

一种基于 LSB 序列局部特征的通用隐写检测方法

钟尚平,徐巧芬,陈羽中,何凤英

(福州大学数学与计算机科学学院,福建福州 350108)

摘要: 基于短重码间距统计的隐写检测方法对 LSB 匹配等隐写技术有良好的检测性能.然而该方法为适应不同的应用场合,需要选择适当的短重码维数.这种一元统计分析方法无法考虑多个特征之间存在的联系,从而影响检测能力.本文分析证明了单个短重码间距统计变量的检测能力规律,给出了可减少检测次数的合理选择短重码维数的方法.基于短重码间距统计特征变量之间的相关性选择特征子集,构造局部特征描述向量,进而提出一种基于 LSB 序列局部特征的通用隐写检测方法.该方法采用 GMM 生成模型描述多维局部特征,并基于全局序列词汇设计融合 GMM 生成模型与 SVM 判别方法的分类器.实验结果表明:本文方法在有效控制虚警率的前提下,对 LSB 匹配隐写和 LSB 替换隐写都有较好的检测性能.

关键词: 通用隐写分析;短重码间距统计;LSB 序列;局部特征;高斯混合模型;全局序列词汇

中图分类号: TP911.73 **文献标识码:** A **文章编号:** 0372-2112 (2013) 02-0239-09

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2013.02.006

A Universal Steganalysis Method Based on Local Features Extracted from LSB Sequences

ZHONG Shang-ping, XU Qiao-fen, CHEN Yu-zhong, HE Feng-ying

(College of Mathematics and Computer Science, Fuzhou University, Fuzhou, Fujian 350108, China)

Abstract: The steganalysis method based on spacing statistics of short duplicate code is an efficient universal detection algorithm for LSB matching steganography. But the steganalysis method must select the appropriate dimension of short duplicate code to meet the different applications. This one-dimension statistical analysis method could not take into account the links between multivariate statistical features, thus may affect the detection capability. In this paper, a detection capability law of a single short duplicate code statistical feature is proved, and a method to reasonable choice of the short duplicate code dimension is presented to reduce the detection number. By analyzing the correlation between the statistical features of short duplicate code spacing statistics, a selected feature subset is described as a vector of local features. Then, a universal steganalysis method based on local features extracted from LSB sequences is proposed. The proposed steganalysis method uses the Gaussian Mixture Model (GMM) to describe the multi-dimensional local features, and designs classifier by integrating GMM generative model and SVM discriminative method based on global sequence vocabulary. The experimental results show that under the premise of effective control of the false alarm rate, the proposed method achieves the best overall detection performance to LSB matching steganography and to LSB replacement steganography.

Key words: universal steganalysis; spacing statistics of short duplicate code; LSB sequence; local features; Gaussian mixture model; global sequence vocabulary

1 引言

目前,对隐写图像的检测主要有两条途径.一是针对某一种具体的隐写方法提出的专用隐写检测方法;二是通用隐写检测方法(盲检测方法),主要有序列随机性检测法和通过提取特征并通过学习训练建立模式分类模型的方法.由于新的隐写方法不断出现,通用隐写检测的重要性日益突出,因此近几年对这一类方法的研究

明显加强^[1].

图像最不重要比特位(Least Significant Bit, LSB)隐写法是出现较早的一种时/空域隐写方法,主要分为替换法和匹配法.对基于替换的 LSB 隐写,已经提出了有效的专用检测方法;对基于匹配的 LSB 隐写,已有的专用检测方法效果一般,且对载体图像的依赖性较强.目前,对 LSB 隐写的通用隐写检测主要采用序列随机性检测方法.另外,王国新等提出了一种基于位平面相关性的

隐写检测方法^[2],该方法通过定义短重码间距统计概念,推导分析了统计分布特征,并采用 Poisson 分布显著性检验检测隐秘信息.此法对 LSB 匹配、LSB 替换、直方图补偿等隐写有良好的检测性能;该方法不仅适用于 LSB 位平面,也可以针对其他的位平面;且在检测过程中不针对具体的隐写算法,具有较强的通用性.

局部特征利用图像中关键点的邻域信息来描述图像,在图像及场景分类领域有着广泛的应用^[3~5].基于局部特征的图像分类方法包括概率生成模型(产生式模型)、判别分类方法以及生成模型与判别方法的结合等.生成模型和判别分类方法都取得了不错的效果,而将这两种方法的结合越来越显示出更大的优越性^[6].目前,生成模型与判别方法的融合方法主要有:第一类融合法为利用关键点包(bag of keypoints)的码本法^[7],该方法仿效文本自动分析领域中的词包^[8](bag of words)思想第二类融合法直接采用基于图像底层局部特征的生成模型^[9]来描述对象,并基于分布核的 SVM 判别方法进行分类^[10];第三种融合方法通过生成模型将原图像特征空间转换到新的特征空间,在新的特征空间上采用判别方法进行识别,如:文献^[11]提出的全局词汇法.

本文利用文献^[2]提出的 LSB 序列短重码的分布特征,仿效用于图像分类等领域的局部特征生成方法和基于局部特征的分类器设计方法:生成模型与判别法的融合法,提出了应用于图像隐写检测的 LSB 序列局部特征,基于全局序列词汇构造融合 GMM 生成模型与判别方法的分类器.另外,针对文献^[2]方法为适应不同的应用场合,需要根据待检测图像的质量和检测率、虚警率的要求选择适当的短重码维数的一元统计分析法,本文通过分析证明单个特征变量的检测能力规律,提出了更为合理的选择短重码维数的方法.

图 1 是本文方法的整体框架图.

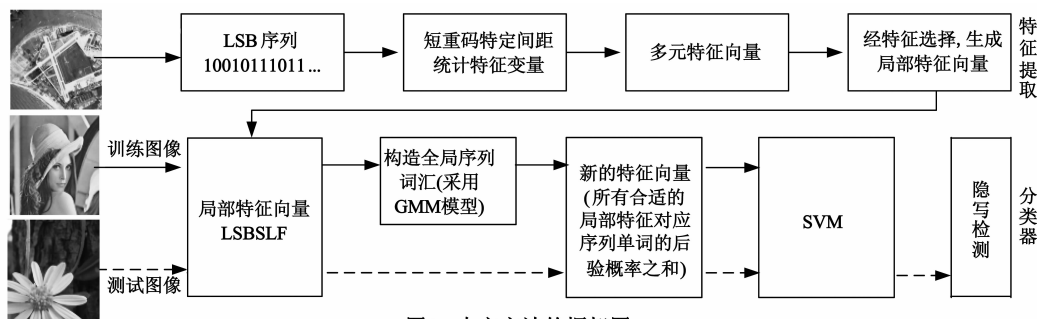


图1 本文方法的框架图

定理 1 自然图像 LSB 序列(长度为 N)在 $d = W$ 处的短重码间距统计结果 $X(W)$ 服从参数为 $N - d - l + 1, q^l$ 的二项分布,近似服从 Poisson 分布,即:

$$p(X(d) = k, d = W) \approx \frac{\lambda^k e^{-\lambda}}{k!}$$

2 LSB 序列中的局部特征

首先根据文献^[2]介绍要用到的术语和定义.

