

一种新的隐写分析方法:IKLDA

陈国明¹, 印 鉴¹, 周端宁², 张 东¹

(1. 中山大学信息科学与技术学院, 广东广州 510275; 2. 美国东华盛顿大学商学院, 华盛顿州斯波坎市 99202)

摘 要: 如何降低高维数据的维数而不损失原有数据的内在信息是机器学习和数据挖掘领域中的热点问题. 本文在图嵌入框架的基础上提出一种新的降维分析算法 IKLDA(improved kernel Linear discriminant analysis), 不仅使得隐藏在图像的信息能被区分出来, 而且大大降低了数据的维数, 理论分析及实验结果表明 IKLDA 的降维隐写分析是有效的, 比其它传统降维方法效果要好, 并且进一步推进了数据挖掘可视化方法在隐写分析的应用.

关键词: 隐写分析; 降维; 图嵌入框架; 可视化

中图分类号: TP399 **文献标识码:** A **文章编号:** 0372-2112 (2009) 08-1762-06

A New Method for Steganalysis Techniques: IKLDA

CHEN Guo-ming¹, YIN Jian¹, ZHOU Duan-ning², ZHANG Dong¹

(1. School of Information Science and Technology, Sun Yat-sen University, Guangzhou, Guangdong 510275, China;

2. College of Business, Eastern Washington University, Spokane WA 99202, USA)

Abstract: Reducing the dimensionality of data without losing intrinsic information is a hotspot in machine learning and data mining. In this paper we propose a new dimensionality reduction algorithms call IKLDA(improved kernel Linear discriminant analysis) on the ground of graph embedding framework. Our method not only can detect the information hidden in digital images but also reduce the dimensionality. Theoretical analysis and experiments show that our new IKLDA algorithm is effective in steganalysis and is more precise than the other traditional dimensional reduction methods. Furthermore, our method promotes development of visualization in the application in steganalysis.

Key words: steganalysis; dimensionality reduction; graph embedding framework; visualization

1 引言

隐写分析是从隐藏信息的载体(如图像)中发现是否有隐藏信息,是信息隐藏技术的反过程.盲隐写分析方法是通过对提取图像特征建立高阶统计模型,利用分类器进行分类判决.这种隐写分析方法不局限于单一的隐写算法,通常对若干隐写算法都适用.将隐写分析当作模式识别两类(即将图像分成嵌入隐藏信息前、后两类)的分类问题,是目前研究的热点.经典的盲隐写分析方法有:J.Farid等^[1]提出的基于小波统计特性的隐写分析方法;通过提取各个小波子带以及多个高阶统计量,组成72维(灰度图像)的特征空间. Shi, Y. Q等^[2]则提出基于 Haar小波统计特性提取小波子带组成特征空间等.常用盲检测的一般思路是选定特征统计量后,对隐藏信息的检测就归结为对检测对象的分类问题,检测方法中的分类是通过确定一个判决门限来区分两类检测

对象,判决门限是经过训练方式确定的.一种较为常用的分类方法是支持向量机^[9](Support Vector Machine, SVM).对于线性不可分的情况,SVM通常把数据投影到一个更高维特征空间.针对于一般情况下高维特征可分性较好但低维特征可分性较差的问题,本文提出一种新方法,使高维融合的特征降到低维后仍然保持一个良好的分类效果,而对降维后的分类结果通过可视化技术直观地展现出来.对于降维方面则通过修正类间与类内距离的权值系数,在图嵌入框架^[3]的基础上提出了一种新的基于核技术的IKLDA算法.

降维方法近年来成为机器学习和数据挖掘中的热点问题.降维方法分为线性方法和非线性方法两大类.其中高效的线性算法有PCA^[11],LDA^[11]等,而非线性降维算法则有ISOMAP^[6],LLE^[8],Laplacian Eigenmap^[7]等多种算法.近年来核(kernel trick)

收稿日期:2008-10-23;修回日期:2009-05-04

基金项目:国家自然科学基金(No.60573097, No.60773198, No.60703133);广东自然科学基金(No.06104916, No.8151027501000021);高等学校博士学科点专项科研基金(No.200500558017);新世纪优秀人才支持计划(No.NCET-06-0727)

技术作为把线性降维方法扩展为非线性降维方法也迅速发展起来. Linear discriminant analysis (LDA) 是一种降维方法, 其目的是使类间的距离最大化的同时类内的距离最小. Laplacian eigenmap 的基本思想: 在高维空间中邻近的点投影到低维空间中也应该离得很近. 通过使用两点间的加权距离作为损失函数, 可求得相应的降维结果. 虽然传统的 LDA 算法在很多实际应用中取得较好的效果, 但它仍然存在局限性, 即不能保持数据的内在属性. 通过在图嵌入框架中发展新的降维新算法可以改进传统 LDA 算法的局限性. 降维的一个目标就是找出嵌入在高维数据中的低维数据空间使得其仍然保留原有数据的内在属性. 如果降维的方法恰当, 那么得到的低维数据在模式分类及数据可视化方面将具有重要价值.

