基于KFD指标聚类的高隐蔽性JPEG隐写分析

黄 炜[1),2)] 赵险峰[2)] 盛任农[3)]

[1)](中国科学院软件研究所 北京 100190)

[2)](中国科学院信息工程研究所信息安全国家重点实验室 北京 100195)

[3)](北京电子技术应用研究所 北京 100191)

摘 要 采用非公开的图像源或算法的隐写行为具有很强的隐蔽性.在这类对隐写者先验不足的场景下聚类分析更为实用.Ker等人比较不同指标不同配置之后,提出基于MMD指标聚类的隐写者识别方法.然而该方法所用MMD指标只考虑两个类样本中心之间的距离,忽略了样本相对中心点的聚合程度对可分性的影响,因而准确率存在提高的空间.为进一步提高现有隐写聚类分析方法的准确率,该文提出用核Fisher鉴别(KFD)指标计算样本间差异度量的聚类方法.首先,提取PEV274校准特征并归一化.然后,计算KFD指标组成距离矩阵.最后,根据样本间差异度量矩阵按重心法自底向上进行层次聚类分析.KFD指标兼顾与最大平均距离(MMD)原理相近的类间方差以及指示样本聚集程度的类内方差,更准确地估算样本间差异.实验结果表明,该文对低嵌入率隐写其准确率最高提高约30%,对高嵌入率准确率降低不超过5%.该文的创新点在于提出了一种更合理的指标和基于该指标聚类隐写分析的方法,比现有方法平均准确率有一定的提高.

关键词 核Fisher鉴别指标;聚类分析;隐写分析;类内方差

中图法分类号 TP309 **DOI号**: 10.3724/SP.J.1016.2012.01951

Steganalysis Based on Clustering via KFD Index Against Highly Undetectable JPEG Steganography

HUANG Wei[1),2)] ZHAO Xian-Feng[2)] SHENG Ren-Nong[3)]

[1)](*Institute of Software, Chinese Academy of Sciences, Beijing* 100190)

[2)](*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100195)

[3)](*Beijing Institute of Electronic Technology and Application, Beijing* 100191)

Abstract It is highly undetectable of steganographers who avoid utilizing public sources of images or steganographic schemes. In such scenario where steganalysers have little priori knowledge about steganographers, clustering is more practical. Ker proposed a MMD-based clustering scheme to distinguish steganographers from innocent actors after comparisons in various configurations and indexes. MMD merely considers the distance between centers of samples from two classes, but ignores the fact that aggregation how samples gather around their centers does affect the separability. Hence, its accuracy needs improvement. To increase the detecting rate further, we propose a clustering based steganalytic scheme using kernel Fisher discriminant indexes (KFDI) as the dissimilarities of samples. We firstly extract the calibration features PEV274 and have them normalized. Then, we calculate the KFD indexes between samples to form the distance matrix. Finally, hierarchical clustering is proceeded with bottom-up iteration where we used the

收稿日期:2012-05-09;最终修改稿收到日期:2012-07-09.本课题得到国家自然科学基金(61170281)、北京市自然科学基金(4112063)、中国科学院战略性先导专项课题(XDA06030601)及中国科学院信息工程研究所创新课题(Y1Z0041101,Y1Z0051101)资助.**黄 炜**,男,1985年生,博士研究生,主要研究方向为隐写分析.E-mail: weihuang@is.iscas.ac.cn. **赵险峰**,男,1969年生,博士,副研究员,主要研究方向为信息隐藏. **盛任农**,男,1969年生,硕士,副研究员,主要研究方向为信号处理.

center of gravity as the center for the new gathered clusters. KFDI considers not only between-class variances that maximum mean discrepancy concentrates on, but also within-class variance that affects the aggregation between classes. Experimental results show that our scheme obtains a high increase in accuracy under low embedding rates, about 30% at most, but a little decrease of no more than 5% under high embedding rates. The key contribution of this paper is to propose a more reasonable indicators and steganalytic method based on the KFDI, and we raised the average accuracy of existing methods.

Keywords kernel Fisher discriminant index; clustering; steganalysis; within-class variances

1 引 言

隐写分析(steganalysis)[1]作为检测隐写行为存在性的一种手段,近年来受到广泛的研究.JPEG图像是最常见的多媒体之一,因而JPEG图像隐写研究具有较好的实用意义.现有JPEG隐写方法包括最低有效位(Least Significant Bits,LSB)、量化方法和在此基础上的改进方法.JPHide[7]基于查表动态选择嵌入位置提高了LSB隐写的隐蔽性.扰动量化(Perturbed Quantization,PQ)[5]选取失真较小的系数进行嵌入,从而达到减小扰动量的目的;PQe[6]在PQ基础上做出改进,选择在DCT能量高的地方嵌入,进一步减少了扰动量.F5[3]通过矩阵编码技术减少改动位置;修改的矩阵编码(Modified Matrix Encoding,MME)[4]进一步寻找失真最小的位置组合进行嵌入,进一步降低扰动量;nsF5[6]在F5基础上进一步改进,消除DCT直方图收缩(shrinkage)效应,是目前隐蔽性最强的JPEG隐写算法之一.

