

基于核主成分分析的数据流降维研究

高宏宾, 侯 杰, 李瑞光

GAO Hongbin, HOU Jie, LI Ruiguang

五邑大学 计算机学院, 广东 江门 529020

School of Computer Science and Technology, Wuyi University, Jiangmen, Guangdong 529020, China

GAO Hongbin, HOU Jie, LI Ruiguang. Research on dimension reduction of data stream based on kernel principal component analysis. Computer Engineering and Applications, 2013, 49(11): 105-109.

Abstract: Theory and implementation of two data stream dimension reduction algorithms, PCA and KPCA, are analyzed. Due to linear PCA and KPCA can not effectively reduce data stream dimension when applied over large scale stream data, a new dimension reduction technique called GKPCA is proposed. With GKPCA, data sets are first partitioned into groups, and then KPCA is applied over each group. Data sets are filtered and regrouped into a new dataset. KPCA is again evaluated over the new data sets. This process is preceding recursively when some reduction threshold is reached which simplifies data stream sampling space and reduces time and space complexity of KPCA. Experimental analysis over different datasets illustrates that GKPCA can reduce data stream dimension excellently with less time consumption.

Key words: Kernel Principal Component Analysis (KPCA); data stream; dimension reduction

摘 要: 分析了数据流降维算法 PCA 和 KPCA 的原理和实现方法。针对在大型数据集上 PCA 线性降维无法有效实现降维且 KPCA 的降维效率差, 提出了一种新的降维策略 GKPCA 算法。该算法将数据集先分组, 对每一组执行 KPCA, 然后过滤重新组合数据集, 再次应用 KPCA 算法, 达到简化样本空间, 降低了时间复杂度和空间复杂度。实验分析表明, GKPCA 算法不仅能取得良好的降维效果, 而且时间消耗少。

关键词: 核主成分分析; 数据流; 降维

文献标志码: A **中图分类号:** TP39 **doi:** 10.3778/j.issn.1002-8331.1110-0295

随着科学技术的发展, 人们不断接触到各种类型的数据, 如图像数据、文本数据、医学信息数据、金融市场交易数据、天文数据等, 大多数情况下这些数据具有多维空间属性, 在统计处理中称之为“高维数据”, 且具有数据流 (Data Stream) 性质。数据的高维度和流的形式给数据处理带来巨大的挑战, 例如, 当处理 256×256 的图片序列时, 通常将图像数据拉成一个向量, 这样, 得到一系列 4 096 维的数据序列, 如果直接对这些数据进行处理, 巨大的计算量将使人们无法忍受, 即“数据维灾”, 同时这些数据通常没有反映出数据的本质特征, 如果直接对它们进行处理, 不会得到理想的结果。所以需要对数据进行降维, 将数据的维数降到合适大小, 同时尽可能多地保留原数据信息, 然后对降维后的数据进行处理^[1]。

降维是将高维模式映射到低维子空间的过程, 其主要

目的有:

- (1) 压缩数据以减少存储量;
- (2) 去除噪声的影响;
- (3) 从数据中提取特征以便于进行分类或聚类;
- (4) 将数据投影到低维空间, 便于理解数据分布。

本文讨论了高维数据降维的定义及常用算法, 接着对核主成分分析 (KPCA) 进行研究, 并提出一种计算效率更高的分组核主成分分析 (GKPCA) 算法, 实验数据表明, 该算法能有效地对数据流进行特征提取, 实现高维数据降维。

1 降维的定义及算法

1.1 降维的定义

若 $\{x_n\}^N \subset R^D$ 为 D 维空间中的一个容量为 N 的数据

基金项目: 广东省自然科学基金 (No.S2011010003681)。

作者简介: 高宏宾 (1960—), 男, 在读博士, 副教授, 硕士生导师, 主要研究领域为数据挖掘与分布式计算、算法设计与分析; 侯杰 (1986—), 男, 硕士研究生, CCF 会员, 研究方向为数据挖掘与分布式计算、算法设计与分析; 李瑞光 (1985—), 男, 硕士研究生, 研究方向为数据挖掘与分布式计算、算法设计与分析。E-mail: tn10000@126.com

收稿日期: 2011-10-17 **修回日期:** 2011-12-06 **文章编号:** 1002-8331(2013)11-0105-05

CNKI 出版日期: 2012-03-08 <http://www.cnki.net/kcms/detail/11.2127.TP.20120308.1521.046.html>