定义 1 把图像最低有效位平面的每一行串接起来,可以构成一个比特序列,称之为 LSB 比特序列^[2].

设灰度 BMP 格式图像的宽和高分别为 W 和 H ,其 LSB 比特序列可看作一个长度为 $N = W \times H$ 的序列 V .

定义 2 如果序列 V 中存在维数相同的子序列(单词) v_i 和 $v_j, i \neq j$,若 $v_i = v_j$,且子序列(单词)的维数 l (即子序列的长度,或子序列包含的比特数)为一个较小的值时(如: $2 \leq l \leq 8$),则称子序列(单词) v_i 和 v_j 为一对短重码; $v_i, v_j (i \neq j)$ 在 V 中的位置差称为短重码间距,记为 $d, d \in \{1, 2, \dots, N - l\}$; 序列 V 中相距为 d 的各种状态短重码个数的累积称为短重码数量,记为 $k, k \in \{0, 1, \dots, N - d - l + 1\}$.

定义 3 短重码间距统计,对待检测序列进行短重码统计,统计短重码数量和间距之间的对应关系,并忽略重码本身的状态和具体位置.统计结果为重码的各个状态在给定间距下的统计累积.

自然图像作为一个局部区域平稳的 Markov 信源,对于其中的某一局部区域,其像素值之间有很强的相关性.也就是说,某一像素的取值只与其一定大小的空间邻域像素的取值相关.自然图像的上述性质反映到 LSB 序列上,有如下统计相关性^[2]:当重码间距接近图像宽度 W 或图像宽度较小的整数倍(如: $2W, 3W$ 等)时,对应的像素为相邻行像素,像素距离较近,会表现出较大的相关性,短重码相等的概率较大,使得统计结果中此间距的短重码明显增多.而自然图像嵌入秘密信息(一般都经过加密处理)后,LSB 序列呈现出秘密信息特征,短重码间距统计没有间距为图像宽度时的统计优势.而且,自然图像和载密图像 LSB 序列中的短重码在特定间距 W 时的统计结果有如下分布规律^[2]:

其中 $\lambda = E[X(d), d = W] = (N - d - l + 1)q^l, \frac{1}{2} \leq q \leq 1$, e 为自然数;而长度为 N 的二值完全随机白噪声序列在 $d = W$ 处的短重码间距统计结果 $X(W)$ 服从参数为 $N - d - l + 1, (1/2)^l$ 的二项分布,近似服从 Poisson 分

布,即:

$$p(X(d) = k, d = W) \approx \frac{\lambda^k e^{-\lambda}}{k!}$$

其中 $\lambda = E[X(d), d = W] = (N - d - l + 1) \left(\frac{1}{2}\right)^l$, e 为自然数.

定理 1 的证明见文献[2].

记序列 V 的维数为 l 短重码在特定间距为 $W, 2W, 3W$ 时的统计特征变量为 $X_l(W), X_l(2W), X_l(3W)$, 其中 $2 \leq l \leq 8$, 显然 $X_l(nW), n = 1, 2, 3$ 与 $X(W)$ 的分布规律是一致的. 另外, 由定理 1 容易定量估计上述单个特征变量的可检测分类能力, 并找到文献[2]方法中, 检测不同隐写图像时, 合理选择短重码维数的方法.

2.1 单个特征变量 $X_l(nW)$ 的检测能力估计及合理选择短重码维数的方法

我们采用经典的如下定义的 Fisher 判别率 (Fisher's Discriminant Ratio, FDR)^[12] 来定量估计单个特征变量 $X_l(nW)$ 的可检测分类能力.

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}, \text{ 其中 } \mu_1, \sigma_1 \text{ 表示自然图像特征}$$

变量 $X_l(nW)$ 的均值和方差; 而 μ_2, σ_2 表示载密图像特征变量 $X_l(nW)$ 的均值和方差. 一般地, FDR 越大, 则单个特征变量 $X_l(nW)$ 的可检测分类能力越大.

由于 Poisson 分布的均值和方差相等, 且都等于参数 λ , 根据定理 1, 可得单个特征变量 $X_l(nW)$ 的 Fisher 判别率:

$$\begin{aligned} FDR(l) &= \frac{(N - d - l + 1)^2 \left(q^l - \left(\frac{1}{2}\right)^l\right)^2}{(N - d - l + 1)^2 \left(q^{2l} + \left(\frac{1}{2}\right)^{2l}\right)} \\ &= \frac{q^{2l} + \left(\frac{1}{2}\right)^{2l} - 2q^l \left(\frac{1}{2}\right)^l}{q^{2l} + \left(\frac{1}{2}\right)^{2l}} \end{aligned} \quad (1)$$

通过比较 $X_l(nW)$ 与 $X_{l+1}(nW)$ 的 Fisher 判别率大小, 可得如下结论:

定理 2 特征变量 $X_l(nW)$ 的 Fisher 判别率 $FDR(l)$ 不大于特征变量 $X_{l+1}(nW)$ 的 Fisher 判别率 $FDR(l+1)$.

证明: 由式(1), 可得:

$$FDR(l+1) = \frac{q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2} - 2q^{l+1} \left(\frac{1}{2}\right)^{l+1}}{q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2}}$$

考虑 $\frac{FDR(l+1)}{FDR(l)}$ 的大小, 若比值大于等于 1, 则定理成立. 由于

$$\frac{FDR(l+1)}{FDR(l)} = \frac{A - 2q^l \left(\frac{1}{2}\right)^l \left(\frac{1}{2} q^{2l+1} + q \left(\frac{1}{2}\right)^{2l+1}\right)}{A - 2q^l \left(\frac{1}{2}\right)^l \left(q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2}\right)},$$

其中

$$A = q^{4l+2} + \left(\frac{1}{2}\right)^{4l+2} + \left(\frac{1}{2}\right)^{2l+2} q^{2l} + \left(\frac{1}{2}\right)^{2l} q^{2l+2}.$$

显然, 要证 $\frac{FDR(l+1)}{FDR(l)} \geq 1$, 只需:

$$2q^l \left(\frac{1}{2}\right)^l \left(\frac{1}{2} q^{2l+1} + q \left(\frac{1}{2}\right)^{2l+1}\right) \leq 2q^l \left(\frac{1}{2}\right)^l \left(q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2}\right).$$