现有隐写分析多基于一组通用隐写特征和一种分类器[8],用已知样本训练后对待测样本进行二类分类,判断待测样本隐写行为存在性.其中,PEV274特征[14]利用校准技术估计原图,对图像裁剪4行4列重做JPEG压缩,并提取了7组统计特性,是目前维度较低且分析能力最强的特征.二类隐写分析可以扩展为多类隐写分析[2],有效地适用已知隐写算法和样本来源(或其范围)的情形.但是,该方法具有两点局限性:(1)在隐写算法范围未知的情况下,分类准确率下降明显;(2)当训练集和测试集样本来源不一致(如训练集和测试集来自不同相机)时,会出现"载体源不匹配"(cover-source mismatch)[10]的情况,准确率大大下降.有经验的隐写者可以利用该局限性来逃避隐写分析,即尽量避免采用公开的算法和图像源,以隐藏秘密信息的存在性.这一类的隐写行为隐蔽性很高,是隐写分析的难点之一.

相对二类或多类隐写分析,在未知算法和图像源范围的场景下的隐写分析方法却少有关注.针对此类高隐蔽性隐写,分析者无法得到适合待测样本的载体和隐写特征,训练和分类过程无从实现,此时,聚类分析(cluster analysis)可以解决这类问题[11].由于隐写是对图像的轻微修改,该环境下截获的样本往往较少,统计上不够稳定.因此,聚类方法并不判断单个样本的隐写存在性,而是找出参与者(actor)的隐写行为存在性.即假定截获的图像来自几个参与者,每个参与者发送一定量的图像,其中部分含有或者完全不含秘密信息.通过聚类分析,可以找出可能有隐写的参与者,以进一步跟踪分析.相对传统隐写分析,该环境有3个特点:

(1)截获样本少;

(2)样本的来源未知,无法获取更多同源样本;

(3)隐写样本的算法及嵌入率未知.

Ker等人[11]提出一种基于最大平均差异(Maximum Mean Discrepancy,MMD)[12,18]准则的隐写聚类分析方法.该方法提取不同参与者所传递图像的PEV274特征[14],将同一个参与者传递的图像整体视为聚类中的一个样本,计算MMD作为样本间差异度量(dissimilarity measure),自底向上进行层次聚类(hierarchical clustering)[9]得到最可能的隐写者.然而,MMD所指示的是两个类的样本中心点在所有可能空间的最大距离,它不能很好地表达两个类间的差异程度.两个类之间是否可以区分,不仅与类间分散程度有关,还受到类内聚合程度的影响.该方法未能考虑类内聚合程度因素,计算样本间差异度量不够准确,其准确率存在提高的空间.

为了进一步提高对高隐蔽性隐写的聚类分析效果,本文提出了一种基于核Fisher鉴别(Kernel Fisher Discriminant,KFD)[13]指标的隐写聚类分析

方法，兼顾样本在类间与类内两方面的聚合程度，以提高隐写分析检测能力. 实验结果表明，该方法在低嵌入率下对隐写者的识别准确率有较大的提升，最高约 30%，同时对高嵌入率隐写的准确率下降不超过 5%. 总体上，KFD 指标(KFD Index，KFDI)更好地反应了两个类的间距，该方法在已知条件较少的情况下能更有效地检测出真正的隐写者.

本文第 2 节介绍相关工作；第 3 节介绍 KFDI 及本文分析算法框架，并比较 KFDI 与 MMD 指标；第 4 节对比本文方法与文献[11]方法，进行实验及实验分析；第 5 节为总结.

2　相关工作

2.1　问题定义

Cachin[1] 提出隐写可以视为一种信息论模型. 令 C 为载体样本集，隐写过程可以视为一个映射关系 $\mathrm{Emb}: C\times E\times D\to S$，其中 E 为秘密信息集，D 为密钥集，S 为隐写样本集，$C, S\subset I$，I 为样本集. 通过 $s=\mathrm{Emb}(c,m,k)$ 将任意 $c\in C, m\in E, k\in D$ 映射到一个隐写样本 $s\in S$. 接受方通过一个对应的映射 $\mathrm{Ext}: S\times D\to E$，获得消息 $m=\mathrm{Ext}(s,k)$.

现有的二类隐写分析通常随机选择一定数量的三元组 (c_i, m_i, k_i) 制备隐写样本 s_i，并提取隐写特征. 隐写特征提取也可以视为一个映射关系 $\mathrm{Fea}: I\to F$，其中 F 为特征集. 对于载体样本，通过 $x=\mathrm{Fea}(c), x\in X$，可以提取载体特征集 X，类似地，对于隐写样本有 $y=\mathrm{Fea}(s), y\in Y$，对于待测样本 $t\in T\subset I$，特征 $z=\mathrm{Fea}(t), z\in Z$. 使用支持向量机等机器学习方法可以利用已知 $x=\{x_1, x_2, \cdots, x_{n_x}\}$ 和 $y=\{y_1, y_2, \cdots, y_{n_y}\}$，判断待测特征 $z=\{z_1, z_2, \cdots, z_n\}$ 是否含有秘密信息. 当 X, Y 与 Z 在选用隐写方法及图像来源方面差异较大时，传统二类或多类隐写分析准确率较低[10].

由于不同型号图像采集设备(如相机或扫描仪等)对图像的成像及加工的过程各有不同，在上述对 T 知识较少的情况下，无法找到与待分析样本来源相近的样本，通过单个样本的判断来识别隐写者变得困难. 不妨假设 T 来自于 N 个参与者 $A_1, A_2, \cdots, A_N$. 其中，有且仅有一个参与者 A_i，其样本数量为 n_i 中有 $\lfloor a\cdot n_i\rfloor$ 个样本含有秘密信息(a 为嵌入密度)，这些样本的嵌入率为 r，其余 $n_i-\lfloor a\cdot n_i\rfloor$ 个样本没有隐写. 另外 $N-1$ 个参与者完全传递载体样本. 本文目的在于仅已知 $A_1, A_2, \cdots, A_N$ 的所有样本情况下，识别最可能的隐写者 A_i.