集合,假设其来自于维数为 $d(d \ll D)$ 的某一数据流形采样。处理目标是研究数据流形合适的低维坐标描述,将原始数据投影到低维空间,获得原数据集合的低维简洁表示。降维问题通常由下列部分组成^[2]:

(1)高维数据空间, Ω^D , 通常为 R^D 的某一个子集。

(2)降维空间(或称低维表示空间), Ω^d 。

(3)映射函数 φ

$$\varphi: \Omega^D \rightarrow \Omega^d; x \rightarrow y = \varphi(x)$$

称 y 为 x 的降维表示。

1.2 降维算法

目前,数据降维方法主要分为两类^[3]:线性降维和非线性降维。

线性降维包括:主成分分析(Principle Component Analysis, PCA)、独立成分分析(Indispensable Component Analysis, ICA)、线性判别分析(Linear Discriminate Analysis, LDA)、投影寻踪(Projection Pursuit, PP)等;PCA是一种基于二阶统计的数据分析方法,该方法在各个变量之间相关关系研究的基础上,用一组较少、互不相关的新变量代替原来较多的变量,而且使这些新变量尽可能多地保留原来复杂变量所反映的信息,具体计算步骤可参看文献[4]。

非线性降维算法包括:核方法、多维尺度方法(Multidimensional Scaling, MDS)、等距离映射算法(ISOMAP)、局部线性嵌入(Locally Linear Embedding, LLE)^[5]等。

核主成分分析作为PCA方法的在处理非线性问题时的扩展,近年来得到了快速的发展,其主要思想是通过某种事先选择的非线性映射函数将输入矢量映射到一个高维线性特征空间,然后在特征空间中使用PCA方法计算主成分^[6]。在具备了PCA所有数学特性的前提下,KPCA还有如下特征:KPCA比PCA提供更优的识别性;KPCA的计算复杂度不因变换空间的维数增加而加大,仅与输入空间的维数相关。目前KPCA的思想已经被世界上广大的科研人员所接受,广泛应用到特征提取、人脸识别^[7]、图像识别^[8]、光谱降维分析^[9-10]、数据分类聚类以及故障监控诊断等领域。

由于KPCA的计算复杂度是和样本数密切相关的,对于一个 $N \times N$ 的核矩阵(N 为样本数),其时间复杂度为 $O(N^3)$,当 N 很大时,KPCA的计算代价会快速增长,因此需要对KPCA算法进行优化,文献[11]介绍了一种将求解PCA的期望最大算法(EM)推广到求解KPCA,但该算法占用存储空间较大,且收敛性不能保证。下面将着重介绍核主成分分析方法与在此基础上的改进算法。

2 核方法与基于KPCA降维算法的改进

2.1 核方法

核方法一种由线性到非线性之间的桥梁,起源于20世纪初叶,是一系列先进性数据处理技术的总称,共同点是这些方法中都用了核映射。核函数方法的基本原理是通过非线性函数把输入空间映射到高维空间,在特征空间中进行数据处理,其关键在于通过引入核函数,把非线性变换后的特征空间内积运算转换为原始空间的核函数计算,从而简化了计算量。其操作过程如图1所示。

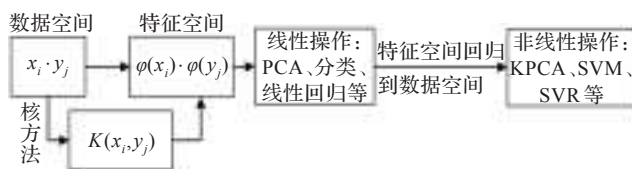


图1 核方法框架示意图

实质上,核方法实现了数据空间、特征空间、类别空间之间的非线性变换。图1中, x_i 和 x_j 为数据空间中的样本点,数据空间到特征空间的映射函数为 φ ,核函数的基础是实现向量的内积变换:

$$(x_i, x_j) \rightarrow K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle$$

在满足 Mercer 条件^[12]下,设输入空间的数据为 $x_i \in R^{d_L} (i=1, 2, \dots, N)$,对于任意对称连续的函数 $K(x_i, x_j)$,存在一个 Hilbert 空间,对映射 $\varphi: R^{d_L} \rightarrow H$,有:

$$K(x_i, x_j) = \sum_{n=1}^{d_F} \varphi_n(x_i) \cdot \varphi_n(x_j)$$

其中 d_F 为 H 的空间维数。这进一步说明,输入空间的核函数实际上与特征空间的内积是等价的,由于在核方法的各种实际应用中,只需要应用特征空间的内积,而不需要了解具体映射 φ 的具体形式,即:在使用核方法时只需要考虑如何选定一个合适的核函数,无需关心与之对应的映射 φ 是否具有复杂的表达式。具体核函数的性质及其构造方法可参看文献[13]。

