_{\alpha,m}(s, t+1) = u, \\ & \quad \mathbf{G}_{\alpha,m}(s+1, t) = v) \end{aligned} \quad (9)$$

其中 $\delta(R)$ 是 R 的特征函数, 即若 R 成立, 则 $\delta(R) = 1$, 否则 $\delta(R) = 0$ 。

接下来, 我们考虑块间特征。对每个 (m, n) ($1 \leq m, n \leq 8$), 提取所有块内 (m, n) 位置上的系数构成矩阵 $\mathbf{G}_{\alpha,m,n}$, $\mathbf{G}_{\alpha,m,n}(k, l) = \mathbf{G}(8k+m, 8l+n)$ ($0 \leq k \leq M-1, 0 \leq l \leq N-1$)。与文献[6]不同, 我们计算该矩阵在上 ($v-$)、下 ($v+$)、左 ($h-$)、右 ($h+$) 4 个方向的差值系数矩阵。

$$\begin{aligned} \mathbf{G}_{\alpha,m,n,h+}(s, t) &= \mathbf{G}_{\alpha,m,n}(s, t) - \mathbf{G}_{\alpha,m,n}(s, t+1), \\ & \quad 0 \leq s \leq M, \quad 0 \leq t \leq N-1 \end{aligned} \quad (10)$$

$$\begin{aligned} \mathbf{G}_{\alpha,m,n,v+}(s, t) &= \mathbf{G}_{\alpha,m,n}(s, t) - \mathbf{G}_{\alpha,m,n}(s+1, t), \\ & \quad 0 \leq s \leq M-1, \quad 0 \leq t \leq N \end{aligned} \quad (11)$$

$$\begin{aligned} \mathbf{G}_{\alpha,m,n,h-}(s, t) &= \mathbf{G}_{\alpha,m,n}(s, t+1) - \mathbf{G}_{\alpha,m,n}(s, t), \\ & \quad 0 \leq s \leq M, \quad 0 \leq t \leq N-1 \end{aligned} \quad (12)$$

$$\begin{aligned} \mathbf{G}_{\alpha,m,n,v-}(s, t) &= \mathbf{G}_{\alpha,m,n}(s+1, t) - \mathbf{G}_{\alpha,m,n}(s, t), \\ & \quad 0 \leq s \leq M-1, \quad 0 \leq t \leq N \end{aligned} \quad (13)$$

使用与块内特征相同的阈值 T , 令所有大于 T 的 $\mathbf{G}_{\alpha,m,n,\beta}(s, t)$ 取值 T , 令所有小于 $-T$ 的 $\mathbf{G}_{\alpha,m,n,\beta}(s, t)$ 取值 $-T$, 其中 $\beta \in \{h+, h-, v+, v-\}$ 。计算 4 个方向

的特征矩阵为

$$B_{h+,\alpha}(u, v) = \frac{1}{64M(N-2)} \sum_{m=1}^8 \sum_{n=1}^8 \sum_{s=0}^{M-2} \sum_{t=0}^{N-2} \delta(\mathbf{G}_{\alpha,m,n,h+}(s, t) = u, \mathbf{G}_{\alpha,m,n,h+}(s, t+1) = v) \quad (14)$$

$$B_{v+,\alpha}(u, v) = \frac{1}{64(M-2)N} \sum_{m=1}^8 \sum_{n=1}^8 \sum_{s=0}^{M-2} \sum_{t=0}^N \delta(\mathbf{G}_{\alpha,m,n,v+}(s, t) = u, \mathbf{G}_{\alpha,m,n,v+}(s+1, t) = v) \quad (15)$$

$$B_{h-,\alpha}(u, v) = \frac{1}{64M(N-2)} \sum_{m=1}^8 \sum_{n=1}^8 \sum_{s=0}^{M-2} \sum_{t=0}^{N-2} \delta(\mathbf{G}_{\alpha,m,n,h-}(s, t+1) = u, \mathbf{G}_{\alpha,m,n,h-}(s, t) = v) \quad (16)$$

$$B_{v-,\alpha}(u, v) = \frac{1}{64(M-2)N} \sum_{m=1}^8 \sum_{n=1}^8 \sum_{s=0}^{M-2} \sum_{t=0}^N \delta(\mathbf{G}_{\alpha,m,n,v-}(s+1, t) = u, \mathbf{G}_{\alpha,m,n,v-}(s, t) = v) \quad (17)$$