故只要下式小于等于 1, 则定理成立:

$$\frac{2q^l \left(\frac{1}{2}\right)^l \left(\frac{1}{2} q^{2l+1} + q \left(\frac{1}{2}\right)^{2l+1}\right)}{2q^l \left(\frac{1}{2}\right)^l \left(q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2}\right)} = \frac{\frac{1}{2} q^{2l+1} + q \left(\frac{1}{2}\right)^{2l+1}}{q^{2l+2} + \left(\frac{1}{2}\right)^{2l+2}},$$

设 $q = \frac{1}{2} + \epsilon, 0 \leq \epsilon \leq \frac{1}{2}$, 则上式为:

$$\frac{\frac{1}{2} q^{2l+1} + \epsilon \left(\frac{1}{2}\right)^{2l+1} + \left(\frac{1}{2}\right)^{2l+2}}{\frac{1}{2} q^{2l+1} + \epsilon q^{2l+1} + \left(\frac{1}{2}\right)^{2l+2}},$$

因为 $\frac{1}{2} \leq q \leq 1$, 故上式小于等于 1 成立.

定理获证.

定理 2 表明: $FDR(l)$ 随 l 单调增加. 实际应用文献[2]方法时, 因为 $X_l(nW)$ 仅仅是近似服从 Poisson 分布, 我们不能仅用 $X_7(nW)$ 或 $X_8(nW)$ 作为每幅图像的检测特征, 但定理 2 保证了: 若短重码维数 l 的选择按从大到小顺序, 则在相同参数阈值约束下, 可较大地减少检测次数, 从而减少运算时间.

另一方面, 文献[2]方法使用了一元(一种维数的短重码特定间距统计特征变量)统计检验的思想方法来检测隐秘信息. 该方法把多种维数的短重码特定间距统计特征变量 $X_l(W)$ 分开分析, 一次处理一个特征变量, 这样做也许比较简单方便, 但由于这种做法忽视了诸多特征变量相互之间可能存在的联系. 因此, 一般会丢失很多对隐写检测有用的信息. 很自然地, 我们会想到采用多元(如: $(X_2(W), X_3(W), \dots, X_8(W))$)等统计检验的思想方法来检测隐秘信息. 因为 $X_l(W), l = 2, 3, \dots, 8$ 相互之间存在的关系影响着生成的特征向量的检测分类能力.

本文不妨以相关性来描述 $X_l(W), l = 2, 3, \dots, 8$ 相互之间存在的关系, 并从相关系数矩阵中获取特征的生成依据.

2.2 特征变量 $X_l(W)$ 之间的相关性与特征选择

下面采用线性相关系数 $\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$ 来衡量特征变量 $X_l(W), 2 \leq l \leq 8$ 两两之间的相关程度.

另外, 文献[13]提出了相关分析用于特征选择的理论依据: 在原特征集中增加或删除完全相关(即相关系

数绝对值为 1) 特征或近似完全相关(即相关系数绝对值近似为 1) 特征, 不影响该特征集的分类能力. 我们将根据此特征选择依据, 在特征集 $(X_2(W), X_3(W), \dots, X_8(W))$ 中选择隐写检测特征向量.

首先我们通过实验求得特征变量 $X_l(W), 2 \leq l \leq 8$ 两两间的相关系数矩阵. 实验中, 我们采用 8000 张 512×512 Bmp 格式图像作为自然图像库 PD1, 这些图片取自 USC-SIPI 的不带水印的图像库^[14]和我们自采集的图像库. 基于区域相关性的 LSB 匹配嵌入方法^[15], 向自然图像(载体图像)中按嵌入率 $\gamma = 100\%$ 嵌入随机比特流. 表 1 为 7 个特征变量 $X_2(W), X_3(W), \dots, X_8(W)$ 的相关系数矩阵, 其中相关系数值的范围在 -1 和 +1 之间; 相关系数值为 0 表示不相关; 相关系数的绝对值越接近 1, 相关程度越高. 明显地, 特征变量 $X_l(W)$ 与 $X_{l+1}(W)$ 之间的相关程度较高, 但其它特征变量两两间的相关程度不太高. 因此, 按照下述特征选择原则^[13], 取门限值 $\tau = 0.9$, 易得特征子集为 $(X_2(W), X_4(W), X_6(W), X_8(W))$:

表 1 随机变量 $X_l(W)$ 之间的相关系数

$X_l(W)$	$X_2(W)$	$X_3(W)$	$X_4(W)$	$X_5(W)$	$X_6(W)$	$X_7(W)$	$X_8(W)$
$X_2(W)$	1	0.910372	0.793519	0.675535	0.563125	0.541216	0.433211
$X_3(W)$	-	1	0.881633	0.700349	0.660745	0.623343	0.456319
$X_4(W)$	-	-	1	0.922478	0.847739	0.773691	0.615627
$X_5(W)$	-	-	-	1	0.948112	0.873011	0.704297
$X_6(W)$	-	-	-	-	1	0.924843	0.807154
$X_7(W)$	-	-	-	-	-	1	0.931741
$X_8(W)$	-	-	-	-	-	-	1

设置一个门限值 $\tau (0 < \tau < 1)$, 按顺序 $i = 2, 3, \dots, 8; j = 2, 3, \dots, 8$ 来考察 7 个特征 $X_2(W), X_3(W), \dots, X_8(W)$. 当 $\rho_{X_i(W), X_j(W)} > \tau$ 时, 比较 $X_i(W), X_j(W)$ 在优劣排序中的次序, 将排序在后的那个特征剔除; 然后继续考察剩下的特征, 特别地, 当遇到两个特征中已有一个特征被剔除时, 无条件保留剩下的那个特征; 全部考察完毕后, 将保留下来的特征构成一个降维的特征子集.

在下一节将述的应用于图像隐写检测的局部特征中, 我们将以特征子集 $(X_2(W), X_4(W), X_6(W), X_8(W))$ 作为局部特征描述向量.

2.3 应用于图像隐写检测的局部特征-LSBSLF

本文将仿效广泛应用于图像分类等领域的局部特征的生成方法, 构建应用于图像隐写检测的局部特征.

设图像 LSB 序列 V 的短重码间距统计结果集合为 $\{X_l(d) | d = 1, 2, \dots, N - l; l = 2, 3, 4, \dots, 8\}$, 则依据定理 1 可知, $d = W, 2W, 3W$ 等是统计相对稳定的关键点. 依据上节结论, 特征子集 $(X_2(W), X_4(W), X_6(W),$

$X_8(W))$ 可作为局部特征描述向量.

