运用统计流形降维的通用型隐写分析算法①

戴良斌, 全笑梅

(北京工业大学 计算机学院, 北京 100124)

摘 要: 统计流形是参数化概率密度函数组成的流形, 将统计流形引入到隐写分析中. 根据图像特征向量的概率密度函数, 将 Fisher 信息距离作为差异度量标准, 然后通过降维将特征嵌入低维欧式空间并采用支持向量机作为分类器. 实验结果表明该算法对 JSteg, F5, MBS1, MBS2, nsF5 等隐写算法都有较好的识别效果.

关键词: 隐写分析; 统计流形; 降维; 特征向量; Fisher 信息距离

Universal Steganalysis Using Statistical Manifold Dimension Reduction

DAI Liang-Bin, QUAN Xiao-Mei

(College of Computer Science, Beijing University of Technology, Beijing 100124, China)

Abstract: Statistical manifold is a manifold of parameterized probability density function. In order to improve steganalysis rate, we propose a using statistical manifold dimension reduction in this paper. The procedure of this algorithm is as follow: first, we use fisher information metric measures the difference among probability density function of image feature vector. Then, we embed characteristics to lower Eucli—dean space by dimension reduction. Finally, we use the support vector machine (SVM) as classifier. The experimental results show that this algorithm is effective to JSteg, F5, MBS1, MBS2 and nsF5 stegano-graphy algorithm.

Key words: steganalysis; statistical manifold; dimension reduction; feature vector; fisher information metric

1 引言

隐写分析是检测作为载体的数字多媒体文件中是否含有隐藏信息的技术, 它与信息隐藏技术相对应. 隐写分析一般分为两种: 一种是针对某一特定隐写算法的检测分析方法, 这类检测方法针对性较强, 但通常随着隐写算法的改进而失效; 另一种是可以检测多种隐写算法的通用型隐写分析, 其一般思路是选定特征向量后, 运用机器学习的方法对检测对象分类. 最简单的通用型隐写分析算法通常只能判断出待检测媒体中是否有隐藏信息. 该方法通常包含两个重要步骤: 从图像中提取特征和利用分类器进行分类判决. 由于通用型隐写分析算法适用范围较广, 而且对新出现的隐写算法也同样有效, 因此通用型隐写分析是该领域目前的研究热点. 通用型隐写分析中分类器的选择对分类效果起重要作用, 现有的分类器包括线性分类器、非线性分类器以及几何分类器等. 支持向量机(Suppport Vector Machine,SVM)[1]是一种较为常用的分类器, 具有较高分类能力和较强鲁棒性. 对于线性不可分的情况, SVM 通常把数据投影到更高维的空间. SVM 的使用促进了特征的高维度化, 使得隐写分析使用更多更复杂的特征成为可能. 并且随着隐写算法的推动, 隐写分析的特征逐步趋于高维多样化.

鉴于高维数据计算困难, 易引发 “维数灾难”问题, 针对高维数据需要进行降维处理. 流形学习是一种非线性降维方法, 其目标是发现数据集中的流形结构, 并在降维的同时尽量保持这些结构信息. 但有些情况高维数据并不存在对应的低维欧式流形, 这时更合理的假设是数据依赖统计流形, 即流形中的每个点均是概率密度函数. 统计流形解决了欧式空间中不易描述的问题, 适合不能用低维欧式流形描述特征的高维数据. 现统计流形降维已在文本分类、人脸识别、纹理识别等研究领域得到应用, Lee et al.[2]将统计流形

① 收稿时间:2014-01-14;收到修改稿时间:2014-02-20

应用到图像分割, K.M. Carter 的 Fisher 信息非参数嵌入[3], 利用非参数方法估算 Fisher 信息距离, 将统计流形降维应用在流式细胞识别和文本分类上.

本文将统计流形引入隐写分析, 具体框架如下: 1)提取图像 DCT 域特征向量; 2)利用 Fisher 信息距离度量特征分布相似度, 估算 Fisher 信息距离并计算流形上测地线距离; 3)针对测地线距离矩阵进行降维; 4)利用分类器对降维后的数据进行分类识别, 判断隐写分类效果.

2 相关知识简介

2.1 算法

基于统计流形降维技术利用信息几何知识, 在降维的同时消除数据间冗余, 便于后期分类处理. 本文使用的算法中, 提取的特征是 DCT 域统计特征, 利用该特征本身是概率密度函数的特点, 将统计流形降维框架应用到隐写分析.