2.2　隐写聚类分析

隐写聚类分析是指通过将同一参与者的隐写样本特征作为一个对象，利用聚类方法找出隐写者的隐写分析方法. 隐写聚类分析过程中，由于载体之间差异(dissimilarity)或距离(distance)较小，最终划为同一个簇(cluster)；同样地，具有一定嵌入强度的隐写样本与载体样本差异较大，最终划为另一簇. 最后，得到可能的隐写者.

隐写聚类分析方法相较于传统二类或多类隐写分析以及一般的聚类分析，存在以下几个特点：

(1) 隐写聚类分析可用的先验知识较少. 由于载体来源和隐写算法范围未知，分析者无法制备类似的样本用于训练；

(2) 维度灾难(Curse of Dimensionality，CoD)[10]. 由于截获同一参与者图像往往较少，很难达到特征数量的数十倍[9]. 这也限定了特征的维度应该尽量较低；

(3) 隐写聚类分析识别隐写者而不是单个隐写样本. 隐写对自然图像修改很少，对单个样本的判断结果并不稳定. 如果对单个图像进行聚类，图像内容差异等因素很可能代替隐写行为存在性，成为影响聚类的主要因素. 因而，隐写聚类分析把同一参与者的样本视为一个整体来判断隐写存在性.

2.3　基于 MMD 的隐写聚类分析

MMD 是度量两个类样本均值在所有映射空间下最大距离的量，其计算公式为

$$\mathrm{MMD}[F,X,Y]=\sup_{f\in F}(E_{x\sim p}f(x)-E_{y\sim q}f(y)).$$

其中，F 是函数 $f: X\to R$ 的集合，R 是实数集，$E_{x\sim p}$ 表示变量 x 服从 p 分布时的期望，$E_{y\sim q}$ 表示变量 y 服从 q 分布时的期望. 由于 MMD 无法直接计算，文献[18]提供了一个近似算法.

Ker 等人经过大量实验比较，提出目前隐写聚类分析中效果较好的 MMD 隐写聚类分析方法. 该方法将 MMD 作为样本间差异度量的计算标准，定义 $\mathrm{MMD}[F, A_i, A_j]$ 为 A_i, A_j 两类之间在函数集 F 下的距离. 接着，进行层次聚类，自底向上不断将样本间差异度量最近的两个类合并，最后找出可能的隐写者. Ker 等人[11] 通过实验比较，计算类间差异时，重心法(centroid clustering)效果最好.

3　基于 KFDI 的隐写聚类分析方法

本节介绍类间方差与类内方差，并将其扩展至

再生核 Hilbert 空间(Reproducing Kernel Hilbert Spaces,RKHS),用核方法求解,称为 KFD 指标(KFDI).接着,从原理上将 KFDI 和 MMD 指标做比较.最后,给出本文方法的实现流程.

3.1 类间方差与类内方差

令 $X_1=\{\boldsymbol{x}_1^{(1)},\boldsymbol{x}_2^{(1)},\cdots,\boldsymbol{x}_{D_1}^{(1)}\}$,$X_2=\{\boldsymbol{x}_1^{(2)},\boldsymbol{x}_2^{(2)},\cdots,\boldsymbol{x}_{D_2}^{(2)}\}$为两个不同分类的样本,并有样本集整体$X=X_1\cup X_2=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_D\}$,则样本均值可以定义为

$$\boldsymbol{\mu}_i=\frac{1}{D_i}\sum_{\boldsymbol{x}\in X_i}\boldsymbol{x},\quad i=1,2.$$

其对于给定向量 $\boldsymbol{w}$ 可被映射为

$$\tilde{\mu}_i=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\mu}_i=\frac{1}{D_i}\sum_{\boldsymbol{x}\in X_i}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x},\quad i=1,2.$$

类间方差(between-class variances)SS_B作为类间分散程度的平方和,其计算公式为

$$\begin{aligned}SS_B&=(\tilde{\mu}_1-\tilde{\mu}_2)^2\\&=\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^{\mathrm{T}}\boldsymbol{w}=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_B\boldsymbol{w}\end{aligned}\quad(1)$$

其中,$\boldsymbol{S}_B=(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^{\mathrm{T}}$称为类间散度矩阵(between-class scatter matrix).

图 1(a)显示了二维情况下的类间方差.图中圆圈代表类 1 的样本,方框代表类 2 的样本,实心圆圈和方框分别是各类的中心点.在向量 $\boldsymbol{w}$ 上的标记分别是这些样本及其中心在该向量上的投影.从图中可以看出,类间方差在作为类 1 和类 2 的差异性大小的评价指标时是非常重要的.一般 SS_B 越大,表明它们的可区分性越高.

类内方差(within-class variances)SS_W 作为类内散度的平方和,其计算公式为

$$\begin{aligned}SS_W&=\sum_{i=1,2}\sum_{\boldsymbol{x}\in X_i}[\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu}_i)]^2\\&=\sum_{i=1,2}\boldsymbol{w}^{\mathrm{T}}\Big[\sum_{\boldsymbol{x}\in X_i}(\boldsymbol{x}-\boldsymbol{\mu}_i)(\boldsymbol{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\Big]\boldsymbol{w}=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_W\boldsymbol{w}\end{aligned}\quad(2)$$

其中,

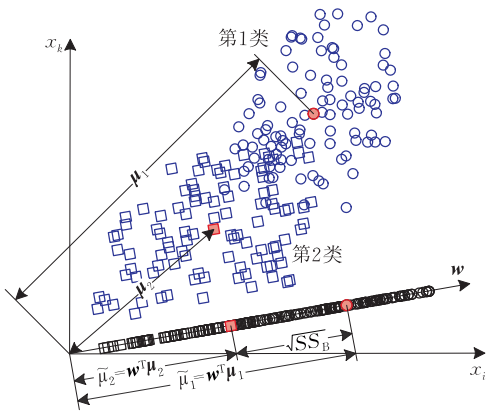
$$\boldsymbol{S}_W=\sum_{i=1,2}\sum_{\boldsymbol{x}\in X_i}(\boldsymbol{x}-\boldsymbol{\mu}_i)(\boldsymbol{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}$$

称为类内散度矩阵(within-class scatter matrix).