根据 2.1 节的结论, 我们认为 \mathbf{G}_0 和 \mathbf{G}_1 同分布。又因为两组特征 $\{\mathbf{A}_{h,0}, \mathbf{A}_{v,0}, \mathbf{A}_{d,0}, \mathbf{A}_{m,0}, \mathbf{B}_{h+,0}, \mathbf{B}_{v+,0}, \mathbf{B}_{h-,0}, \mathbf{B}_{v-,0}\}$ 与 $\{\mathbf{A}_{h,1}, \mathbf{A}_{v,1}, \mathbf{A}_{d,1}, \mathbf{A}_{m,1}, \mathbf{B}_{h+,1}, \mathbf{B}_{v+,1}, \mathbf{B}_{h-,1}, \mathbf{B}_{v-,1}\}$ 分别由 \mathbf{G}_0 和 \mathbf{G}_1 通过相同的方式得到, 因此相应的每对特征也是同分布的。将这些同分布的特征合并, 得到

$$\left. \begin{aligned} \mathbf{A}_\beta &= (\mathbf{A}_{\beta,0} + \mathbf{A}_{\beta,1})/2, \quad \beta \in \{h, v, d, m\} \\ \mathbf{B}_\beta &= (\mathbf{B}_{\beta,0} + \mathbf{B}_{\beta,1})/2, \quad \beta \in \{h+, v+, h-, v-\} \end{aligned} \right\} \quad (18)$$

进一步, 观察式(2)-式(17)特征的产生过程可以看出一些矩阵是相等的, 如 $\mathbf{A}_{h,0} = \mathbf{A}_{v,1}$, $\mathbf{A}_{v,0} = \mathbf{A}_{h,1}$, $\mathbf{B}_{h+,0} = \mathbf{B}_{v+,1}$, $\mathbf{B}_{v+,0} = \mathbf{B}_{h+,1}$, $\mathbf{B}_{h-,0} = \mathbf{B}_{v-,1}$, $\mathbf{B}_{v-,0} = \mathbf{B}_{h-,1}$, 从而结合式(18)有 $\mathbf{A}_h = \mathbf{A}_v$, $\mathbf{B}_{h+} = \mathbf{B}_{v+}$ 和 $\mathbf{B}_{h-} = \mathbf{B}_{v-}$ 。此外, 由于 \mathbf{B}_{h+} 和 \mathbf{B}_{h-} 可以分别由图像 \mathbf{F}_0 和其水平翻转图像通过相同的计算方式得到, 根据图像各向同性的假设, 二者也是同分布的, 因而可以合并。最终, 基于图像各向同性所设计的新特征集合为 $\{\mathbf{A}_h, \mathbf{A}_d, \mathbf{A}_m, \mathbf{B}_h\}$, 其中 $\mathbf{B}_h = (\mathbf{B}_{h+} + \mathbf{B}_{h-})/2$ 。

3 特征比较

3.1 维数比较

事实上, 新特征集合中的矩阵 \mathbf{A}_m 和 \mathbf{B}_h 都具有对称性:

(1)从式(5)和式(9)观察 $\mathbf{A}_{m,0}$ 和 $\mathbf{A}_{m,1}$ 的统计过程可知, 每产生一对 $(\mathbf{G}_{0,m}(s+1, t), \mathbf{G}_{0,m}(s, t+1)) = (u, v)$, 就会同时产生一对 $(\mathbf{G}_{1,m}(s+1, t), \mathbf{G}_{1,m}(s, t+1)) = (-v, -u)$, 这即是说 $\mathbf{A}_{m,0}$ 和 $\mathbf{A}_{m,1}$ 关于 $u = -v$ 对称, 结合式(18)可知 \mathbf{A}_m 关于 $u = -v$ 对称;