本文使用的算法主要分为 5 步, 根据输入图像集(训练图像 train、测试图像 predict), 首先提取图像集 DCT 域特征向量, 文中使用 Fridrich 提出的图像校准[4]过程提取特征: 通过将 JPEG 图像解压到空域, 减去左上角四行四列, 对裁剪后的图像利用原图像的量化矩阵进行压缩, 得到校准后的图像, 然后提取原图像和校准图像的 DCT 域全局直方图作为特征向量; 第 2 步, 估算特征向量集的 Fisher 信息距离矩阵 D; 第 3 步: 根据 Fisher 信息矩阵 D 计算统计流形上测地线距离 G; 第 4 布: 针对矩阵 G 降维; 第 5 步: 利用低维数据训练 SVM, 训练好的分类器将对测试图像集进行是否隐写分类判断.

2.2 Fisher 信息距离估计

统计流形是一种特殊的流形, 该流形中每一个元素都是参数化的概率密度函数. 根据信息几何理论, 概率密度函数间的距离用 Fisher 信息距离度量. Fisher 信息距离的数学定义是对数似然函数对总体分布参数导数的方差, $F=\{f_\theta(x)\mid\theta\in\Theta\}$ 设为一个由参数分布簇组成的函数, 其中 $\theta=(\theta_1,\theta_2,\cdots\theta_n)\in\Theta\subset R^n$ 为分布参数, 定义

$$I(\theta)_{i,j}=E[(\frac{\partial}{\partial\theta_i}lnf_\theta(x))(\frac{\partial}{\partial\theta_j}lnf_\theta(x))] \quad (1)$$

则矩阵 $I(\theta)$ 为 Fisher 信息矩阵. 通过定义可知, 计算 Fisher 信息距离需要已知统计流形的参数, 但在实际应用中这些参数通常是未知的, 因此根据定义直接求解 Fisher 信息距离比较困难. 但无论统计流形上的参数如何变化, Fisher 信息距离始终是流形上的一致度量. 这说明当无法获得流形上特定参数时, 可以通过恰当的信息度量来近似估算 Fisher 信息距离[3].

现已经有多种度量函数可以估算 Fisher 信息距离, 如: KL 散度、Hellinger 距离和 cosine 距离等. KL 散度是信息论中重要的度量, 假设 p(x)、q(x)为两个随机变量的概率密度函数, 则其 KL 散度计算式为

$$KL(p//q)=\int p(x)\log\frac{p(x)}{q(x)}dx \quad (2)$$

KL 散度又称 KL 距离, 常用来衡量一个概率密度相对于另一个概率密度的相对熵. KL 散度与 Fisher 信息度量有如下关系: 当 $p\to q$ 时, $\sqrt{2KL(p||q)}\to D_F(p||q)$ 需要说明的是, KL 散度不满足对称性, 也不满足三角不等式, 为满足对称性可将 KL 散度定义为

$$D_{KL}(p,q)=[KL(p,q)+KL(q,p)]/2 \quad (3)$$

因此最终可得 $D_F(p||q)\approx\sqrt{D_{KL}(p,q)}$, KL 散度适合估算值变化较大的连续概率分布函数间距离, 且随着两个概率分布差异的增大, 其 KL 散度也随之增大; 但由于不稳定性, KL 散度不适合从有限样本点计算的概率分布函数, 这时可用 Hellinger 距离估算. Hellinger 距离也常用来衡量两个概率分布的相似度, 其数学计算公式为

$$D_H(p//q)=\sqrt{\int(\sqrt{p(x)}-\sqrt{q(x)})^2dx} \quad (4)$$

Hellinger 距离满足对称性和三角不等式, 且与 Fisher 信息度量有如下关系: 当 $p\to q$ 时 $2D_H(p||q)\to D_F(p||q)$, Hellinger 距离常用来估算多项式概率分布点间距离; cosine 距离常用来度量球面分布的点间距离, 其数学计算公式为

$$D_c(p//q)=2\arccos\int\sqrt{p(x)\times q(x)}dx \quad (5)$$

且 cosine 距离与 Fisher 信息度量的关系为 $p\to q$ 时 $D_c(p||q)\to D_F(p||q)$. 当然, 除 KL 散度、Hellinger 距离、cosine 距离外, 也可以用其他信息距离估算 Fisher 信息距离. 鉴于提取的图像 DCT 域全局直方图具有多项式分布特性, 本文选择 Hellinger 距离估算

Fisher 信息距离.