2.2 KPCA 基本原理

对于输入空间的 M 个样本 $x_k (k=1, 2, \dots, M)$, $x_k \in R^N$,使 $\sum_{k=1}^M x_k = 0$,其协方差矩阵 C 为:

$$C = \frac{1}{M} \sum_{j=1}^M x_j x_j^T \quad (1)$$

对于经典的PCA算法,通过求解特征方程 $\lambda v = Cv$, λ 为 C 的一个特征值, v 为相应的特征向量,以此来获得贡献率较大的特征值和与之对应的特征向量。现引入非线性映射函数 φ ,使输入空间中样本点 x_1, x_2, \dots, x_M 变换为特征空间中的样本点 $\varphi(x_1), \varphi(x_2), \dots, \varphi(x_M)$,假设:

$$\sum_{k=1}^M \varphi(x_k) = 0 \quad (2)$$

则在特征空间 F 中,其协方差矩阵为:

$$\bar{C} = \frac{1}{M} \sum_{j=1}^M \varphi(x_j) \varphi(x_j)^T \quad (3)$$

因此,特征空间 F 的特征值和特征向量求解即求解方程:

$$\lambda v = \bar{C}v, v \in F \setminus \{0\} \quad (4)$$

从而有:

$$\lambda(\varphi(x_k) \cdot v) = \varphi(x_k) \cdot \bar{C}v \quad (5)$$

特征向量 v 可以线性表示为:

$$v = \sum_{i=1}^M \alpha_i \varphi(x_i) \quad (6)$$

由式(3)(5)(6)得:

$$\lambda \sum_{i=1}^M \alpha_i (\varphi(x_k) \cdot \varphi(x_i)) = \frac{1}{M} \sum_{i=1}^M \alpha_i (\varphi(x_k) \cdot \sum_{j=1}^M \varphi(x_j)) (\varphi(x_j) \cdot \varphi(x_i))$$
$$k=(1,2,\cdots,M)$$
(7)

定义 $M \times M$ 的矩阵 K :

$$K_{i,j} = \varphi(x_i) \cdot \varphi(x_j)$$
(8)

则式(7)可转化为:

$$M\lambda\alpha = K\alpha$$
(9)

求解式(9)即可获得映射空间上的特征值和特征向量。测试样本在 F 空间向量 V^k 上的投影为:

$$V^k \cdot \varphi(x) = \sum_{i=1}^M \alpha_i^k (\varphi(x_i) \cdot \varphi(x)) = \sum_{i=1}^M \alpha_i^k K(x_i, x)$$
(10)

如果特征空间中数据不满足中心化条件,则需要对和矩阵进行修正,用下式中的 \bar{K} 代替式(9)中的 K :

$$\bar{K}_{i,j} = K_{i,j} - \frac{1}{M} \sum_{m=1}^M l_{im} K_{mj} - \frac{1}{M} \sum_{n=1}^M l_{in} K_{nj} + \frac{1}{M^2} \sum_{m,n=1}^M l_{im} K_{mn} l_{nj}$$
(11)

式(11)中,对于所有的 $i,j, l_{i,j} = 1^{[14]}$ 。

2.3 KPCA 实现步骤

(1)从数据流中获取 m 条数据记录(每条记录有 n 个属性),将其表示成 $m \times n$ 维的矩阵:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

(2)选取适当的核函数,计算核矩阵 K ,本文中选取高斯径向基(RBF)核函数:

$$K(x,y) = \exp(-\frac{\|x-y\|^2}{\sigma^2})$$

(3)修正核矩阵 K ,得 KL 。

(4)计算 KL 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 和特征向量 v_1, v_2, \cdots, v_n 。

(5)将特征值按降序排列,并调整与其对应的特征向量。

(6)利用 Gram-Schmidt 正交法单位化特征向量,得到 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 。

(7)计算特征值的累积贡献率 B_1, B_2, \cdots, B_n ,根据给定的提取效率 p ,如果 $B_t \geq p$,则提取 t 个主分量 $\alpha_1, \alpha_2, \cdots, \alpha_t$ 。