(2)从式(10), 式(12), 式(14)和式(16)观察 $\mathbf{B}_{h+,\alpha}$ 和 $\mathbf{B}_{h-,\alpha}$ ($\alpha = 0, 1$) 的统计过程可知, 每产生一对 $(\mathbf{G}_{\alpha,m,n,h+}(s, t), \mathbf{G}_{\alpha,m,n,h+}(s+1, t)) = (u, v)$, 就会同时产生一对 $(\mathbf{G}_{\alpha,m,n,h-}(s+1, t), \mathbf{G}_{\alpha,m,n,h-}(s, t)) = (-v, -u)$, 这即是说 $\mathbf{B}_{h+,\alpha}$ 和 $\mathbf{B}_{h-,\alpha}$ 关于 $u = -v$ 对称, 再根据式(18)有 $\mathbf{B}_h = (\mathbf{B}_{h+} + \mathbf{B}_{h-})/2 = (\mathbf{B}_{h+,0} + \mathbf{B}_{h-,0})/4 + (\mathbf{B}_{h+,1} + \mathbf{B}_{h-,1})/4$, 从而得到 \mathbf{B}_h 关于 $u = -v$ 对称。

根据对称性, 原本都是 $(2T+1)^2$ 维的两个矩阵 \mathbf{A}_m 和 \mathbf{B}_h 在隐写分析中实际用到的维数均降为 $(T+1)(2T+1)$ 。参照文献[4]关于阈值选择的结论选择 $T = 4$, 然后比较传统的 Markov 特征和新特征如表 1。

由表 1 可以看出, 文献[4,6]设计的特征一共 6 个矩阵, 维数是 $324 + 162 = 486$, 而本文设计的特征含 4 个矩阵, 维数是 $207 + 45 = 252$ 。与文献中的特征相比, 本文设计的块内特征的维数降低了 36%, 块间特征的维数降低了 72%。

3.2 设计方法比较

传统的 Markov 特征分别给出了几个方向的转移概率矩阵作为特征, 却没有考虑各个方向的特征矩阵的关系以及这些矩阵所服从的分布的特点。本文则从图像各向同性的假设入手, 将空域特征矩阵的同分布关系以及分布的对称性扩展到 DCT 域。我们进一步分析本文与文献[4,6]的特征矩阵的关系, 可以看到:

(1)根据 2.2 节的结论 $\mathbf{A}_{v,0} = \mathbf{A}_{h,1}$ 以及式(6)可知, \mathbf{A}_h 实际上由 $\mathbf{A}_{h,0}$ 与 $\mathbf{A}_{v,0}$ 合并得到;

(2)观察式(4), 式(8)可知 $\mathbf{A}_{d,0} = \mathbf{A}_{d,1}$, 于是有 $\mathbf{A}_d = \mathbf{A}_{d,0}$;

表 1 Markov 特征的维数比较

	块内特征		块间特征	
	文献[4]	本文	文献[6]	本文
矩阵类型	对应 $\mathbf{A}_{h,0}, \mathbf{A}_{v,0}, \mathbf{A}_{d,0}, \mathbf{A}_{m,0}$ 的转移概率	$\mathbf{A}_h, \mathbf{A}_d, \mathbf{A}_m$	对应 $\mathbf{B}_{h+,0}, \mathbf{B}_{v+,0}$ 的转移概率	\mathbf{B}_h
特征维数	$4(2T+1)^2$	$(5T+3)(2T+1)$	$2(2T+1)^2$	$(T+1)(2T+1)$
$T = 4$ 维数	324	207	162	45

(3)根据 3.1 节的结论 $\mathbf{A}_{m,0}$ 和 $\mathbf{A}_{m,1}$ 关于 $u = -v$ 对称, 因此 \mathbf{A}_m 可以仅通过计算 $\mathbf{A}_{m,0}$ 得到;

(4)根据 3.1 节的结论可知, $\mathbf{B}_{h+\alpha}$ 和 $\mathbf{B}_{h-\alpha}$ ($\alpha = 0, 1$) 关于 $u = -v$ 对称, 结合 2.2 节的结论 $\mathbf{B}_{v+,0} = \mathbf{B}_{h+,1}$ 可知, \mathbf{B}_h 可以仅通过计算 $\mathbf{B}_{h+,0}$ 和 $\mathbf{B}_{v+,0}$ 得到。