2.3 统计流形上测地线距离估计

统计流形的数学定义如下，设 S 是一簇由参数 $\theta=(\theta_1,\theta_{2.}...\theta_n)$ 确定的概率密度函数所组成的集合，即 $S=\{p(x|\theta)|\theta\in\Theta\subseteq R^n\}$，称 S 为统计流形，其中 θ 不仅是 S 的参数，而且还是 S 的坐标系. 统计流形中每一个参数 θ 对应一个概率密度函数 $p(x|\theta)$，同时一个概率密度函数对应于统计流形中的一个点.

通常，欧式空间中任意两点间的距离是连接两点的直线长度，该直线距离通常用 L_2 范式度量.

流形上两点间的距离用测地线距离度量，测地线距离指沿着流形连接两点最短的路径距离，该最短路径即测地线路径. 通常测地线路径是一条曲线路径，但随着两点的接近，测地线距离逐渐收敛于欧氏距离，因此可以用欧式距离近似流形中两近邻点之间的测地线距离. 但当两端点距离较远时，由于空间扭曲，测地线距离不能直接用两点间的欧式距离近似度量，但可以用沿着流形连接两点间最短的欧式距离之和近似测地线距离. 如图 1，p 和 q 为近邻点，两点间距离可用欧式距离近似(直线)，p 和 r 不是近邻点，则两点间距离为测地线距离(直线段总和). 计算测地线距离矩阵通常借助于邻域图的最短路径，该过程如下：首先构建一个邻域图，若两点互为邻域点，连接两点，并将其权值设为两点间的欧式距离，否则将其权值设为∞. 其次计算邻域图中顶点间的最短路径，得到最短路径矩阵，该最短路径矩阵即测地线矩阵.

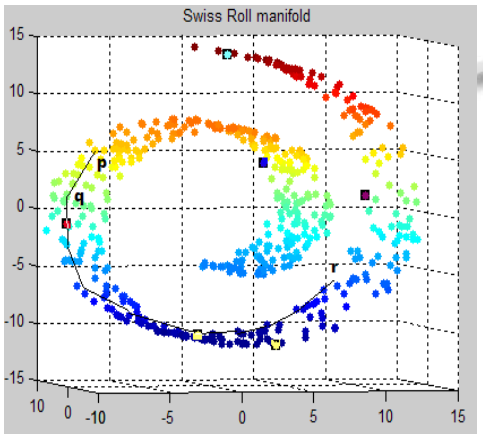

图 1 欧式距离与测地线距离

统计流形上测地线距离的计算和流形上测地线距离的计算类似，图 2 所示为统计流形中两点间的测地线距离. 与流形中测地线距离计算不同的是，需要用 Fisher 信息距离估算统计流形中两近邻点之间的测地线距离. 当统计流形中两点距离较远时，其测地线距

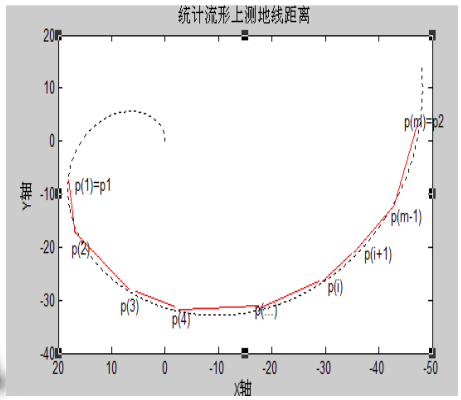

图 2 统计流形上测地线距离

离为连接两点的最小 Fisher 信息距离路径. 即给定由 n 个参数 $p_\theta=\{\theta_1,...,\theta_n\}$ 参数化的概率密度函数，p1、p2 之间的测地线距离计算公式为

$$G_F(\mathrm{p}_1,\mathrm{p}_2;P)\approx\min_{m,P}\sum_{i=1}^{m}\hat{D}_F(\mathrm{p(i)},\mathrm{p(i+1)})$$
$$\mathrm{p(i)}\to\mathrm{p(i+1)}\forall i \tag{6}$$