不妨称应用于图像分类等领域的局部特征为传统局部特征; 称本文提出的应用于图像隐写检测的基于 LSB 序列的局部特征为 LSBSLF (LSB Sequence Local Features). 比较传统局部特征与 LSBSLF 的特点, 有: 传统局部特征的原始处理对象为图像的像素值, 而 LSBSLF 的原始处理对象为图像的 LSB 序列; 另外, 评估传统局部特征的性能指标一般分析其是否具有不变性和鲁棒性等, 而评估 LSBSLF 的性能指标一般分析其对不同类型的图像是否统计稳定等.

3 生成模型与判别方法相融合的分类器

考虑到采用全局词汇法^[11]来融合生成模型与判别方法, 其性能更为有效. 本文将采用文献[11]的思想设计生成模型与判别方法相融合的分类器.

3.1 采用 GMM 生成模型描述多维局部特征

虽然我们已知一元随机变量 $X_l(W), X_l(2W), X_l(3W)$ 近似服从 Poisson 分布, 但我们并不能据此推断多维 LSBSLF 向量: $(X_2(W), X_4(W), X_6(W), X_8(W))$ 等的多元分布密度. 而 GMM 作为高斯概率密度函数的一个线性组合, 只要有足够多数目的混合分量, 就可以任意逼近一个密度函数. 我们将采用 GMM 来描述多维 LSBSLF 向量的多元分布密度函数, 并建立生成模型.

一个 M 阶高斯混合模型的概率密度函数是由 M 个高斯概率密度函数加权求和得到, 即: $P(\mathbf{X}/\lambda) = \sum_{i=1}^M w_i b_i(\mathbf{X})$, 其中, \mathbf{X} 是一个 D 维随机向量; 对于给定的时间序列 $\mathbf{X} = \{X_t\}, t = 1, 2, \dots, T, b_i(X_t), i = 1, \dots, M$ 是子分布; $w_i, i = 1, \dots, M$ 是混合权重, 满足以下条件:

$\sum_{i=1}^M w_i = 1$. 每个子分布是 D 维的联合高斯概率分布, 可表为:

$$b_i(\mathbf{X}) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}_i|^{1/2}} \times \exp\left\{-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{X} - \boldsymbol{\mu}_i)\right\},$$

其中, $\boldsymbol{\mu}_i$ 是均值向量; $\boldsymbol{\Sigma}_i$ 是协方差矩阵. 完整的高斯混合模型 λ 由下式表示:

$$\lambda = \{w_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}, i = 1, \dots, M.$$

通过适当地选择 GMM 的均值、(协)方差和概率权重, 就可以完成对一个概率密度函数的建模. 通常对角化的协方差矩阵就足以形成较好的近似, 并且能有效地减少需要估计的未知变量的数目. 本文采用 Figueiredo-Jain Algorithm^[16]求解高斯混合模型, 以便针对自然图像和具有不同嵌入率的载密图像, 自适应地确定 GMM 高斯子分布的数量.

3.2 用全局序列词汇构造新的特征向量

首先我们以视觉领域中的视觉单词和视觉词汇表作为类比对象,定义全局序列单词和词汇表。

定义 4 以 λ_0 和 λ_1 表示载体图像和载密图像局部特征的 GMM 模型. 设 λ_0 的子分布为 $b_i^0(\mathbf{X})$, $i = 1, \dots, M_0$; λ_1 的子分布为 $b_i^1(\mathbf{X})$, $i = 1, \dots, M_1$. 把每个子分布 $b_i^j(\mathbf{X})$ 定义为全局序列单词. 把全局序列单词集合:

$\mathbf{B} = \{b_i^0(\mathbf{X}) \mid i = 1, 2, \dots, M_0\} \cup \{b_i^1(\mathbf{X}) \mid i = 1, 2, \dots, M_1\}$ 称为全局序列词汇表. 全局序列词汇表的势(集合的元素的个数)为 $M_0 + M_1$.

本文在一幅图像中取三个局部特征向量: $\mathbf{X}_{1W} = (X_2(W), X_4(W), X_6(W), X_8(W))$, $\mathbf{X}_{2W} = (X_2(2W), X_4(2W), X_6(2W), X_8(2W))$ 和 $\mathbf{X}_{3W} = (X_2(3W), X_4(3W), X_6(3W), X_8(3W))$, 并用上述局部特征对不同全局序列单词的后验概率^[11]之和构成新的特征向量:

$$\mathbf{Y} = (y_1, y_2, \dots, y_{M_0 + M_1}).$$

这个新的特征向量作为支持向量机(SVM)^[17]的输入. 新的特征向量具有如下优点:(1)用 GMM 模型的子分布作为全局序列单词,并用上述局部特征对不同全局序列单词的后验概率之和构成新的特征向量,使得这个新的特征向量具有比直接使用局部特征向量 \mathbf{X}_{1W} 、 \mathbf{X}_{2W} 和 \mathbf{X}_{3W} 具有更丰富的全局结构信息;(2)基于 GMM 模型的后验概率描述中,不仅可以使得各个特征更平滑地反映词汇的分布,而且允许一个特征对多个单词做出反映,因此可以编码更多的判别分类信息。

3.3 隐写检测算法

Step1: 对载体图像和载密图像的局部特征建立 GMM 模型;

Step2: 构造全局序列词汇表 \mathbf{B} ;

Step3: 在每一张图像中取三个局部特征向量 \mathbf{X}_{1W} 、 \mathbf{X}_{2W} 和 \mathbf{X}_{3W} , 求这些局部特征对词汇表 \mathbf{B} 中的每个单词的后验概率之和, 构成新的特征向量 \mathbf{Y} . 以 \mathbf{Y} 作为 SVM 模型的输入;

Step4: 以大量的载体图像和载密图像作为训练样本, 按 Step3 训练 SVM 模型;

Step5: 对每一张测试图像, 用 Step4 训练生成的 SVM 模型进行检测分类。

4 实验结果与分析

4.1 实验建立

实验使用了两个数据集: PD1 和 PD2. PD1(本文 2.2 节已使用)包含 8000 张 512×512 Bmp 格式图像, 取自 USC-SIPI 的不带水印的图像库^[14]和我们自采集的图像. PD2 包含 6000 张 1024×1024 Bmp 格式图像, 取自我们自采集的图像. 数据集中的图片主题包括人物, 风

景, 建筑, 动植物等. 为了满足实验的需求, 我们对图片进行了灰度化处理. 自采集的图像来自不同的摄影者, 不同的相机和不同的场景. 因为目前互联网上开放的图像库经常都带有水印信息, 无法用来做隐写检测实验, 故为了准备多个不同规格(如: 不同大小等)的数据集, 通常需要自己采集部分或全部图像。

实验中采用两类 LSB 隐写算法(LSB 匹配法^[15]和 LSB 替换法^[18])把服从均匀分布的加密序列嵌入到图像的最低位平面中, 且分别按十种嵌入率 100%, 90%, ..., 20%, 10% 等比例隐写秘密信息。