综上所述, 如同传统的 Markov 特征, 新特征仍然是从 $\mathbf{A}_{h,0}, \mathbf{A}_{v,0}, \mathbf{A}_{d,0}, \mathbf{A}_{m,0}$ 计算得到的, 不需要用到转置图像 \mathbf{G}_1 以及块间特征 $\mathbf{B}_{h-,0}$ 和 $\mathbf{B}_{v-,0}$ 。我们之所以像 2.2 节那样介绍特征的提取过程是为了更清楚地揭示特征矩阵之间的同分布关系以及部分矩阵所具有分布对称性结构。注意到文献中出现了许多种提取 Markov 特征的方法, 其区别反映在不同的提取源: 量化后的 DCT 系数图像^[7], 量化后的 DCT 系数绝对值图像^[8], 量化后的 DCT 系数绝对值的差值图像^[4,6], 我们给出应用各向同性假设设计新特征的一般方法:

(1)对于块内特征, 以相同的方式分别从提取源图像及其转置图像中计算和提取特征;

(2)对于块间特征, 除已选方向上的特征提取外, 增加与每个已选方向相反的方向上的特征提取;

(3)分别分析提取出来的块内和块间特征的同分布关系, 将同分布的特征合并为同一特征;

(4)分析合并后特征的对称性, 确定最终的隐写检测特征维数。

这一方法的优点在于, 其过程与提取源的选择无关, 也可以用于更高阶的 Markov 特征, 具有广泛的适用性。如果仅考虑文献[4]中提到的 4 个方向的块内特征和文献[6]中使用的水平和垂直方向的块间特征, 则提取特征的具体过程可以简化为如下步骤:

(1)提取块内特征: 将水平和垂直方向的特征矩阵合并, 对角线方向的特征矩阵保持不变, 将负对角方向的特征矩阵关于 $u = -v$ 对称位置上的数据合并。

(2)提取块间特征: 将水平和垂直方向的特征矩阵合并, 将合并后得到的特征矩阵关于 $u = -v$ 对称位置上的数据合并。

3.3 计算复杂度比较

在隐写检测的过程中, 特征提取时间是衡量检测方法优劣的重要指标之一。降低特征提取环节的计算复杂度可以提高隐写检测的效率, 因此有必要分析新设计和传统设计的计算复杂度, 进一步比较二者的差别。观察可知, 新设计与传统设计在提取过程中前半段的计算过程完全相同, 即计算块内和块间各个方向的共生矩阵。之后, 传统设计仅要求

对各个方向矩阵做归一化求取联合或转移概率, 而新设计则还要考虑合并同分布的特征。因此我们的分析从得到共生矩阵之后开始。

在传统设计中, 对于一个 $(2T+1)^2$ 维的共生矩阵, 归一化求取联合概率需要 $(2T+1)^2$ 次乘法, 而计算转移概率则需要 $2T(2T+1)$ 次加法和 $(2T+1)^2$ 次乘法。因此该特征的联合概率形式计算总共需要 $6(2T+1)^2$ 次乘法, 而转移概率形式的计算需要 $12T(2T+1)$ 次加法和 $6(2T+1)^2$ 次乘法。