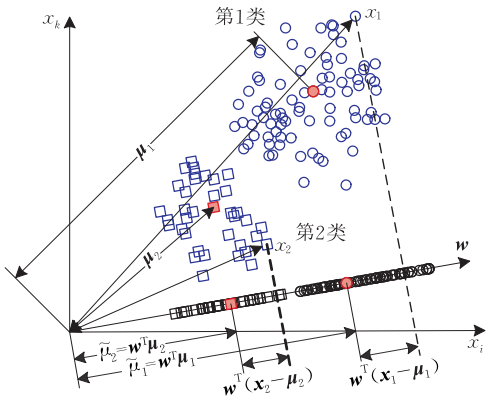
图 1(b)显示了二维情况下的类内方差.可以看出类内方差表明了两个类的内聚性,即两个类的样本对其中心的聚合程度.当 SS_B 一定时,SS_W 越小,界限越清楚,可分离性就越好.

在线性情况下,Fisher 鉴别指标为类间方差和类内方差在向量 $\boldsymbol{w}$ 上的比值,即

$$J(\boldsymbol{w})=\frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_B\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_W\boldsymbol{w}}\quad(3)$$



(a) 类间方差示意图



(b) 类内方差示意图

图 1 KFDI 类间方差与类内方差示意图

3.2 KFD 指标

现实世界的数据复杂度高,线性鉴别的能力往往有限.核(kernel)的思想[15]最初用于支持向量机(Support Vector Machine,SVM)[8]和其它算法.下面介绍基于核的鉴别分析及 KFDI.

令 Φ 作为到空间 F 的非线性映射.在空间 F 的线性鉴别,即是最大化

$$J(\boldsymbol{w})=\frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_B^{\Phi}\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_W^{\Phi}\boldsymbol{w}}.$$

式(1)和(2)被替代为

$$\boldsymbol{S}_B^{\Phi}=(\boldsymbol{\mu}_1^{\Phi}-\boldsymbol{\mu}_2^{\Phi})(\boldsymbol{\mu}_1^{\Phi}-\boldsymbol{\mu}_2^{\Phi})^{\mathrm{T}},$$

$$\boldsymbol{S}_W^{\Phi}=\sum_{i=1,2}\sum_{\boldsymbol{x}\in X_i}(\Phi(\boldsymbol{x})-\boldsymbol{\mu}_i^{\Phi})(\Phi(\boldsymbol{x})-\boldsymbol{\mu}_i^{\Phi})^{\mathrm{T}}.$$

其中,$\boldsymbol{\mu}_i^{\Phi}=\frac{1}{D_i}\sum_{j=1}^{D_i}\Phi(\boldsymbol{x}_j^{(i)})$.

如果 F 维度很高,甚至是无限的,上式就不可能直接求解.考虑到只计算$\langle\Phi(\boldsymbol{x})\cdot\Phi(\boldsymbol{y})\rangle$,而不关心 $\Phi(\boldsymbol{x})$或 $\Phi(\boldsymbol{y})$的具体数值,因此,不需将数据显式映射到 F,可以转而寻求某种算法,直接计算点乘$\langle\Phi(\boldsymbol{x})\cdot\Phi(\boldsymbol{y})\rangle$.为此,研究人员引入核的概念,定义核为一个对称正定有限矩阵的函数 $k:X\times X\mapsto R$.

令 F 为一个 RKHS，为求得 RKHS F 中的 $J(\boldsymbol{w})$，只需找到式(3)只包含点乘公式的一种形式. 在将点乘替换为一个特定的核函数，即令 $\langle \Phi(\boldsymbol{x}) \cdot \Phi(\boldsymbol{y}) \rangle = k(\boldsymbol{x},\boldsymbol{y})$. 函数 k 可以是高斯函数或其它形式. 由可再生核的理论可知，任意解 $\boldsymbol{w} \in F$ 必须存在于 F 的训练样本范围之内. 因此，可以找到一个 $\boldsymbol{w}$ 的展开形式

$$\boldsymbol{w} = \sum_{i=1}^{D} \alpha_i \Phi(\boldsymbol{x}_i).$$

根据 $\boldsymbol{\mu}_i^{\Phi}$ 的定义，可得

$$\boldsymbol{w}^{\mathrm{T}} \boldsymbol{\mu}_i^{\Phi} = \frac{1}{D_i} \sum_{j=1}^{D} \sum_{k=1}^{D_i} \alpha_j \langle \Phi(\boldsymbol{x}_j) \cdot \Phi(\boldsymbol{x}_k^{(i)}) \rangle = \frac{1}{D_i} \sum_{j=1}^{D} \sum_{k=1}^{D_i} \alpha_j k(\boldsymbol{x}_j, \boldsymbol{x}_k^{(i)}) = \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{M}_i.$$