对载体图像和嵌入不同比例秘密信息的载密图像, 用本文提出的方法进行隐写检测, 并与文献[2]方法、频率检测、串行检测、Poker 检测和游程检测^[19]等方法进行比较分析. 另外, 本文方法也与仅采用生成模型或仅采用判别方法的下述两种检测方法进行了比较分析: 一是基于 GMM 模型的方法^[20], 不妨称为 The GMM-Based Method; 二是直接采用 \mathbf{X}_{1W} 、 \mathbf{X}_{2W} 和 \mathbf{X}_{3W} 作为特征向量, 并以 SVM 作为分类器, 不妨称为 The SVM Classifier。

本文使用了 SVM 工具软件包 LibSVM^[21]. 核函数采用 Gauss 径向基核, 并通过经典的交叉验证法^[17]实现参数优化. 在使用文献[2]方法、频率检测、串行检测、Poker 检测和游程检测等方法进行隐写检测时, 设定的判决显著性系数同为 0.01。

4.2 实验结果

4.2.1 对 PD1 数据集的实验结果

我们将 PD1 数据集(图像大小为 512×512)分为两部分: 训练集 5000 幅, 测试集 3000 幅. 本实验中, 我们做了十组实验, 每一组实验从测试集中随机选取 100 幅图像进行检测分析. 表 2 为针对 LSB 匹配隐写的不同检测算法的十组实验平均检测率(%), 实验结果表明: 在判决显著性系数设定为 0.01 的条件下, 频率检测、串行检测、Poker 检测和游程检测的虚警率达到了较高的 27.4% ~ 65.3%; 另外, 本文算法在虚警率与文献[2]方法、The GMM-Based Method 和 The SVM Classifier 相当的情况下, 只有在嵌入率为 70% 的实验中检测正确率略低于 The SVM Classifier, 其余实验中的检测正确率均为最好. 表 3 为针对 LSB 替换隐写的十组实验平均检测率(%), 实验结果表明: 在判决显著性系数设定为 0.01 的条件下, 频率检测、串行检测、Poker 检测和游程检测的虚警率达到了较高的 35.9% ~ 62.4%; 另外, 本文算法在虚警率与文献[2]方法、The GMM-Based Method 和 The SVM Classifier 相当的情况下, 只有在嵌入率为 10% 的实验中检测正确率略低于 The SVM Classifier, 其余实验中的检测正确率均为最好。

4.2.2 对 PD2 数据集的实验结果

我们将 PD2 数据集(图像大小为 1024×1024)分为

两部分:训练集 4000 幅,测试集 2000 幅.本实验中,我们做了十组实验,每一组实验从测试集中随机选取 100 幅图像进行检测分析.表 4 为针对 LSB 匹配隐写的不同检测算法的十组实验平均检测率(%),实验结果表明:在判决显著性系数设定为 0.01 的条件下,频率检测、串行检测、Poker 检测和游程检测的虚警率达到了较高的

27.6%~58.1%;另外,本文算法在虚警率与文献[2]方法、The GMM-Based Method 和 The SVM Classifier 相当的情况下,只有在嵌入率为 40% 的实验中检测正确率略低于 The SVM Classifier,其余实验中的检测正确率均为最好.表 5 为针对 LSB 替换隐写的十组实验平均检测率(%),实验结果表明:在判决显著性系数设定为 0.01 的

表 2 针对 LSB 匹配隐写的不同检测算法在 PD1 数据集上的检测结果

Steganalysis Methods		Embedding Ratio										
		100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	Cover
1	The Proposed Method	99.2	98.3	96.6	86.5	69.2	55.9	36.4	24.1	22.7	17.3	13.3
2	The Method of Ref. [2]($l=2$)	88.9	86.5	85.3	77.9	50.0	33.4	26.8	19.2	15.7	10.0	13.1
3	The Method of Ref. [2]($l=3$)	88.4	86.5	85.4	79.4	51.0	35.6	27.0	19.5	16.8	10.1	13.3
4	The Method of Ref. [2]($l=4$)	88.5	87.7	86.3	80.1	52.3	36.7	27.7	20.1	16.3	11.1	13.4
5	The Method of Ref. [2]($l=5$)	90.2	88.9	87.2	79.9	52.2	36.5	27.0	20.2	16.3	10.6	13.0
6	The Method of Ref. [2]($l=6$)	88.7	87.9	87.9	80.7	53.5	37.0	28.2	20.0	16.8	10.7	13.3
7	The Method of Ref. [2]($l=7$)	89.9	90.4	87.4	80.8	55.5	37.9	28.2	20.2	16.6	11.7	13.1
8	The Method of Ref. [2]($l=8$)	89.5	88.0	87.4	81.9	59.3	38.3	28.3	20.4	17.1	12.3	13.2
9	The GMM-Based Method	92.7	90.9	90.5	80.7	61.3	50.2	35.8	20.3	17.5	12.5	13.7
10	The SVM Classifier	97.9	97.1	96.6	87.4	66.2	54.5	36.3	24.1	22.2	17.0	13.4
11	0,1 Frequency Test Method	100.0	92.5	87.7	85.9	83.6	79.5	76.1	72.4	70.3	68.2	65.3
12	Serial Test Method	99.6	90.4	86.9	80.8	76.1	62.5	54.9	53.0	51.8	55.5	54.9
13	Poker Test Method	91.6	90.9	89.3	81.2	74.7	46.6	38.4	36.8	34.4	30.1	29.3
14	Runs test Method	98.9	97.3	91.2	80.6	61.0	40.8	27.5	26.5	25.3	23.6	27.4