在新设计中, 对于水平和垂直方向的特征(包括块内和块间), 我们分析两种情况, 一是图像行列数目不同, 即 $M \neq N$; 二是图像行列数目相同, 即 $M = N$ 。对于第 1 种情形, 由于水平和垂直方向的块内特征矩阵归一化系数不同, 两者需要分别做归一化, 共花费 $2(2T+1)^2$ 次乘法, 合并则需要 $(2T+1)^2$ 次加法; 类似地, 水平和垂直方向的块间特征矩阵归一化需要 $2(2T+1)^2$ 次乘法, 合并需要 $(2T+1)^2$ 次加法, 此外, 利用对称性降维还需要 $T(2T+1)$ 次加法。对于第 2 种情形, 由于水平和垂直方向的块内特征矩阵归一化系数相同, 两者先合并再做归一化, 只需要 $(2T+1)^2$ 次加法和 $(2T+1)^2$ 次乘法; 类似地, 水平和垂直方向的块间特征矩阵 $(2T+1)^2 + T(2T+1) = (3T+1)(2T+1)$ 次加法和 $(T+1)(2T+1)$ 次乘法。对于负对角方向的块内特征矩阵, 可以先利用对称性降维, 然后再归一化, 这样仅需要 $T(2T+1)$ 次加法和 $(T+1)(2T+1)$ 次乘法。如此计算则情况 1 需要 $2(2T+1)^2 + 2T(2T+1) = (6T+2)(2T+1)$ 次加法和 $5(2T+1)^2 + (T+1) \cdot (2T+1) = (11T+6)(2T+1)$ 次乘法, 情况 2 需要 $(6T+2)(2T+1)$ 次加法和 $3(2T+1)^2 + 2(T+1)(2T+1) = (8T+5)(2T+1)$ 次乘法。情况 2 的乘法次数更少, 因而具有更低的计算复杂度。这事实上给出了图像行列数相等时新设计特征的一个快速提取算法。无论哪种情况, 与传统设计相比, 新设计相对其转移概率形式的加法和乘法次数均有所减少, 相对于其联合概率形式的加法次数增加, 但乘法次数减少。

我们使用 Matlab 软件统计了不同情况下提取特征所花费的时间。以文献[6]中的特征提取源为例, 传统设计中, 提取转移概率和联合概率的时间分别为 2525 s 和 2485 s; 新设计中, 情形 1 和情形 2 的特征提取时间分别为 2321 s 和 2255 s。也就是说, 新方法的特征提取计算复杂度低于传统方法。

4 实验与结果

4.1 实验设置

BOWS2^[14]是一个常用的隐写分析实验图像库,

包含了 10700 幅尺寸为 512×512 的未压缩的 256 色灰度图像。本文从中选择了 5000 幅图像用于生成载体载密图像集。实验用到了当前流行的一些隐写算法, 包括了 Outguess^[15], MB1^[16], MB2^[17], F5^[18], nsF5^[19], PQ^[20], PQ_t^[19], PQ_e^[19], YASS^[21], MME2^[22], MME3^[22]。按照算法与图像内容的关系, PQ_t, PQ_e, MME2, MME3 属于自适应算法。按照输入图像的格式, PQ, PQ_t, PQ_e 要求输入图像质量因子为 85, 而输出的载密图像质量因子为 70, MME2, MME3 则要求输入为未压缩的原始图像。按照修改位置, YASS 与其它算法都不相同。具体地讲, 该算法并不直接修改 DCT 系数, 而是使用图像的空域表示并重新选择分块嵌入。这种方法使得用于估计载体的校准技术不再有效。总之, 我们综合使用这些特点各异的算法从而对特征的检测性能做一个全面的评估。

在实验中, 如果我们在载体载密图像生成的过程中使用不同的 JPEG 压缩器, 则会影响到隐写检测的结果^[19], 这样将可能会干扰到我们对降维效果的评价。因此本文对同种算法的载体载密图像均使用相同的 JPEG 压缩器, 即隐写算法代码提供的压缩器, 以消除该因素带来的影响。除 MB1, MB2, F5, nsF5, YASS 使用质量因子 75 的载体载密图像, 其它算法均按照代码中的要求或相关文献中使用的质量因子。对于 YASS 算法, 我们按照其原始代码的要求, 使用最大嵌入设置。在实验中给出了 4 种隐藏质量因子(QFh)情况下的检测结果。对于其它算法, 生成载密图像时选择的嵌入率指标均为 bpac (bits per non-zero AC coefficient, 平均每个非零交流系数负载的比特数)。

SVM(Support Vector Machine)分类器^[23]是隐写分析中常用的分类器。我们使用高斯核

$$k(x, y) = \exp(-\gamma \|x - y\|_2^2), \quad \gamma > 0 \quad (19)$$

的 SVM 分类器对每种算法的每种嵌入率的载体载密图像特征进行二分类实验。选择 30% 的载体图像及相应的载密图像特征作为训练样本, 剩余的图像特征作为测试样本。SVM 的核参数 γ 和惩罚因子 C 满足

$$(C, \gamma) \in \{(2^i, 2^j) \mid i \in \{-5, -4, \dots, 15\}, \\ j \in \{-15, -14, \dots, 3\}\} \quad (20)$$