其中，点乘被核函数替代，得到

$$(\boldsymbol{M}_i)_j = \frac{1}{D_i} \sum_{k=1}^{D_i} \alpha_j k(\boldsymbol{x}_j, \boldsymbol{x}_k^{(i)}).$$

则式(2)可被重新写为

$$SS_B^{\Phi} = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{S}_B^{\Phi} \boldsymbol{w} = \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{\alpha} \tag{4}$$

其中，$\boldsymbol{M} = (\boldsymbol{M}_1 - \boldsymbol{M}_2)(\boldsymbol{M}_1 - \boldsymbol{M}_2)^{\mathrm{T}}$. 与上式相似，可得

$$SS_W^{\Phi} = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{S}_W^{\Phi} \boldsymbol{w} = \boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{N} \boldsymbol{\alpha} \tag{5}$$

其中 $\boldsymbol{N} = \sum_{j=1,2} \boldsymbol{K}_j (\boldsymbol{I} - \boldsymbol{1}_{D_j}) \boldsymbol{K}_j^{\mathrm{T}}$，$\boldsymbol{K}_j$ 是一个 $l \times D_j$ 的矩阵，$(\boldsymbol{K}_j)_{n,m} = k(x_n, x_m^{(j)})$，$\boldsymbol{I}$ 为单位矩阵，$\boldsymbol{1}_{D_j}$ 是元素全为 $1/D_j$ 的矩阵.

结合式(4)、(5)，KFD 度量可通过最大化

$$J(\boldsymbol{\alpha}) = \frac{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\mathrm{T}} \boldsymbol{N} \boldsymbol{\alpha}} \tag{6}$$

得到. 在有限样本的情况下，KFDI 定义为类间方差和类内方差在任意映射下的最大比率，即为 RHKS F 下的 $J(\boldsymbol{\alpha})$ 的最大值.

式(6)可通过求解 $\boldsymbol{N}^{-1}\boldsymbol{M}$ 的最大特征值得到. 因而，KFDI 的计算公式为

$$\mathrm{KFDI}[F, X, Y] = v(\boldsymbol{N}^{-1}\boldsymbol{M}),$$

其中，$v(\boldsymbol{A})$ 为矩阵 $\boldsymbol{A}$ 的最大特征值.

如果 $\boldsymbol{N}$ 非正定，可能无法求得 $\boldsymbol{N}^{-1}$. 此时，可以把 $\boldsymbol{N}$ 和多个单位矩阵相加，例如，可以使用 $\boldsymbol{N}' = \boldsymbol{N} + u \cdot \boldsymbol{I}$ 代替 $\boldsymbol{N}$(建议取 $u = 0.1$). 该方法可视为对 $\boldsymbol{\alpha}$ 的范数 $\|\boldsymbol{\alpha}\|^2$ 做规范化(normalization)，降低基于样本估计的特征值所带来的偏差[13].

3.3　KFDI 和 MMD 比较

KFDI 计算类间方差与类内方差的比值在所有映射中的最大值. 其中，类间方差是两个类的中心点在其投影方向的差值平方. MMD 计算的是两个类中心点在所有映射空间上的距离最大值，其物理意义相当于类间方差.

本文从 BOSSBase 图库①随机选取 1000 张图像作为载体，计算不同算法嵌入率下载体与隐写样本间的 MMD 与 KFDI(表 1). 从表中可以看出，MMD 与类间方差在数值上存在一定程度的关系，一般 MMD 越大，类间方差也越大.

表 1　不同算法嵌入率下 MMD 与 KFDI 对照(1000 对样本)

算法	嵌入率/bpac	MMD	类间方差	类内方差	KFDI
JPHide[7]	0.05	0.00641	0.00042	0.18002	0.00231
	0.10	0.01348	0.00085	0.20269	0.00420
	0.15	0.03381	0.00192	0.17544	0.01095
	0.20	0.07652	0.00412	0.15862	0.02597
	0.30	0.16211	0.01021	0.12998	0.07854
MME3[4]	0.05	0.00016	0.00002	0.07583	0.00022
	0.10	0.00280	0.00021	0.13262	0.00160
	0.15	0.00568	0.00044	0.14264	0.00311
	0.20	0.05339	0.00398	0.16342	0.02433
	0.30	0.40974	0.05248	0.08361	0.62767
PQe[6]	0.05	0.00014	0.00003	0.29449	0.00010
	0.10	0.00050	0.00011	0.23891	0.00044
	0.15	0.00107	0.00023	0.20833	0.00112
	0.20	0.00179	0.00042	0.20399	0.00204
	0.30	0.00918	0.00093	0.15838	0.00584

然而，不同的类对其中心的聚合程度并不一样. 当两个类的样本间差异度量相同时，样本等中心聚合程度大的类，其区分度也较高. 类内方差指示的即是样本对中心的聚合程度. 图 1(a)与(b)具有相同的类间方差，但图 1(b)中类内方差较小，其区分度明显高于图 1(a). 从表 1 中也可以看出，类内方差在一定范围内浮动，起到一定的补充作用.

将类间方差除以类内方差，有利于突出聚合程度较高的类相对于其它类之间的分离程度，因而在度量样本间差异度量上具有更好的适用性.

3.4　本文分析流程

本文提出的基于 KFDI 的隐写聚类分析方法，其实现流程如图 2 所示. 具体步骤如下.