本文的贡献是: (1) 在图嵌入框架的基础上提出了一种新的基于核技术的有监督学习算法 IKLDA 来解决隐写分析这个分类问题, 成功的把原始图像和隐密图像区分开来, 理论分析和实验结果表明, 新方法比传统的降维算法有更高的分类准确度; (2) 提出一种可行有效的先融合特征后再降维的隐写分析方法; (3) 估算数据集特征空间的内在维数, 推动数据挖掘可视化技术在隐写分析中的应用.

2 IKLDA 算法

2.1 LDA 算法及其改进

LDA 寻找一个线性映射使得在低维数据空间中最大化类间间距的同时最小化类内的距离. 假设有 l 个样本 $x_1, x_2, \dots, x_l \in R^d$, 属于 C 个类别. 当讨论有监督学习问题时, 设与样本点对应的类别标签是 $Z_i \in (1, 2, \dots, C)$. X 是所有样本点的矩阵. 设 $Y \in R^r (1 \leq r \leq d)$ 是高维样本点 X 的低维表示, r 是低维空间的维数. 先考虑线性降维的方法. 定义一个 $d \times r$ 的转换矩阵 a , 嵌入在高维样本中的低维空间可以表示成为: $Y = a^T X$.

LDA 的目标就是: 寻找一个转换矩阵 a , 使得数据点类间距离最大的同时类内距离最小.

$$a_{lda} = \operatorname{argmax} \frac{a^T S_b a}{a^T S_w a} \quad (1)$$

$$S_b = \sum_{k=1}^c l_k (\mu^k - \mu) (\mu^k - \mu)^T \quad (2)$$

$$S_w = \sum_{k=1}^c \sum_{i=1}^{l_k} (x_i^k - \mu^k) (x_i^k - \mu^k)^T \quad (3)$$

$$S_t = \sum_{i=1}^l (x_i - \mu) (x_i - \mu)^T \quad (4)$$

$$S_t = S_b + S_w$$

其中 μ 是所有样本的平均值, l_k 是第 k 类样本的数目, μ^k 是第 k 类样本的平均值, x_i^k 是 k 类中的第 i 个

样本. S_w 称为类内散射矩阵, 是类内距离的度量; S_b 称为类间散射矩阵, 是类间距离的度量. 所以, LDA 就是要找一个转换矩阵 a_{lda} , 使类间散射矩阵最大的同时类内散射矩阵最小.

要求解式(1)中的 a_{lda} 就转化成为求解一个特征方程的问题. 设 $\{\Phi_k\}_{k=1}^d$ 成为 $S_b \Phi = \lambda S_w \Phi$ 特征值问题的特征向量, 对应着特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$. 设 $a_{lda} = \{\Phi_1, \Phi_2, \dots, \Phi_d\}$ 为转换矩阵. 嵌入在高维数据中的低维子空间能够表示成为: $Y = a_{lda}^T X$.

为了把 LDA 算法融入图嵌入框架中, 类内散射矩阵 S_w 可以转换成数据点对间关系的表达式. 由于 μ^k 是第 k 类均值, 令

$$\mu^k = \frac{1}{l_k} \sum_{j=1}^{l_k} x_j^k, (\mu^k)^T = \frac{1}{l_k} \sum_{i=1}^{l_k} x_i^k, \text{代入式(3):}$$

$$\begin{aligned} S_w &= \sum_{k=1}^c \sum_{i=1}^{l_k} (x_i^k - \mu^k) (x_i^k - \mu^k)^T \\ &= \sum_{k=1}^c \sum_{i=1}^{l_k} x_i^k (x_i^k)^T - \sum_{k=1}^c \frac{1}{l_k} \sum_{i=1}^{l_k} x_i^k (x_j^k)^T \\ &= \frac{1}{l_k} \sum_{k=1}^c \sum_{i,j=1}^{l_k} x_i^k (x_i^k)^T - \frac{1}{l_k} \sum_{k=1}^c \sum_{i,j=1}^{l_k} x_i^k (x_j^k)^T \\ &= \frac{1}{2 \cdot l_k} \sum_{i,j=1}^{l_k} (x_i (x_i)^T + x_j (x_j)^T - x_i (x_j)^T - x_j (x_i)^T) \\ &= \frac{1}{2 \cdot l_k} \sum_{i,j=1}^{l_k} (x_i - x_j) (x_i - x_j)^T \end{aligned} \quad (5)$$