这些参数值的选择由训练样本和网格的搜索算法确定。对于测试结果, 我们用检测错误率来评价特征的检测能力。

4.2 隐写检测结果

为验证新的设计方法的有效性, 我们给出 3 组

对照:

(1)通过比较文献[6]特征和按照新设计得到的相应的 252 维特征, 验证新设计对检测性能的影响;

(2)通过比较文献[7]特征和按照新设计得到的相应的 154 维特征, 验证 4 方向合并降维的不合理;

(3)通过比较文献[8]特征和按照新设计得到的相应的 114 维特征, 验证去掉两对角特征矩阵的不合理。

为了公平, 我们对新特征矩阵的截断阈值选择均与相应文献中的推荐阈值相同。由于以上提及的文献涉及到多种特征提取源, 因此这些比较也体现了新设计方法广泛的适用性。观察第(1)个对照组的检测结果可知, 新设计特征仅在 MME 算法 0.15 bpac 时的检测性能有所降低, 而在其它算法的检测中性能普遍有所提高, 在个别算法的检测错误率降低了 3%以上(表 2 中粗体标记)。这意味着新设计不仅能够降低特征的维数, 而且其对同分布特征的合并过程使得特征的检测性能更加鲁棒。观察第(2)个对照组可知, 将 4 方向合并的特征即使借助了差值校准方法(见文献[7]), 其检测性能仍然不如新设计的特征。该特征在 F5, PQ, YASS 系列算法的许多嵌入率情况下, 错误率大大高于新特征(表 2 中粗体标记)。观察第(3)个对照组可知, 将两对角特征矩阵去掉的做法同样会使特征的性能降低。这种差别尽管在修改量化后 DCT 系数的嵌入算法的检测中比较小, 但在 YASS 算法的检测中却表现得尤为明显(表 2 中粗体标记)。

此外, 我们增加了对应于 3 组新特征的 Cartesian 校准特征以观察校准方法对新设计特征的检测性能的影响。数据显示, 校准特征对提高第(1), (3)两组新特征的检测性能的帮助较小, 但对提高第(2)组新特征的检测性能的帮助较明显, 在某些情况下其检测错误率降低了 3%以上(表 2 中粗体标记)。这意味着校准方法的作用可能会随着提取源的变化而变化。由于校准技术带来了成倍的特征, 这使分类过程的计算复杂度大大增加, 而大多情况下, 其在提高特征检测性能方面却并不显著, 因此, 我们认为如何更好地使用校准技术可能成为一个新的研究方向。

4.3 分类时间

由于新设计降低了特征维数, 因此它在分类时间上的开销也远小于传统方法。图 1 展示了 F5 和 PQ 在嵌入率为 0.05 bpac, 0.1 bpac 时的分类时间比较。横轴为训练图像样本在整个图像集上所占比例。纵轴为分类时间。可以看出, 随着训练样本的增多, 新设计相对于传统设计在分类时间方面的优势也越加明显。

表2 Markov 特征的错误率比较(%)