(8)计算已修正的核矩阵 KL ,在提取出的特征向量上的投影 $Y = KL \cdot \alpha$,其中 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_t)$;所得的投影 Y 即为数据经 KPCA 降维后所得数据。

2.4 GKPCA 算法的提出

由于 KPCA 是基于样本的,算法的时间复杂度和空间复杂度与数据流维数无关,但是与样本记录数目却密切相关,对于小型高维数据流,KPCA 能够很好地完成降维,随着样本记录数目的增多,KPCA 的时间复杂度和空间复杂度也会增加。

为了降低计算时间与内存消耗,提出一种分组的 KPCA 算法,即将原始样本先分成若干组,对每一组数据进行过滤,然后再运用 KPCA 算法。

设原始样本记录数为 N ,现将其分为 G 组,则每组记录数目为:

$$G_1 = \lceil \frac{N}{G} \rceil$$

第 $i(i=1,2,\cdots,G)$ 组中与第一投影方向相对应的系数向量为 α^i :

$$\alpha^i = (\alpha_1^i, \alpha_2^i, \cdots, \alpha_G^i), i=1,2,\cdots,G$$

对每一组进行数据过滤,设过滤阈值为 ζ, α^i 按降序排列,设排序后为 β, β 满足: $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_G$,经过数据过滤后,设保留下来的记录数目为 $Left_i(i=1,2,\cdots,G)$,则 $Left_i$ 由下式决定:

$$\frac{\sum_{k=1}^{Left_i} \beta_k^2}{\|\alpha_i\|^2} \geq \zeta$$

ζ 越大,过滤掉的样本记录越少,否则越多。这样,每组数据过滤后形成新的样本集,并对新样本集进行 KPCA 算法,形成最终的投影方向。GKPCA 算法的实现步骤为:

- (1)重新采样并进行分组,形成 G 组样本。
- (2)对每一组执行 KPCA,并对样本进行过滤。
- (3)组合过滤后的样本,形成新的样本数据集,并再次执行 KPCA 算法,对样本进一步过滤。
- (4)获得 GKPCA 的特征投影。

3 实验与分析

实验环境:本文在 Matlab 下对 PCA、KPCA 和 GKPCA 进行模拟仿真,配置环境为 Pentium4, 2.8 GHz CPU,内存 1 GB, Windows XP 操作系统,采用的数据流为“SPECTF 心脏数据”(http://archive.ics.uci.edu/ml/machine-learning-databases/spect/),该数据集来源为美国俄亥俄州医学院,曾应用在“知识发现方法自动心脏显像诊断”(Artificial Intelligence in Medicine)。数据集描述心脏单个质子发射计算机断层显像(SPECT)影像诊断。每个病人分为两类:正常(1)和不正常(0)。测试数据集包含了 187 条实例,每条实例有 45 个属性,一组二进制数据和 44 组连续数据,连续数据取值为[0,100]。

测试结果:实验选取提取效率 $p=0.95$,过滤阈值 $\zeta=0.8$,分组数为 3,分析了 PCA、KPCA 和 GKPCA 算法的降维效果和计算时间,并进行了对比,图 2~图 5 分别显示了原始数据集分布、PCA 降维后数据分布、KPCA 降维后数据分布和 GKPCA 降维后的数据分布。

通过降维结果分布图的对比与分析,可以看出,三种算法都有效地实现了数据降维,且相对于 PCA 和 KPCA 算法,GKPCA 算法的降维效果更好。

表 1 对比了三种算法降维后的数据集维数和计算时间,从表 1 中可以看出,KPCA 算法有效地实现了数据降

表 1 数据集 Spect 降维结果

数据集或算法	实例数×维数	降维数目	计算时间/s
原始数据集	187×45	—	—
PCA	187×27	18	0.050
KPCA	187×24	21	1.290
过滤后样本	51×45	—	—
GKPCA	51×19	26	0.521