同理类间散射矩阵 S_b 可以转换成数据点对间关系的

$$\begin{aligned} \text{表达式. 由于 } \mu \text{ 是所有样本的均值, 令 } \mu &= \frac{1}{l} \sum_{j=1}^l x_j \mu^T = \\ &= \frac{1}{l} \sum_{i=1}^l x_i, \text{代入式(2),} \end{aligned}$$

$$\begin{aligned} S_b &= S_t - S_w = \sum_{i=1}^l (x_i - \mu) (x_i - \mu)^T - S_w \\ &= \sum_{i=1}^l x_i x_i^T - \frac{1}{l} \sum_{j=1}^l x_j x_j^T - S_w = \frac{1}{l} \left(\sum_{i=1}^l x_i x_i^T - \sum_{j=1}^l x_j x_j^T \right) - S_w \\ &= \frac{1}{2 \cdot l} \sum_{i,j=1}^l (x_i - x_j) (x_i - x_j)^T - \frac{1}{2 \cdot l} \sum_{i,j=1}^l (x_i - x_j) (x_i - x_j)^T \end{aligned} \quad (6)$$

在式(5)和式(6)的推导过程中, 其实加入了一个隐含条件, 就是 x_i 和 x_j 属于同一类别. 对于一个分类问题, 近邻点可能具有相同的类别标签. 在图嵌入框架中利用 Laplacian 的在高维空间中邻近的点投影到低维空间中也应该离得很近的这一基本思想. 从而把 LDA 算法精髓中的类间距离大; 类内距离小这一基本思想融合到图嵌入框架中来求解问题. 从图嵌入角度来看, 给定一个样本的集合 $x_i (i = 1, 2, \dots, l)$, 用 K 近邻带权图 $G\{X, W\}$ 来表示近邻点之间的关系. X 是顶点的集合. 如果两个顶点是近邻关系, 那么用一条边来连接这两个顶点 i, j . $W_{i,j}$ 的每个元素表示顶点点对间相似度的

取值,其中非零值表示这条边的权值:

$$W_{i,j} = \begin{cases} 1, & \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, & \text{else} \end{cases} \quad (7)$$

其中 $N_k(x_i)$ 表示 x_i 近邻点的集合.对式(7)进行如下改进,邻接相似度矩阵 $W_{i,j}$ 可以表示成为基于高斯分布的欧氏距离:

$$W_{i,j} = \begin{cases} \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma_i \sigma_j}\right), & \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, & \text{else} \end{cases} \quad (8)$$

其中 $\sigma_i = \|x_i - x_i(K)\|$, $x_i(K)$ 是 x_i 的第 K 个近邻.

图 G 的对角阵 D 和拉普拉斯矩阵 L 定义如下:

$$L = D - W, D_{ii} = \sum_{j=1}^n W_{i,j} \quad (9)$$

在图嵌入框架中的算法就是要找合适的低维向量来表达图 G 顶点的相似度关系,低维空间的向量表示成为 $Y = [y_1, y_2, \dots, y_n]^T$. 其中 y_i 是顶点 x_i 的低维表示. 原数据空间近邻点在嵌入空间中有相同的属性,低维空间中的点受到原始特征空间的相似度关系限制. 求解这个问题变成寻找一个低维空间表示的降维问题:

$$Y = \underset{Y^T B Y = d}{\operatorname{argmin}} \sum_{i \neq j} \|y_i - y_j\|^2 W_{i,j} = \underset{Y^T B Y = d}{\operatorname{argmin}} Y^T L Y \quad (10)$$

其中 d 是一个常量, B 是一个限制矩阵,引入 B 的目的是避免目标函数出现的多个取值问题. 这种基于图的保持相似度的属性就是使得对于近邻点 x_i 和 x_j 来说, y_i 和 y_j 的距离更短. 从而在图嵌入框架中,保持原始数据特征空间向低维空间映射的属性. 由高维空间向低维空间的线性映射 $Y = X^T W$ 及式(10)可以得到待求解的线性映射函数 w :

$$W = \underset{w^T B w = d}{\operatorname{argmin}} \sum_{i \neq j} \|W^T x_i - W^T x_j\|^2 W_{i,j} = \underset{w^T B w = d}{\operatorname{argmin}} w^T X L X^T w \quad (11)$$