算法	嵌入率	文献[6] 486 维	本文 252 维	校准 504 维	文献[7] 98 维	本文 154 维	校准 308 维	文献[8] 72 维	本文 114 维	校准 228 维
Outguess	0.02	27.37	25.61	22.83	16.43	16.53	15.50	29.60	27.47	24.94
	0.05	6.33	4.91	3.57	3.17	1.56	1.60	8.21	6.54	4.56
MB1	0.02	35.23	32.80	31.46	34.71	35.97	33.11	31.73	30.67	29.69
	0.05	14.90	12.77	11.91	18.00	15.13	12.41	13.06	11.41	10.56
MB2	0.02	32.30	29.74	29.34	33.16	34.39	28.79	29.34	28.31	27.49
	0.05	11.60	9.89	8.94	15.56	12.39	8.93	9.14	8.17	7.27
F5	0.05	29.56	23.04	21.79	26.99	22.83	19.23	20.07	19.96	18.61
	0.10	11.99	5.69	5.04	9.30	5.33	3.53	4.00	3.94	3.10
nsF5	0.10	32.49	27.53	26.91	31.56	26.81	24.57	24.53	24.46	23.31
	0.15	22.36	16.07	15.33	21.34	15.77	12.96	14.17	13.37	11.80
PQ	0.05	21.97	16.64	16.54	28.36	20.01	16.57	16.77	15.40	13.84
	0.10	11.53	6.20	5.19	16.14	6.56	5.14	6.06	5.24	4.46
PQt	0.15	31.77	30.37	30.50	37.54	32.53	30.51	29.84	29.59	28.30
	0.20	26.59	23.61	24.93	31.96	26.00	23.96	24.43	23.01	22.60
PQe	0.15	35.23	30.80	30.81	39.71	32.84	29.71	32.71	31.20	30.44
	0.20	29.11	26.04	25.23	36.43	27.07	24.30	27.14	25.67	24.07
MME2	0.15	47.64	49.86	49.90	43.00	41.39	38.83	41.70	39.84	38.69
	0.20	23.89	23.34	18.27	25.27	22.71	18.49	22.34	20.39	17.79
MME3	0.15	49.46	49.77	49.87	44.40	42.73	41.47	42.07	41.54	39.80
	0.20	29.20	27.86	23.77	31.50	28.77	24.90	27.67	26.11	23.13
YASS, QFh=60		5.94	4.44	4.04	16.19	6.80	6.30	7.30	4.47	4.20
YASS, QFh=65		9.79	8.37	7.17	28.14	14.64	12.70	14.90	8.66	7.51
YASS, QFh=70		14.07	12.54	11.41	32.73	19.90	17.07	21.64	13.89	12.63
YASS, QFh=75		19.46	17.46	18.03	34.36	22.13	21.77	27.31	19.94	18.74

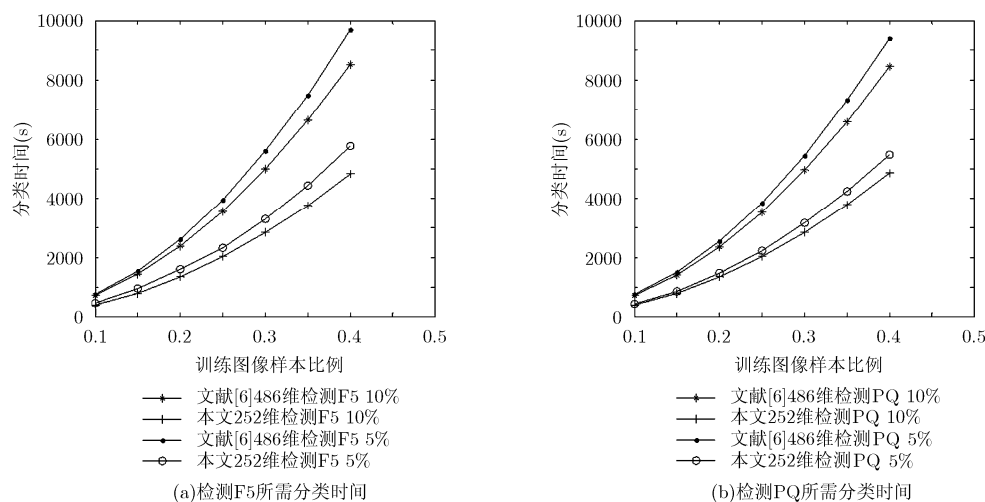


图1 分类时间比较

5 结束语

本文将空域特征设计所用到的各向同性统计假设推广到 DCT 域, 从而给出了一种能有效降维的特征设计方法。新方法具有广泛的适用性, 能够应

用到不同提取源不同阶数的 Markov 特征的设计中。实验结果表明, 新的设计方法无论在特征提取时间还是分类时间方面的开销均小于传统的设计方法, 且隐写检测的结果表明, 该方法不仅能降低特征的