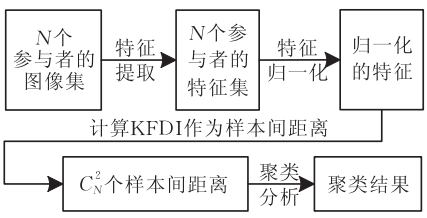


图 2　基于 KFDI 聚类的 JPEG 隐写分析实现流程

① BOSSBase v1.0 Image base. ftp://camnepserver.felk.cvut.cz/RAWs/. 2011

(1)对 N 个参与者的图像集 $\boldsymbol{A}=\{A_1,A_2,\cdots,A_N\}$ 提取 PEV-274 特征[14],得到特征集 $\boldsymbol{z}=\{\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_N\}$,其中 $\boldsymbol{z}_i$ 为 $n_i\times d$ 的矩阵,PEV274 的特征维度 $d=274$;

(2)计算特征集 $\boldsymbol{z}$ 每一维特征的均值 μ_d 和样本标准差 s_d,按下式对特征矩阵规范化:

$$\boldsymbol{z}'_i=\begin{cases}\dfrac{\boldsymbol{z}_i-\mu_d}{s_d}, & s_d\neq 0\\ 0, & s_d=0\end{cases};$$

(3)计算隐写特征集 $\boldsymbol{z}'$ 两两之间的样本间差异度量 $d_{i,j}=\mathrm{KFDI}[F,\boldsymbol{z}'_i,\boldsymbol{z}'_j]$,得到 N^2 个样本间差异度量,选用高斯内核 $k(\boldsymbol{x},\boldsymbol{y})=\exp(-\gamma(\boldsymbol{x}-\boldsymbol{y})^2)$,其中,$\gamma=-\eta^2$,$\eta$ 为 $\boldsymbol{x}$,$\boldsymbol{y}$ 样本间 L_2 距离(L_2 divergences);

(4)根据样本间差异度量矩阵 $\boldsymbol{d}_{i,j}$ 按重心法自底向上进行层次聚类分析,以两个类的重心之间距离为两类间的距离,得到初步的聚类结果;

(5)如果聚类结果最后合并的两类有一类只有一个参与者,则该参与者为隐写者;如果聚类结果最后合并的两类都包含多个参与者,则距离另一类最远的参与者为隐写者.

4 实 验

4.1 实验配置

为验证所提方法的有效性,本文采用 BOSSBase v1.0 图库及自制图库共来自 6 台相机(表 2)的图像,各相机选取前 100 幅用以模拟 6 个参与者,假设每个参与者各使用一种相机.将这些图像压缩至 JPEG 质量因子 90,并用开源隐写算法 nsF5[6] 和 JPHide[7] 进行隐写,其中嵌入率密度 a 与嵌入率 r(单位为 bpac(bit per non-zero AC coefficient))相等,$a,r\in[0.1,0.7]$,记为 $a\times r$.

表 2 每个参与者对应的图像信息

参与者	相机型号	分辨率	图像来源
A1	Canon EOS 400D Digital	2592×3888	BOSSBase
A2	Canon EOS Digital Rebel XSi	4272×2848	BOSSBase
A3	M9 Digital Camera	5212×3468	BOSSBase
A4	Nikon D70	2000×3008	BOSSBase
A5	Pentax K20D	4672×3104	BOSSBase
A6	Canon EOS 450D	4272×2848	自制图库

实验用自然图像来模拟非隐写者,并从嵌入率为 r 的隐写图像中随机选择其中的 $\lfloor\alpha\cdot n\rfloor$ 幅代替对应自然图像,用以模拟隐写者.依次遍历每个参与者的不同嵌入率配置,记录其成为隐写者时,用本文方法识别正确的概率.计算概率时,对每个隐写者随机选取隐写图像重复 100 次.

4.2 实验结果

基于 KFDI 的隐写聚类分析并与 MMD 的对照实验结果如表 3 所示,其中准确率较高的情况用粗体显示.从表中可以看出,在两种隐写算法中的低嵌入率部分,KFDI 检测率高于 MMD 最高 30%,而在高嵌入率部分,检测率降低不超过 5%.

造成该现象的主要原因是在低嵌入率情况下,隐写样本与载体样本的特征之间差异性较小,而样本中心的聚合程度差异较大.此时考虑类内方差的因素有利于提高检测能力.而在高嵌入率时,MMD 聚类准确率已较高,隐写样本与载体样本的特征之间差异性较大.此时类内方差对样本间差异度量的影响较小,考虑类内方差的因素无助于提高检测率.

表 3 不同隐写算法及嵌入配置下的识别率(100 次实验)

隐写算法	分析方法	识别率						
		0.1×10%	0.2×20%	0.3×30%	0.4×40%	0.5×50%	0.6×60%	0.7×70%
		(嵌入负载(bpac)×嵌入密度)						
nsF5	文献[11]方法	0.2283	0.3283	0.6233	**0.9767**	**0.9917**	**0.9933**	**0.9950**
	本文方法	**0.2650**	**0.5900**	**0.9367**	0.9567	0.9683	0.9717	0.9733
JPHide	文献[11]方法	0.6167	0.6367	0.6500	0.8117	**0.9833**	**0.9900**	**0.9933**
	本文方法	**0.6733**	**0.7517**	**0.8683**	**0.9350**	0.9500	0.9567	0.9650

表 4 展示了当嵌入负载 0.30 bpac 及嵌入密度为 30%情况下的检测结果.从表中可知,采用文献[11]方法时,A2、A3 和 A4 多被识别为 A3,相对地本文方法采用 KFDI 作为样本间差异度量,结果就会有较好的区分度.