2.2 基于核的 LDA 技术与图嵌入框架结合的 IKL-DA 算法

以上描述的算法为线性算法,它在实际应用中分类问题可能会对某些高维非线性数据失效,从而不能发现流形中的内在几何结构. 由此,我们采用再生核希尔伯特空间 Reproducing Kernel Hilbert Space (RKHS) 技术来扩展传统的 LDA 算法. 先考虑一个非线性映射 $\Phi(x)$ 从 R^n 到再生核希尔伯特空间的映射. $\langle \Phi(x), \Phi(y) \rangle = K(x, y)$ 对应某一变换空间中的内积. K 是核函数,不同的核函数代表不同的学习机器. 常用的核函数为:

$$\begin{cases} \text{高斯核函数: } K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right); \\ \text{多项式核函数: } K(x, y) = (c + \langle x, y \rangle)^d; \\ \text{Sigmoid 核函数: } K(x, y) = \tanh(\langle x, y + \alpha \rangle); \end{cases} \quad (12)$$

本文的核函数采用高斯核函数 $K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$, $\sigma > 0$. 核技术需要计算数据点对间的内积 $K_{i,j} = K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$. 由核函数与式(11)可得映射函数 α :

$$\alpha = \underset{w^T K B k = d}{\operatorname{argmin}} \sum_{i \neq j} \|\alpha^T k_i - \alpha^T k_j\|^2 W_{i,j} = \underset{w^T K B k = d}{\operatorname{argmin}} w^T K L K^T w \quad (13)$$

其中 k_i 表示核矩阵中第 k 列向量,解(13)就是解决广义特征值分解问题:

$$K B K \alpha = \lambda K L K \alpha \quad (14)$$

其中 $L = D - W$, $D_{ii} = \sum_{j=1}^n W_{i,j}$, $W_{i,j}$ 可由式(8)得到

由式(5)得到类内散射矩阵 S_w 点对关系的权重系数

$$S_w^{\text{weight}} = \frac{1}{l_k} \quad (15)$$

由式(6)得到类间散射矩阵 S_b 点对关系的权重系数

$$S_b^{\text{weight}} = \left(\frac{1}{l} - \frac{1}{l_k}\right) \quad (16)$$

在解式(14)的过程中,由于 $K L K$ 的退化问题^[5] 给解广义特征值分解问题带来困难,所以通常加入调节系数 ϵ :

$$K B K \alpha = \lambda (K L K + \epsilon I) \alpha \quad (17)$$

其中 ϵ 是一个很小的常量.

IKLDA 算法的核心思想在于把基于核的 LDA 技术与图嵌入框架中 Laplacian 相结合. 从而使得新算法具备:(1)数据点类间距离最大的同时类内距离最小;(2)高维空间的近邻点投影到低维空间也是近邻点. 在融合的过程中,式(8)中的邻接图相似度矩阵 $W_{i,j}$ 意味着同一个类内相距较远的点对将对 LDA 算法中的类内散射矩阵 S_w 与类间散射矩阵 S_b 影响很小. 由于不同类别间点对关系的叠加强对 S_w 、 S_b 产生影响,所以我们加入权重因子进行调节.

基于 LDA 算法改进的基础在于对 S_w 、 S_b 权值系数的定义,例如 loog^[14] 提出 approximate pairwise accuracy criterion (aPAC) 度量准则; Sugiyama^[4] 提出 LFDA 算法,其中定义 $S'_w = \frac{1}{l_k} S_w$, $S'_b = \frac{1}{l} S_b + (1 - \frac{l_k}{l}) e_i e_i^T + \frac{1}{l} e_i e_i^T$; Yan, S. Hu^[15] 提出的 ICBKM 聚类算法中定义: $S'_w =$

$$l - \sum_{k=1}^c \frac{1}{l_k} e_k e_k^T, S'_b = \sum_{k=1}^c \frac{1}{l_k} e_k e_k^T - \frac{1}{l} e e^T$$