维数, 还能使特征的检测性能更加鲁棒。从实验结果可以看出, 现有方法对自适应算法的检测性能还比较弱。下一步, 我们将对自适应算法展开研究以寻找更好的检测方法。

致谢 感谢中山大学黄方军博士对我们的帮助! 他提供的隐写算法代码使我们得以顺利地完论文的工作。

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] Kodovský J, Fridrich J, and Holub V. Ensemble classifiers for steganalysis of digital media[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(2): 432-444.
- [3] Kodovský J and Fridrich J. Steganalysis of JPEG images using rich models[C]. *Proceedings of SPIE Media Watermarking, Security, and Forensics*, Burlingame, 2012: 1-13.
- [4] Shi Y Q, Chen C, and Chen W. A Markov process based approach to effective attacking JPEG steganography[C]. *Proceedings of Information Hiding Workshop*, Alexandria, 2006: 249-264.
- [5] Pevný T and Fridrich J. Merging Markov and DCT features for multi-class JPEG steganalysis[C]. *Proceedings of SPIE Security, Steganography, and Watermarking of Multimedia*, San Jose, 2007: 1-14.
- [6] Chen C and Shi Y Q. JPEG image steganalysis utilizing both intrablock and interblock correlations[C]. *IEEE International Symposium on Circuits and Systems*, Washington, 2008: 3029-3032.
- [7] Davidson J and Jalan J. Steganalysis using partially ordered Markov models[C]. *Proceedings of Information Hiding Workshop*, Calgary, 2010: 118-132.
- [8] Liu Q, Sung A H, and Qiao M. Neighboring joint density-based JPEG steganalysis[J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(2): 1-16.
- [9] 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.
- [10] Liu Q, Sung A H, Qiao M, *et al.* An improved approach to steganalysis of JPEG images[J]. *Information Sciences*, 2010, 180: 1643-1655.
- [11] Fridrich J and Kodovský J. Rich models for steganalysis of digital images[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(3): 868-882.
- [12] Fridrich J, Kodovský J, Goljan M, *et al.* Steganalysis of content-adaptive steganography in spatial domain[C]. *Proceedings of Information Hiding Workshop*, Prague, 2011: 102-117.
- [13] Kodovský J, Fridrich J, and Holub V. On dangers of overtraining steganography to an incomplete cover model[C]. *Proceedings of ACM Workshop on Multimedia & Security*, New York, 2011: 69-76.
- [14] Bas P and Furon T. BOWS-2 Jul. 2007 [DB/OL]. <http://bows2.gipsa-lab.inpg.fr>. 2007.
- [15] Provos N. Defending against statistical steganalysis[C]. *Proceedings of USENIX Security Symposium*, Washington, 2001: 13-17.
- [16] Sallee P. Model-based steganography[C]. *Proceedings of Digital Watermarking Workshop*, Seoul, 2003: 154-167.
- [17] Sallee P. Model-based methods for steganography and steganalysis[J]. *International Journal of Image and Graphics*, 2005, 5(1): 167-190.
- [18] Westfeld A. High capacity despite better steganalysis (F5—a steganographic algorithm)[C]. *Proceedings of Information Hiding Workshop*, Pittsburgh, 2001: 289-302.
- [19] Kodovský J, Fridrich J, and Pevný T. Statistically undetectable JPEG steganography: dead ends, challenges, and opportunities[C]. *Proceedings of ACM Workshop on Multimedia & Security*, Dallas, 2007: 3-14.
- [20] Fridrich J, Goljan M, and Soukal D. Perturbed quantization steganography using wet paper codes[C]. *Proceedings of ACM Workshop on Multimedia & Security*, Magdeburg, 2004: 4-15.
- [21] Solanki K, Sarkar A, and Manjunath B S. YASS: Yet another steganographic scheme that resists blind steganalysis[C]. *Proceedings of Information Hiding Workshop*, Saint Malo, 2007: 16-31.
- [22] Huang F, Huang J, and Shi Y Q. New channel selection rule for JPEG steganography[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(4): 1181-1191.
- [23] Vapnik V N. *The Nature of Statistical Learning Theory*[M]. New York: Springer-Verlag, 1995: 138-146.

张 昊: 男, 1984年生, 博士生, 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。