对于其中一次实验,聚类层次图如图 3 所示.从图中可以看出,在一定的嵌入率下,隐写者在聚类中会被划分为最后一层,从而暴露其隐写行为.需要指出的是,在嵌入率较低的情况下,聚类时更倾向于判定为某个特定的相机(表 4 的 A6).对于 MMD 指标,尽管隐写者是 A4,最后判定为 A3(图 3(b)),和没有隐写者时的情况大致相同(图 3(a)).而在 KFDI 下,其准确地判断为 A4.

表 4　本文方法与文献[11]方法隐写者识别结果(100 次实验)

(a) nsF5，文献[11]方法，0.30 bpac×30%

隐写者	判定结果					
	A1	A2	A3	A4	A5	A6
A1	87	0	4	0	0	9
A2	0	100	0	0	0	0
A3	0	0	100	0	0	0
A4	0	0	100	0	0	0
A5	0	0	12	0	88	0
A6	0	0	1	0	0	99

(b) nsF5，本文方法，0.30 bpac×30%

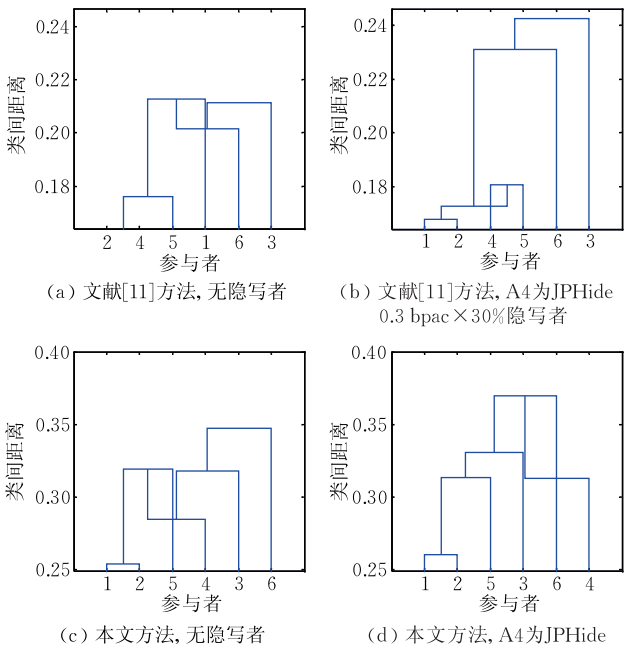
隐写者	判定结果					
	A1	A2	A3	A4	A5	A6
A1	89	0	0	0	0	11
A2	0	90	1	0	0	9
A3	0	0	95	0	0	5
A4	0	0	0	96	0	4
A5	0	0	0	0	92	8
A6	0	0	0	0	0	100

(c) JPHide，文献[11]方法，0.30 bpac×30%

隐写者	判定结果					
	A1	A2	A3	A4	A5	A6
A1	0	0	3	0	0	97
A2	0	96	1	0	0	3
A3	0	0	99	0	0	1
A4	0	0	100	0	0	0
A5	0	0	0	0	95	5
A6	0	0	0	0	0	100

(d) JPHide，本文方法，0.30 bpac×30%

隐写者	判定结果					
	A1	A2	A3	A4	A5	A6
A1	59	0	0	0	0	41
A2	0	89	0	0	0	11
A3	0	0	91	0	0	9
A4	0	0	1	94	0	5
A5	0	1	6	1	88	4
A6	0	0	0	0	0	100

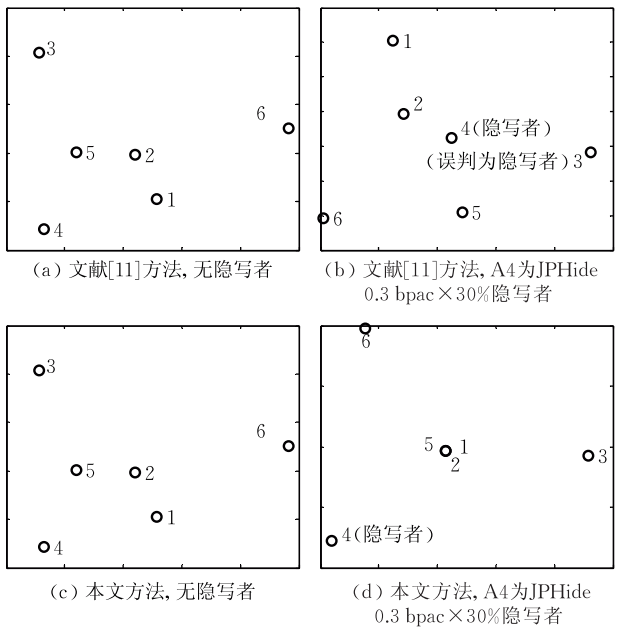


(a) 文献[11]方法，无隐写者　(b) 文献[11]方法，A4为JPHide 0.3 bpac×30%隐写者

(c) 本文方法，无隐写者　(d) 本文方法，A4为JPHide 0.3 bpac×30%隐写者

图 3　本文方法与文献[11]方法在有无隐写者时的层次图比较

本文将各个参与者之间的 KFDI 作为距离，用多维标度法(MultiDimensional Scaling，MDS)[17]作图，得到图 3 实验对应各个参与者的大致平面分布(图 4)．从图中可以看出，在一定的嵌入率下，隐写者总是和普通参与者距离较大．此时，在距离足够大的时候，隐写行为可以被检测．