由式(15), (16), (17), 我们重新定义了类间与类内距离的权值系数:

$$\begin{cases} S'_w = \alpha \sqrt{\frac{1}{l_k}} K L K, \text{ 其中 } 0 < \alpha < 1, D_{ii} = \sum_{j \neq i} W_{ij}, L = D - W \\ S'_b = (1 - \alpha) \left(\sqrt{\frac{1}{l}} - \sqrt{\frac{1}{l_k}} \right) K L K, \text{ 同上} \end{cases} \quad (18)$$

式(17)可以写成如下形式:

$$S'_{\mu}\alpha = \lambda(S'_{\omega} + \epsilon I)\alpha \quad (19)$$

综上所述:由式(8),(9),(18),(19)及高斯核函数可得 IKLDA 算法:

$$\left((1-\alpha)\left(\sqrt{\frac{1}{l}} - \sqrt{\frac{1}{l_k}}\right)K L K\right)\beta = \lambda\left(\alpha\sqrt{\frac{1}{l_k}}K L K + \epsilon I\right)\beta \quad (20)$$

其中 $K = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right)$, $D_{ii} = \sum_{i \neq j} W_{i,j}$, $L = D - W$, $W_{i,j}$ 可由式(8)得到.

设 $\{\beta_i\}_{i=1}^n$ 是式(20)的广义特征向量而 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ 是式(20)对应的特征值.我们可得到降维后的特征空间 $Y = \beta^T K$.

3 实验

实验使用 1813 幅 256×256 的灰度 jpg 原始图像,每一幅都进行 steghide^[12]隐写,信息嵌入率为 90%.在实验的数据集上我们先采用了特征融合的方法.数据集中共有 150 维特征,前 72 维是由 Farid^[1]的隐写分析方法而得,后 78 维由 Shi^[2]的隐写分析方法而得.72 维的意义如下:对原始图像一次小波分解后水平、垂直、对角方向子带分别记为 $H1$ 、 $V1$ 、 $D1$;第二层分解后的三个子带记为: $H2$ 、 $V2$ 、 $D2$;第三层分解后的三个子带记为: $H3$ 、 $V3$ 、 $D3$.对三层小波系数的预测误差子带记为: $EH1$ 、 $EV1$ 、 $ED1$, $EH2$ 、 $EV2$ 、 $ED2$, $EH3$ 、 $EV3$ 、 $ED3$.前四阶统计量记为: mean, var, kur, ske. 所以有, 1: meanV, 2: meanH, 3: meanD, 4: varV, 5: varH, 6: varD, 7: kurV, 8: kurH, 9: kurD, 10: skeV, 11: skeH, 12: skeD, 13: meanEV, 14: meanEH, 15: meanED, 16: varEV, 17: varEH, 18: varED, 19: kurEV, 20: kurEH, 21: kurED, 22: skeEV, 23: skeEH, 24: skeED, 然后 25 到 48 为第二层小波分解各量的统计量, 49 到 72 维为第三层小波分解各量的统计量, 排列顺序同第一层.78 维意义: 1 到 3: 为原始图像的前三阶统计量; 4 到 6 为预测误差图像的前三阶统计量; 7 到 42 维为原始图像进行前三层 haar 小波分解各子带的前三阶统计量, 排列顺序为: meanLL1, varLL1, kurLL1, meanHL1, varHL1, kurHL1, meanLH1, varLH1, kurLH1, meanHH1, varHH1, kurHH1, meanLL2, varLL2, kurLL2, meanHL2, varHL2, kurHL2, meanLH2, varLH2, kurLH2, meanHH2, varHH2, kurHH2, meanLL3, varLL3, kurLL3, meanHL3, varHL3, kurHL3, meanLH3, varLH3, kurLH3, meanHH3, varHH3, kurHH3; 43 到 78 维为预测误差图像进行前三层 haar 小波分解后各子带的前三阶统计量, 排列顺序同 7 到 42 维.前 1 到 1813 行数据为原始图像特征向量, 后 1814 到 3626 行数据为隐写图像特征向量, 并且前面分别用 +1 和 -1 标示出来.

3.1 数据内在维数的估计问题

对于一个降维问题来讲,在实际应用中选取将特征向量降到多少维是一个很难回答的问题.选取特征过少,不足以描述图像多方面的特性,将导致误检率或漏检率过高;选取特征太多,特征之间的冗余性和依赖性较大,从而区分性不好.一般来说内在维数是表达出数据信息所需的最少参数的数目.我们使用相关维估计法^[13]来估计数据集的内在维数:

$$C(r) = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n c, \text{ 其中 } c = \begin{cases} 1, & \text{if } \|x_i - x_j\| \leq r \\ 0, & \text{if } \|x_i - x_j\| > r \end{cases} \quad (21)$$

我们用 $C(r)$ 来估计数据的内在维数 d :

$$d = \lim_{r \rightarrow 0} \frac{\log c(r)}{\log r} \quad (22)$$

上述方程一般不容易求解,实际中一般使用如下公式:

$$\bar{d} = \frac{\log(c(r_2) - c(r_1))}{\log(r_2 - r_1)} \quad (23)$$

通过式(23)用上文的数据集计算得到内在估计的维数取值为 2.91741.