其中 $\hat{D}_F$ 是 Fisher 信息距离的估计值，估算式为，

$$D_F(\mathrm{p}_1,\mathrm{p}_2)\approx\min_{m,\{\theta(1)...\theta(m)\}}\sum_{i=1}^{m}D_F(\mathrm{p}(\theta_{(i)}),\mathrm{p}(\theta_{(i+1)}))$$
$$,\mathrm{p}(\theta_{(i)})\to\mathrm{p}(\theta_{(i+1)}) \tag{7}$$

且满足 $p(\theta_{(1)})=p_1,p(\theta_{(m)})=p_2,\{\theta_{(1)}...\theta_{(m)}\}\in p_\theta,m\le n$. 统计流形中测地线矩阵的计算也需要借助于最短路径矩阵.

2.4 降维

文中使用的降维方法是拉普拉斯特征映射(Laplacian Eigenmaps, LEM)[5]，其基本思想是高维空间中距离很近的点投影到低维空间距离也应该很近. 具体步骤如下：首先构建邻接图 G，若数据点 x_i 在数据点 x_j 邻域内，则 $G_{i,j}$=1，否则 $G_{i,j}$=0；其次计算权值矩阵 W，若 x_i、x_j 邻接，则，$W_{ij}=e^{-\frac{D_x(i,j)^2}{t}}$ 否则 $W_{ij}=0$. 最后通过权值矩阵 W 计算拉普拉斯矩阵 L，L 等于 W 与 W 的对角矩阵 Φ 之差. 则低维嵌入坐标为 $[V_1,...V_d]^T$，其中 $V_1,...V_d$ 对应特征方程 $Ly=\lambda\Phi y$

的 d 个最小特征值所对应特征向量.

2.5 SVM 分类器

SVM 是一种基于统计学习理论的分类方法, 其最大特点是结构风险最小化原则. SVM 是为解决二分类问题产生的, 具有较强的泛化能力.

针对二分类问题, SVM 通过寻找一个超平面作为两类训练数据的分割, 以保证最小的分类错误率. 当线性可分时, SVM 寻找使两类样本的分类间隔最大的最优超平面, 设 n 个样本 $x_i \in R^n$, 所属类别, $y_i \in \{1,-1\}$ 最优超平面满足方程 $wx+b=0$, 且使两类样本分类间隔最大. 最优超平面的问题可归结为满足约束条件 $y_i(x_i gw+b)-1 \geq 0$ 的 $\min\limits_{w,b}\frac{1}{2}||w||^2$.

当线性不可分时, SVM 通过核函数把样本数据投影到高维特征空间, 在高维空间中构造超平面.

现有的常用核函数有: 线性核函数、$k(x,y)=x^T y+c$ 多项式核函数 $k(x,y)=(\alpha xy+c)^d$、高斯核函数 $k(x,y)=exp(-\frac{||x-y||^2}{2\sigma^2})$.