表 3 针对 LSB 替换隐写的不同检测算法在 PD1 数据集上的检测结果

Steganalysis Methods		Embedding Ratio										
		100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	Cover
1	The Proposed Method	100.0	99.5	97.3	86.7	72.6	58.7	37.9	25.0	22.4	16.2	11.3
2	The Method of Ref. [2]($l=2$)	94.3	92.3	91.2	73.3	52.3	36.3	29.5	19.4	15.9	10.1	11.1
3	The Method of Ref. [2]($l=3$)	94.7	93.6	92.3	73.9	53.5	37.2	29.5	19.8	16.4	11.0	11.6
4	The Method of Ref. [2]($l=4$)	92.9	92.6	91.8	74.1	53.8	37.9	29.8	20.0	16.3	11.3	11.4
5	The Method of Ref. [2]($l=5$)	90.0	89.1	90.5	74.8	54.9	38.1	29.7	20.2	16.6	11.6	11.0
6	The Method of Ref. [2]($l=6$)	91.3	87.5	90.7	75.3	56.1	38.7	30.1	20.2	16.6	11.7	11.3
7	The Method of Ref. [2]($l=7$)	88.7	86.3	88.6	76.1	61.3	39.7	30.2	20.3	16.9	11.6	11.1
8	The Method of Ref. [2]($l=8$)	89.3	87.4	86.2	77.7	62.2	40.2	30.5	21.1	16.9	11.8	11.2
9	The GMM-Based Method	96.4	93.7	89.7	80.9	66.3	47.8	29.5	20.3	15.8	14.5	12.4
10	The SVM Classifier	98.7	97.3	94.2	82.3	70.6	54.4	33.1	23.6	20.3	16.3	11.7
11	0,1 Frequency Test Method	100.0	94.9	90.5	87.5	84.3	78.3	71.4	65.5	64.5	63.6	62.4
12	Serial Test Method	99.7	94.8	90.6	86.6	81.1	68.0	60.3	54.2	52.8	51.8	50.7
13	Poker Test Method	91.8	95.8	90.9	84.3	77.1	57.6	48.3	43.6	41.6	40.2	39.4
14	Runs test Method	98.7	99.1	96.2	86.2	75.6	55.0	46.0	38.9	37.3	36.3	35.9

表 4 针对 LSB 匹配隐写的不同检测算法在 PD2 数据集上的检测结果

Steganalysis Methods		Embedding Ratio										
		100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	Cover
1	The Proposed Method	100.0	100.0	100.0	91.1	82.8	67.6	50.3	39.1	30.2	19.4	8.4
2	The Method of Ref. [2]($l=2$)	100.0	98.5	95.6	89.0	74.2	50.1	34.1	23.0	19.1	10.3	9.1
3	The Method of Ref. [2]($l=3$)	93.7	93.2	94.1	89.6	75.5	50.6	35.5	23.7	19.5	10.3	8.6
4	The Method of Ref. [2]($l=4$)	100.0	91.1	87.3	90.1	76.0	51.2	35.5	24.6	20.0	11.4	8.7
5	The Method of Ref. [2]($l=5$)	93.5	89.1	86.2	90.5	76.3	52.0	36.2	25.3	20.0	11.5	8.5
6	The Method of Ref. [2]($l=6$)	88.4	88.2	85.5	89.2	77.5	53.0	36.6	25.5	21.2	12.0	8.9
7	The Method of Ref. [2]($l=7$)	87.4	86.3	86.1	85.3	78.6	54.7	37.0	26.0	23.2	12.5	9.0
8	The Method of Ref. [2]($l=8$)	87.0	86.2	86.7	85.8	79.6	55.9	39.1	26.5	24.0	13.2	8.9
9	The GMM-Based Method	94.9	92.4	92.1	87.5	75.4	60.3	48.3	35.5	23.4	15.1	8.3
10	The SVM Classifier	97.3	95.0	94.7	89.3	81.1	64.4	52.5	37.2	29.5	18.1	8.8
11	0,1 Frequency Test Method	100.0	86.4	80.2	77.3	74.4	71.1	66.3	64.5	59.7	60.0	58.1
12	Serial Test Method	93.5	79.0	74.2	61.4	59.0	47.2	44.4	42.6	41.9	40.5	38.1
13	Poker Test Method	83.5	88.4	82.9	61.6	59.5	33.4	30.7	33.6	28.2	29.3	27.3
14	Runs test Method	100.0	96.1	85.8	73.9	52.3	33.6	31.8	28.2	28.2	26.4	27.6

表 5 针对 LSB 替换隐写的不同检测算法在 PD2 数据集上的检测结果

Steganalysis Methods	Embedding Ratio										
	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	Cover
1 The Proposed Method	100.0	100.0	100.0	92.4	81.3	64.6	48.1	38.5	29.3	18.3	8.3
2 The Method of Ref. [2]($l=2$)	95.3	95.5	93.3	87.1	70.4	48.0	28.2	22.4	20.9	11.2	8.1
3 The Method of Ref. [2]($l=3$)	98.4	97.1	90.7	91.9	74.0	48.5	30.8	23.0	21.3	12.0	8.5
4 The Method of Ref. [2]($l=4$)	96.7	95.8	91.7	90.6	75.0	50.2	32.3	23.4	21.5	12.4	8.2
5 The Method of Ref. [2]($l=5$)	95.1	93.4	93.5	90.7	76.6	51.2	32.7	24.1	22.0	12.5	8.8
6 The Method of Ref. [2]($l=6$)	95.1	91.7	91.9	87.2	77.1	52.3	33.1	24.3	22.2	12.7	8.2
7 The Method of Ref. [2]($l=7$)	92.6	92.2	89.3	84.6	77.8	53.7	34.6	25.7	23.1	13.1	8.7
8 The Method of Ref. [2]($l=8$)	90.5	91.2	86.6	84.8	78.2	54.8	35.5	27.1	24.6	13.9	8.4
9 The GMM-Based Method	95.2	94.0	92.7	90.2	76.6	58.3	42.2	33.6	24.6	14.5	9.1
10 The SVM Classifier	98.8	97.4	92.7	91.3	79.1	61.3	44.3	37.5	26.8	17.2	8.8
11 0,1 Frequency Test Method	100.0	86.5	81.4	77.4	74.6	71.5	66.2	64.1	59.3	60.0	58.2
12 Serial Test Method	94.3	80.8	77.5	63.6	58.0	46.5	45.1	43.1	41.5	40.3	38.7
13 Poker Test Method	88.0	87.6	83.8	62.5	58.3	36.3	35.6	33.7	31.0	30.1	27.7
14 Runs test Method	100.0	96.8	95.9	93.9	91.8	34.4	31.5	30.1	26.9	26.4	27.8

条件下,频率检测、串行检测、Poker 检测和游程检测的虚警率达到了较高的 27.8% ~ 58.2%;另外,本文算法在虚警率与文献[2]方法、The GMM-Based Method 和 The SVM Classifier 相当的情况下,所有嵌入率实验中的检测正确率均为最好。

4.3 性能分析

4.3.1 本文方法与其它检测方法的性能比较

在数据集 PD1 和 PD2 上的实验结果表明:(1)在判决显著性系数设定为 0.01 的条件下,频率检测、串行检测、Poker 检测和游程检测等序列随机性检测方法也取得了一定的检测正确率,但是这些方法的虚警率都偏高,甚至是“草木皆兵”。实际上,通过调整判决显著性系数,可以降低这些方法的虚警率,但同时这些方法的检测正确率也显著降低。实验表明,通过调整判决显著性系数,使这些序列随机性检测方法的虚警率与本文算法和文献[2]方法相当时,其检测正确率将明显低于本文算法和文献[2]方法(限于篇幅,实验结果本文不列出,类似结果可参见文献[2]中的表 1)。而本文方法与文献[2]方法、GMM 模型法、SVM 分类法的虚警率相当,均处在较低的可控范围,且有较好的检测正确率;(2)无论是对 LSB 匹配隐写,还是对 LSB 替换隐写,本文方法在不同嵌入率情形下都取得了较为明显的优于文献[2]方法的实验结果;(3)无论是对 LSB 匹配隐写,还是对 LSB 替换隐写,本文方法在不同嵌入率情形下都略优于 GMM 模型法和 SVM 分类法的实验结果。图 2 和图 4 为针对 LSB 匹配隐写和替换隐写技术,本文提出的检测方法和文献[2]方法(对不同的 l ,取最大平均检测率)、GMM 模型法、SVM 分类法等在三组数据集上