3.2 在隐写分析中实施可视化技术

我们通过实施核技术前的 LDA 算法可得到相对应的特征值,这些特征值可以提供关于数据集内在维数的有用信息.特征值图中前 d 个特征值取值一般较高, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, 它们变化的趋势较大, d 可以看成数据集的内在维数.剩余特征值一般变化趋势较小,它们可代表维数约简模型中的冗余噪声.图 1 横坐标表示维数,纵坐标表示特征值.从图 1 我们观察到:当 LDA 约简到 30 维特征的时候,只有前面 3 维所对应的特征值变化趋势较大,余下维数所对应的特征值变化趋缓.结合相关维估计法所计算出来的维数 2.91741,我们把数据集的内在维数确定为 3 维.由此可见,在隐写分析中是可以实施可视化技术的.然而 IKLDA 算法也有局限性:用 IKLDA 算法经过核技术映射后取得的特征值却未能观察到这种大的变化趋势,而且也与传统 LDA 算法中维数不超过数据集中的类别数稍有区别.

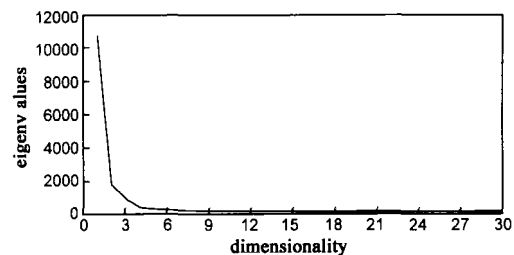


图1 LDA 维数(30维)与特征值关系图

3.3 IKLDA 与其它降维算法比较的特征云图

下面的实验是通过降维后得到的 3 维原始图像及

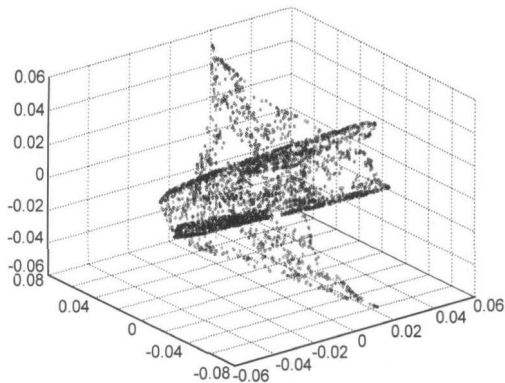


图2 Laplacian eigenmap降维特征云图

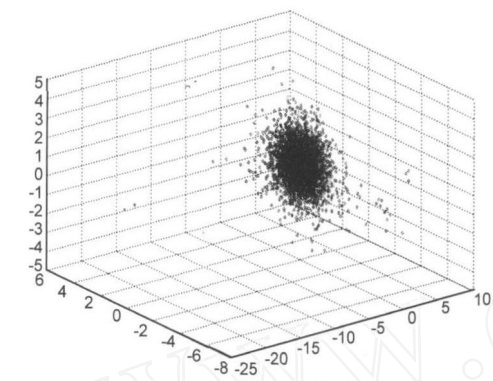


图3 LDA降维特征云图

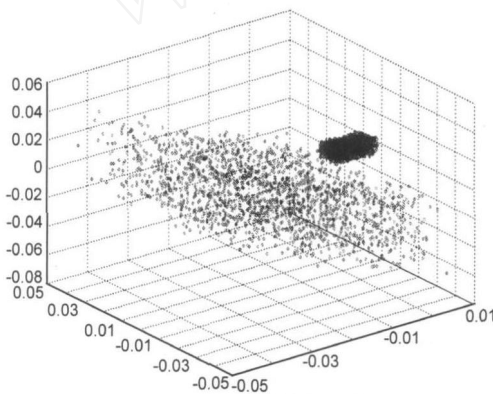


图4 IKLDA降维特征云图

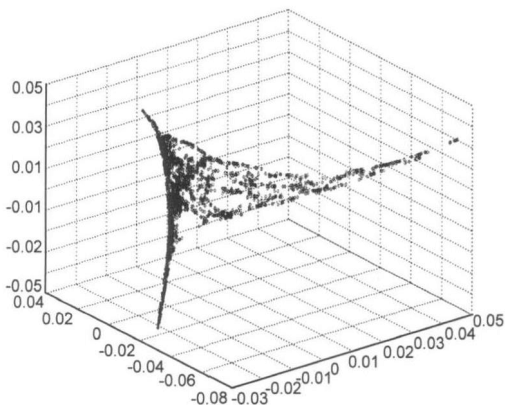


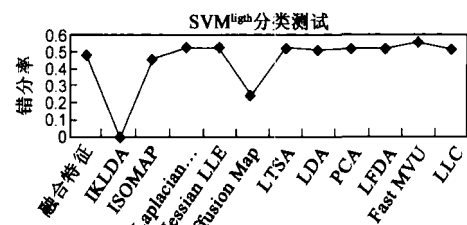
图5 LFDA降维特征云图

其隐密图像样本点集构成的特征云图,实验中把高维特征空间降到3维以实现可视化分析比较.从图4的实验结果来看,本文所用的IKLDA算法已经完全能够把