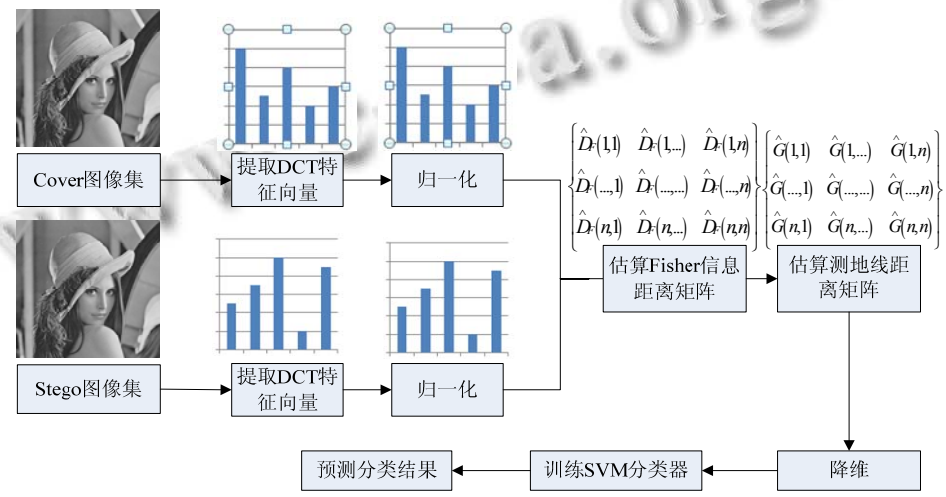

图 3　操作流程图

表 1

隐写方法		嵌入率 10%	25%	50%	75%	100%
JSteg	虚警率	4.2%	0%	0%	0%	0%
	漏检率	1.2%	0%	0%	0%	0%
	正确率	94.6%	100%	100%	100%	100%
F5	虚警率	20.4%	14.2%	8.4	1.2%	0%
	漏检率	18.8%	8%	4.6%	1%	0%
	正确率	60.8%	79.8%	87%	97.8%	100%
MBS1	虚警率	13.4%	6%	4.6%	1.4%	1.8%
	漏检率	30.6%	21.2%	5.8%	3.6%	3.8%
	正确率	56%	72.8%	89.6%	95%	94.4%
MBS2	虚警率	11.6%	8.8%	4.4%	2.8%	1.4%
	漏检率	27.4%	13.8%	6.4%	2.2%	3.4%
	正确率	61%	77.4%	89.2%	95%	95.2%
nsF5	虚警率	10.6%	5.8%	5%	1.6%	0%
	漏检率	35.6%	7.6%	3.8%	3.4%	0%
	正确率	53.8%	86.6%	91.2%	95%	100%

3　实验

3.1 实验步骤

实验使用 BossBase 作为数据库, 该数据库中包含 10000 幅 512*512 图像, 每一幅都进行 F5[6]、JSteg、MBS1[7]、MBS2[7]、nsF5[8]隐写, 嵌入率分别为 10%、25%、50%、75%和 100%. 随机选取 2000 张图片作为训练集(1000 张 cover 图像, 1000 张 stego 图像), 选取 500 张图像用作预测集(250 张 cover 图像, 250 张 stego 图像). 使用校准图像 DCT 域全局直方图[4]作为特征向量, 降维方法使用 LEM. 针对降维结果分类, SVM 分类器[9]选择高斯核函数.

3.2 实验结果及分析

从实验结果可以看出, 对于每一种隐写算法, 随着隐写嵌入率的增大, 正确率也随之增大. 在 10%的嵌入率情况下, JSteg 算法的识别率可达到 94.6%; 当嵌入率提高到 25%时, 对 F5、nsF5、MBS1 和 MBS2

算法正确检测率达到 80%左右. 基于统计流形降维可得到数据集的紧致向量, 增强数据间差异度, 提高分类检测率. 实验结果表明该方法针对多种隐写算法都有较高的分类检测率, 且针对不同嵌入率算法, 本方法都有较强的稳定性和正确性.

4 结论

本文把统计流形降维引入到隐写分析, 针对图像高维特征向量, 进行了基于统计流形的降维操作. 实验证明本算法已能将载体图像和隐秘图像分开, 在保持低虚警率和漏检率的同时, 提高了正确识别率, 因此针对高维特征向量进行统计流形降维, 在隐写分析领域具有可行性和实用性.

参考文献

1 Cortes C, Vapnik V. Support vector network. Machining Learning, 1995, 20(3): 273–297.

2 Lee SM, Abbott AL, Araman PA. Dimensionality reduction and clustering on statistical manifolds. Proc. IEEE Conf. on Computer Vision and Pattern Recognition. June 2007. 1–7.

3 Carter KM, Raich R, Finn WG, Hero AO. Fine: Fisher information non-parametric embedding. IEEE Trans. on Pattern Analyis and Machine Intelligence, 2008.

4 Fridrich J. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. 6th International Workshop. IH 2004 Toronto, Canada. May 2004.

5 Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Dietterich TG, Becker S, Ghahramani Z, eds. Advances in Neural Information Processing Systems. MIT Press. 2002.

6 Fridrich J, Goljan M, Hogea D. Steganalysis of JPEG images: Breaking the F5 algorithm. Proc. of 6th International Workshop on Information Hiding. Lecture Notes in Computer Science. Springer-Verlag. 2004, 2578: 310–323.

7 Sallee P. Model-based steganography. International Workshop on Digital Watermarking. Seoul, Korea. Springer-Verlag Press, 2003. 154–167.

8 Fridrich J, Pevn T, Kodovsk J. Statistically undetectable JPEG steganography: Dead ends, challenges, and opportunities. In: Dittmann J, Fridrich J, eds. Proc. of the 9th ACM Multimedia & Security Workshop. Dallas, TX. September 20–21, 2007. 3–14.

9 Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans. on Intelligent Systems and Technology, 2011, 27: 1–27.