的平均检测率的直观比较。

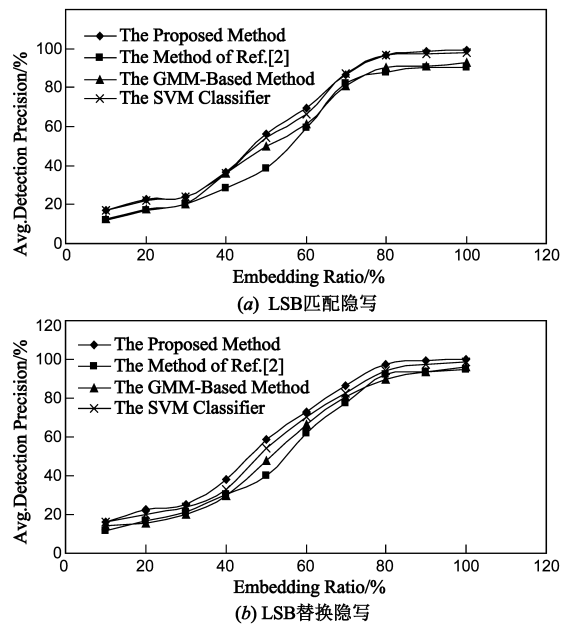


图2 针对不同的隐写技术,本文提出的检测方法和文献[2]方法(对不同的 l ,取最大平均检测率)、GMM模型法、SVM分类法等在三组数据集上的平均检测率比较

4.3.2 定理 2 的实验结果分析

对嵌入率为 10% ~ 70% 的隐写图像的检测率,有如下结果:文献[2]方法随着选择的短重码维数 l 的增大,检测率也增大,即直接选择 $X_8(W)$ 作为假设检验变量为最优选择。但对嵌入率为 80% ~ 100% 的隐写图像的检测率,得到的实验结果与检测嵌入率为 10% ~ 70% 的隐写图像的结果不同。这些实验结果进一步论证了定理 2:对嵌入率较低(如:10% ~ 70% 等)的隐写图

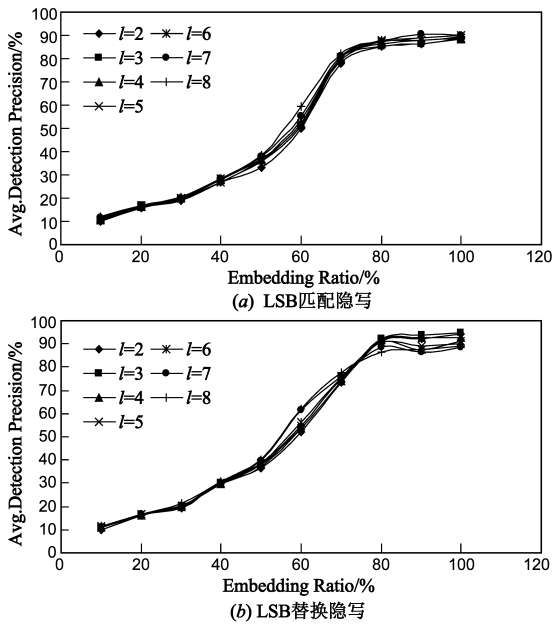


图3 针对不同的隐写技术,文献[2]检测方法对不同 l 在PD1数据集上的平均检测率比较

像,检验变量 $X_l(W)$ 服从 Poisson 分布的近似度高,用文献[2]方法检测隐写图像时,若对短重码维数 l 的选择按从大到小顺序,则在相同参数阈值约束下,可较大地减少检测次数(或仅需一次),从而减少运算时间.针对不同的隐写技术和不同的数据集,文献[2]检测方法对不同 l 的平均检测率的直观比较见图3和图5.

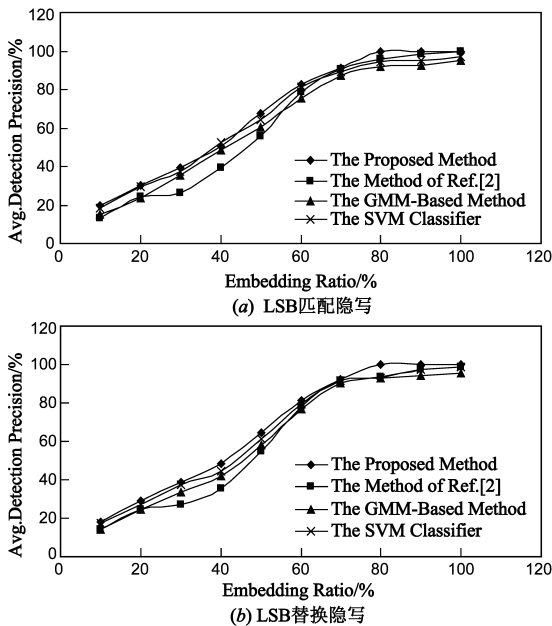


图4 针对不同的隐写技术,本文提出的检测方法和文献[2]方法(对不同的 l ,取最大平均检测率)、GMM模型法、SVM分类法等,在PD2数据集上的平均检测率比较

4.3.3 两种数据集上的实验结果比较

数据集 PD1 和 PD2 分别包含了 512×512 、和 1024×1024 不同大小的图像.本文涉及的通用隐写检测方

法所采用的特征向量是基于统计的,而更大的图像的统计特征更为明显.故无论是从本文方法与其它检测方法的性能比较,还是从定理2的实验结果中基本体现了较大图像的统计优势,也进一步验证了本文结论和本文提出方法的有效性.