两类数据区分开来,类间的距离相对较大,而其中的一个类别数据的类内距离较大,而另一个类别数据的类内距离较小.虽然图2中Laplacian eigenmap从几何结构上能看出两类数据是有分别的,呈现出对称性,但特征空间中存在多数样本点重叠不可分,类间距离较小,两个类别各自的类内距离较大.但考虑其几何结构的对称性,如果同其他方法结合,将有改善分类效果的希望.实验结果证明我们的IKLDA算法比其他降维算法Isomap、Laplacian eigenmap、Hessian LLE、Diffusion Map、LTSA、LDA、PCA、FastMVU、LFDA、MDS及LLC分类效果都要好.由于篇幅的关系,图2~图5只列出其中的4幅特征云图进行比较.通过本文采用的算法,不但把高维特征向量降到了3维,以可视化的方式展现出结果,而且具备较高的隐写区分效果.

3.4 用支持向量机测试融合的特征及若干降维算法降维后的特征分类效果

我们用支持向量机SVM^{Light}作为分类器测试特征融合及若干降维算法降维后的分类效果,比较它们的错分率.SVM^{Light}使用RBF核函数,参数Gamma设置为0.3,参数error penalization设置为0.5.在3626个样本集合中,使用3000个样本来训练分类器,余下626个样本作为测试数据.我们得到的错分率测试结果如图6.从SVM^{Light}的分类效果看我们的算法IKLDA的错分率最小,分类效果最好;其次是Diffusion Map,错分率也较低;而原始融合特征和ISOMAP比其他降维算法Laplacian eigenmap、Hessian LLE、LTSA、LDA、LFDA、MDS、PCA、FastMVU、LLC分类效果稍好,但差别不大.



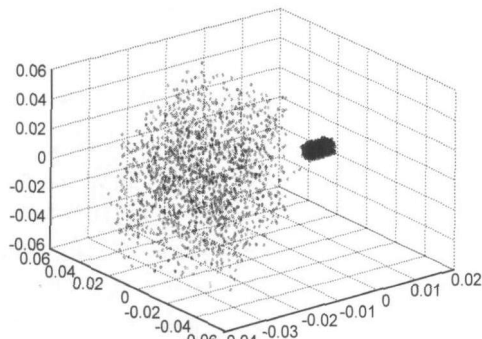
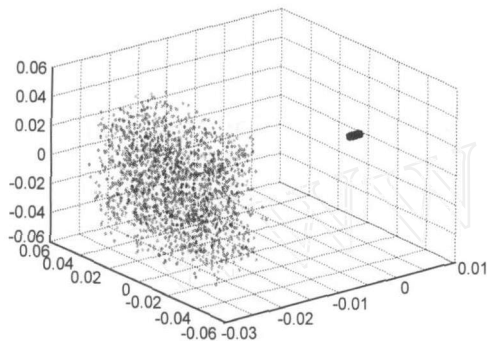
融合的特征空间和各降维算法得到的特征空间