(a) 文献[11]方法，无隐写者　(b) 文献[11]方法，A4为JPHide 0.3 bpac×30%隐写者

(c) 本文方法，无隐写者　(d) 本文方法，A4为JPHide 0.3 bpac×30%隐写者

图 4　本文方法与文献[11]方法在有无隐写者时的 MDS 图比较

5　结束语

载体来源和隐写算法范围未知往往造成隐写分析准确率的下降，高级用户可能据此采用不公开的载体源或算法来逃避隐写分析．在该情况下，Ker 经过实验提出了一种基于 MMD 进行隐写聚类分析的有效方法．然而，MMD 只考虑类与类的中心点之间的距离却忽略了样本聚集程度因素对计算样本间差异度量的影响．为了解决该问题并进一步提高隐写分析准确率，本文提出了一种用核 Fisher 鉴别指标提高隐写聚类分析的方法．该方法不仅考虑了类间方差对聚类效果的主要作用，还考虑类内方差在计算样本间差异度量时的补充作用，较大地提高了相对低嵌入率情况下的检测能力，最高提高约 30%，同时，在高嵌入率情况检测率降低不超过 5%．

参　考　文　献

[1] Cachin C. An information-theoretic model for steganography//Aucsmith D ed. Proceedings of the 1st International Workshop on Information Hiding. LNCS 2939. Springer,

1998: 35-49

[2] Fridrich J, Pevný T. Multi-class blind steganalysis for JPEG images//Proceedings of the SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII. San Jose, 2006: 0O01-0O13

[3] Westfeld A. F5-a steganographic algorithm: high capacity despite better steganalysis//Moskowitz ed. Proceedings of the 4th International Workshop on Information Hiding (IH'2001). LNCS 2137. Berlin. Springer-Verlag, 2001: 289-302

[4] Kim Y, Duric Z, Richards D. Modified matrix encoding technique for minimal distortion steganography//Camenisch eds. Proceedings of the 8th International Workshop on Information Hiding (IH'2006). LNCS 4437. Springer, 2007: 314-327

[5] Fridrich J, Goljan M, Soukal D. Perturbed quantization steganography. ACM Multimedia Systems, 2005, 11(2): 98-107

[6] Fridrich J, Pevný T, Kodovský J. Statistically undetectable JPEG steganography: dead-ends, challenges, and opportunities//Proceedings of the 9th ACM Workshop on Multimedia and Security. Dallas, TX, 2007: 3-14

[7] JP Hide & Seek. http://linux01. gwdg. de/~ alatham/stego. html, 2011

[8] Chang C C, Lin C. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST), 2011, 2(3): 27

[9] Theodoridis S. Pattern Recognition. 4th Edition. Beijing: China Machine Press, 2009

[10] Fridrich J, Kodovský J. Breaking HUGO—The process discovery//Proceedings of the 13th International Workshop on Information Hiding. Prague, 2011: 85-101

[11] Ker A D, Pevný T. A new paradigm for steganalysis via clustering//Proceedings of the Society of Photo-optical Instrumentation Engineers. San Francisco, CA, 2011: 0U01-0U13

[12] Pevný T, Fridrich J. Benchmarking for steganography//Proceedings of the 6th Information Hiding (5284). Santa Barbara, CA, 2008: 251-267

[13] Mika S, Ratsch G, Weston J, Scholkopf B, Muller K. Fisher discriminant analysis with kernels//Hu Y, Larsen J, Wilson E, Douglas S eds. Neural Networks for Signal Processing IX. NJ: IEEE Press, 1999: 41-48

[14] Pevný T, Fridrich J. Merging Markov and DCT features for multi-class JPEG steganalysis//Delp E J, Wong P W eds. Proceedings of the SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX. San Jose, CA, 2007: 3-1-3-14

[15] Muller K-R, Mika S, Ratsch G, Tsuda K, Scholkopf B. An introduction to kernel-based learning algorithms. IEEE Transactions on Neural Networks, 2001, 12(2): 181-201

[16] Fisher R A. The statistical utilization of multiple measurements. Annals of Eugenics, 1938, 8: 376-386

[17] Borg I, Groenen P. Modern Multidimensional Scaling: Theory and Applications. 2nd Edition. New York: Springer-Verlag, 2005: 207-212

[18] Gretton A, Borgwardt K, Rasch M, Scholkopf B, Smola A. A kernel method for the two-sample-problem//Jordan M I, Cun Y L, Solla S A. Advances in Neural Information Processing Systems, Cambridge: MIT Press, 2007, 19: 513-520

HUANG Wei, born in 1985, Ph. D. candidate. His main research interests include information hiding and steganalysis.

ZHAO Xian-Feng, born in 1969, Ph. D., associate professor. His main research interests focus on information hiding.

SHENG Ren-Nong, born in 1969, M. S., associate professor. His main research interests focus on signal processing.

Background

Steganography is an art and science of hiding information such that its presence cannot be detected, while steganalysis refers to the analysis of intercepted signals to determine whether they contain hidden messages. Since researchers have made much effort on steganalysis with many conditions known, blind steganalysis which have no idea of what kind of steganographic scheme or embedding rates are used. Ker proposed an approach on hierarchical clustering with centroid linkage via maximum mean discrepancy (MMD) to decide the most suspicious actor. But the accuracies need improvement. In this paper, we proposed clustering via kernel Fisher discriminant index (KFDI) to achieve higher accuracies, especially in low embedding rates. The authors of this paper engaged in the field of information hiding in the National Natural Science Foundation of China (NSFC) and other projects funded by government agencies. We explored the accuracy of detecting steganography in more realistic scenarios, in NSFC No. 61170281, etc. We have obtained some patents and published papers. The contributions of this paper are on three folds: (1) The impact of within-class variances in measuring the distance between clusters is emphasized; (2) A more accurate steganalytic clustering is proposed; (3) It's more practically useful when there is little knowledge about the steganographers.