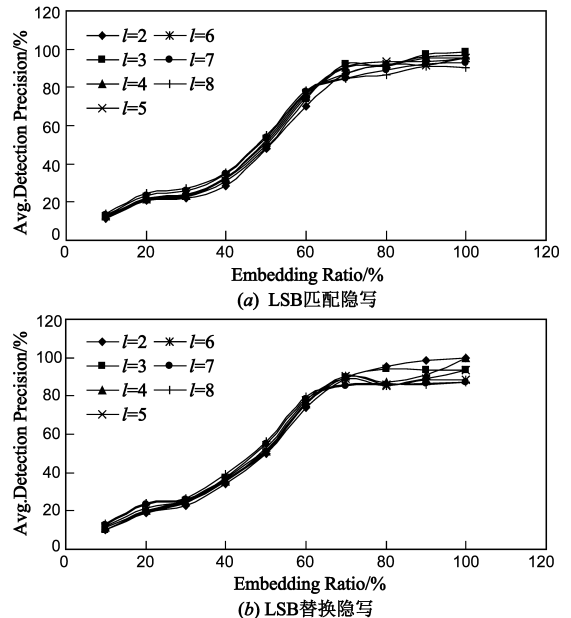


图5 针对不同的隐写技术,文献[2]检测方法对不同 l 在PD2数据集上的平均检测率比较

5 结论

本文提出了一种基于 LSB 序列局部特征的通用隐写检测方法.它有如下特点:

- (1)局部特征(LSBSLF)提取算法的原始处理对象为图像的 LSB 序列;
- (2)基于全局序列词汇设计了融合 GMM 生成模型与判别方法的分类器;
- (3)采用多元统计分析思想考虑多个特征间存在的联系.为如下需求提供了一种提高性能的一般性解决框架:为适应不同的应用场合,需要选择适当的特征变量作检测分类.

另一方面,本文分析证明了单个短重码间距统计变量的检测能力规律,提出了文献[2]方法中合理选择短重码维数的方法,对嵌入率为 10%-70% 的隐写图像的检测有效.

本文方法继承了文献[2]方法具有的通用性等优点.但这两种方法一样,依然存在如下不足^[2]:若隐写算法不在载体图像位平面自相关性强的区域嵌入信息,则检测算法失效;另外,检测性能与自然图像 LSB 序列本身的分布有关,若自然图像位平面本身很近似服从均匀分布,则检测算法失效.实际上,如何降低载

体图像内容本身对通用检测性能的影响一直是个公开课题^[22].

参考文献

- [1] 王朔中,张新鹏,张卫明.以数字图像为载体的隐写分析研究进展[J].计算机学报,2009,32(7):1247-1263.
Wang Shuo-zhong, Zhang Xin-peng, et al. Recent advances in image-based steganalysis research[J]. Chinese Journal of Computers, 2009, 32(7): 1247-1263. (in Chinese)
- [2] 王国新,平西建,许漫坤,等.一种基于短重码间距统计的隐写分析方法[J].中国科学 F 辑:信息科学,2009,39(4):416-421.
Wang Guo-xin, Ping Xi-jian, Xu Man-kun, et al. A steganalysis method based on spacing statistics of short duplicate code [J]. Science in China Series F: Information Sciences, 2009, 39(4): 416-421. (in Chinese)
- [3] K Mikolajczyk, C Schmid. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [4] J Zhang, M Marszalek, S Lazebnik, et al. Local features and kernels for classification of texture and object categories: A comprehensive study[J]. International Journal of Computer Vision, 2007, 73(2): 213-238.
- [5] 高常鑫,桑农.整合局部特征和滤波器特征的空间金字塔匹配模型[J].电子学报,2011,39(9):2034-2038.
Gao Chang-xin, Sang Nong. Unifying local features and filter-bank features in the spatial pyramid matching model[J]. Acta Electronica Sinica, 2011, 39(9): 2034-2038. (in Chinese)
- [6] Mario Fritz, Bastian Leibe, Barbara Caputo. Integrating representative and discriminant models for object category detection [A]. Proceeding of the 10th IEEE Int Conf on Computer Vision [C]. Beijing: IEEE Computer Society, 2005. 1363-1370.
- [7] Gabriella Csurka, Christopher R. Dance, Lixin Fan, et al. Visual categorization with bags of keypoints [A]. Proceeding of the ECCV Int Workshop on Statistical Learning in Computer Vision [C]. Prague, Czech Republic: Springer, 2004. 59-74.
- [8] Blei D, Ng A, Jordan M. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3(5): 993-1022.
- [9] Wikipedia. The free encyclopedia: Mixture model [EB/OL]. http://en.wikipedia.org/wiki/Mixture_model, 2011-11-3.
- [10] J Farquhar, S Szedmak, Hongying Meng, et al. Improving "bag-of-keypoints" image categorization [R]. Technical report, University of Southampton, 2005.
- [11] 郭立君,赵杰煜,史忠植.生成模型与判别方法相融合的图像分类方法[J].电子学报,2010,38(5):1141-1145.
Guo Li-jun, Zhao Jie-yu, et al. Image categorization of Integrated generative models and discriminative methods[J]. Acta Electronica Sinica, 2010, 38(5): 1141-1145. (in Chinese)
- [12] S Theodoridis, K Koutroumbas. Pattern Recognition [M]. Singapore: Elsevier Pte Ltd, 2006. 149-150.
- [13] CHENG Yong-qing, ZHUANG Yong-ming, YANG Jing-yu. Optimal fisher discriminant analysis using the rank decomposition[J]. Pattern Recognition, 1992, 25(1): 101-111.
- [14] USC SIPI. The USC-SIPI Image database [EB/OL]. <http://sipi.usc.edu/services/database/database.html>, 2011-11-20.
- [15] 陈铭,张茹,刘凡凡,等.基于区域相关性的 LSB 匹配隐写分析[J].通信学报,2010,31(3):1-11.
Chen Ming, Zhang Ru, Liu Fan-fan, et al. Steganalysis of LSB matching based on regional correlation [J]. Journal on Communications, 2010, 31(3): 1-11. (in Chinese)
- [16] M. Figueiredo, A. Jain. Unsupervised learning of finite mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(3): 381-396.
- [17] 邓乃扬,田英杰.支持向量机-理论,算法与拓展[M].北京:科学出版社,2009. 81-111.
- [18] 王朔中,张新鹏,张开文.数字密写和密写分析[M].北京:清华大学出版社,2005. 20-23.
- [19] Menezes M, Oorschot P V, Vanstone S. Handbook of Applied Cryptography [M]. CRC Press, 1996. 169-187.
- [20] 赵力.语音信号处理[M].北京:机械工业出版社,2010. 228-230.
- [21] Chih-Chung Chang, Chih-Jen Lin. LIBSVM-A library for support vector machines [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2011-10-3.
- [22] LUO Xiang-yang, WANG Dao-shun, WANG Ping, et al. A review on blind detection for image steganography [J]. Signal Processing 88, 2008: 2138-2157.

作者简介



钟尚平 男,1969 年出生,博士,副教授.主要研究方向为网络信息安全、模式识别等.
E-mail: spzhong@fzu.edu.cn



徐巧芬 女,1989 年出生,硕士研究生.主要研究方向为多媒体安全、模式识别等.