图6 SVM^{Light}分类错分率

3.5 IKLDA 算法中关键参数的设定

为了展示IKLDA算法如何工作,以下实验是IKLDA算法中关于参数的设置及其分类效果的实验.在式(20)中 ϵ 是一个很小的常量,先固定 a 给 ϵ 设置不同的取值,可以得到不同的分类效果特征云图:随着 ϵ 的取值减少,正负两类间的类间距离趋大,类内距离趋小,分类效果趋好.但随着 ϵ 的取值进一步减少,出现逆转:类内距离增大,类间距离减少.之后随着 ϵ 的取值减少,再次出现类内距离减少,类间距离增大;然后再固定 ϵ ,给 a 设置不同取值,观察特征云图的变化:随着 a

的增大,其中一个类别数据的类内距离不断变小,而另一个类别数据的类内距离不断变大;而两类数据的类间距离有变大的趋势.由于篇幅的关系,图 7、图 8 也列出其中两幅特征云图.

图7 $\varepsilon=0.0001$ 图8 $\alpha=0.6$

4 结论

本文在图嵌入框架的基础上提出一种新的降维算法 IKLDA,在盲隐写分析的实验中取得很好效果.本文区别于其他盲检测方法在于:通用盲检测的一般思路是选取检测对象的多个特征统计量,由统计量构成特征矢量集,再利用神经网络、聚类算法或 SVM 从试验数据中构造检测模型;而本文先通过 Farid 和 Shi 的特征向量融合、再确定数据集的内在维数,然后对高维特征向量采用 IKLDA 降维算法降至 3 维,最后采用可视化的技术展现隐写分析的分类结果.实验证明了本文的方法已能将干净图像与隐密图像清楚地区分开,而且比传统降维算法分类效果要好.因此将多个特征量融合后再进行降维的方法,具备较高的可分性.

参考文献:

- [1] Farid H, et al. Steganalysis using color wavelet statistics and one-class support vector machines[A]. In Proc SPIE-Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VI[C]. San Jose: SPIE, 2004. 35 - 45.
- [2] Shi Y Q, et al. Effective steganalysis based on statistical moments of wavelet characteristic function [A]. In Proc IEEE ITCC2005 [C]. Las Vegas: IEEE Computer Society Press,

2005. 768 - 773.

- [3] Yan S, et al. Graph embedding and extensions: A general framework for dimensionality reduction [J]. IEEE Transactions, 2007, PAMI-29(1): 40 - 51.
- [4] Sugiyama M. Local fisher discriminant analysis for supervised dimensionality reduction [A]. In Proc ICML2006 [C]. Pittsburgh: ACM Press, 2006. 905 - 912.
- [5] Friedman H. Regularized discriminant analysis [J]. Journal of the American Statistical Association, 1989, 84 (405): 165 - 175.
- [6] Tenenbaum J B, et al. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290 (5500): 2319 - 2323.
- [7] Belkin M, et al. Laplacian eigenmaps and spectral techniques for embedding and clustering [A]. In Proc NIPS2002 [C]. Vancouver: MIT Press, 2002. 585 - 591.
- [8] Belkin M, et al. Laplacian eigenmaps for dimensionality reduction and data representation [R]. Technical Report, University of Chicago, 2002.
- [9] Cortes C, et al. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273 - 297.
- [10] He X, et al. Face recognition using Laplacian faces [J]. IEEE Transactions, 2005, PAMI - 27(3): 328 - 340.
- [11] Martinez A, et al. PCA versus LDA [J]. IEEE Transactions, 2001, PAMI-23(2): 228 - 233.
- [12] Savoldi A, et al. Blind multi-class steganalysis system using wavelet statistics [A]. IHH-MSP2007 [C]. Los Alamitos: IEEE Computer Society Press, 2007. 93 - 96.
- [13] Kawaguchi A, et al. Estimating the correlation dimension from a chaotic system with dynamic noise [J]. Japan Statist, 2005, 5(2): 287 - 302.
- [14] Loog M, et al. Multiclass linear dimension reduction by weighted pairwise fisher criteria [J]. IEEE Transactions, 2001, PAMI - 23(7): 762 - 766.
- [15] Yan S, et al. Nonlinear discriminant analysis on embedded manifold [J]. IEEE Transactions, 2007, CSVT-17(4): 468 - 477.

作者简介:



陈国明 男, 1977 年生于广东广州, 中山大学信息科学与技术学院讲师, 博士研究生. 主要从事机器学习、数据挖掘及模式识别的研究.
E-mail: isscgm@mail.sysu.edu.cn

印鉴 男, 1968 年生于湖北仙桃, 中山大学信息科学与技术学院教授、博士生导师. 主要从事机器学习、数据挖掘及模式识别的研究.