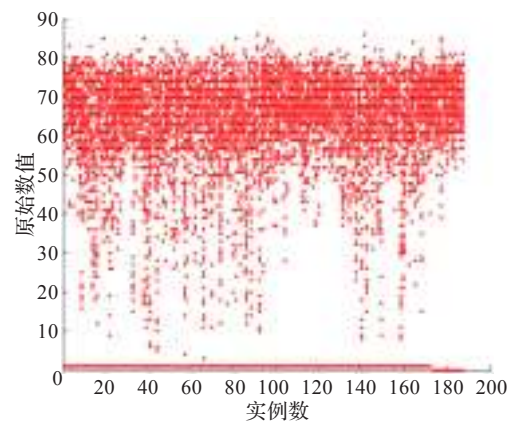


图2 原始数据分布图

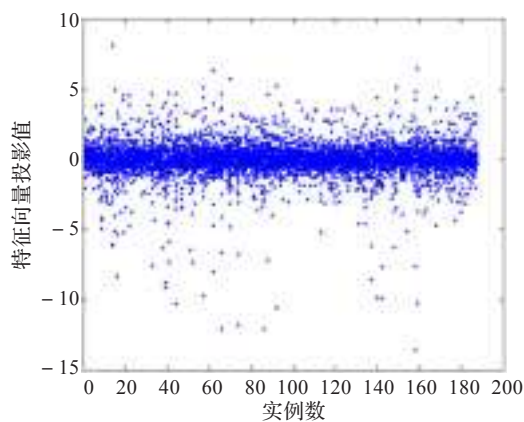


图3 经过PCA降维后的数据分布

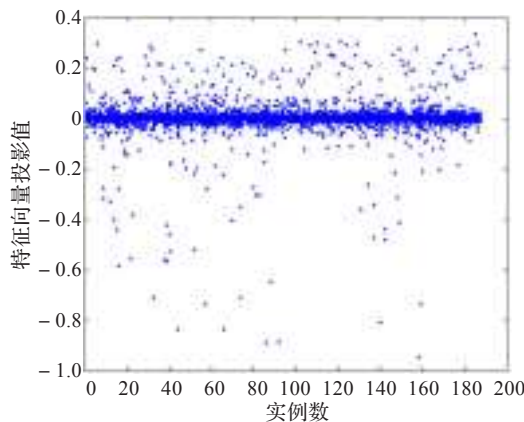


图4 经过KPCA降维后的数据分布

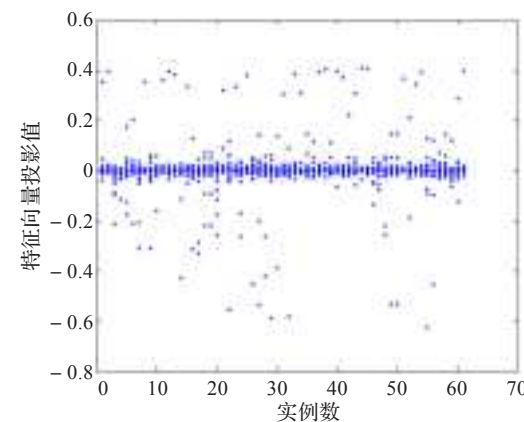


图5 经过GKPCA降维后的数据分布

表2 测试样例的降维结果

实例数×维数	189×13	378×13	567×13	756×13	945×13	1 134×13	1 323×13
PCA	189×11	378×11	567×11	756×11	945×12	1 134×12	1 323×12
KPCA	189×9	378×9	567×9	756×9	945×9	1 134×10	1 323×10
GKPCA	84×7	169×8	266×8	346×8	429×8	481×8	579×9

维,但时间开销较大,GKPCA 算法不仅实现了数据简约,而且缩短了计算时间。

以上测试是在小规模数据集上进行的,对于较大的数据集,采用了“阿拉伯口语数字数据集”(http://archive.ics.uci.edu/ml/machine-learning-databases/00195/),该数据集显示了阿拉伯口语数字的梅尔频率倒普系数时间序列,数据来自阿拉伯本地的44位男性和44位女性。该数据集曾应用于语音识别等领域^[15]。数据集包含两部分,一部分用于实验测试,共有87 063条实例,另一部分用于实验训练,共有269 856条实例,两个数据集中每行都含有13个梅尔频率属性。本实验截取测试数据集的部分数据,逐步增加测试用例,来检验PCA、KPCA与GKPCA的运行效率。

实验分别采取了 $189 \times n (1 \leq n \leq 7)$ 个测试样例,选取提取效率 $p=0.95$,过滤阈值 $\zeta=0.8$,分组数为9,时间消耗如图6所示,降维效果如表2所示。

由图6可以看出,随着数据量的增大,PCA算法处理时间基本上没有发生变化,但KPCA的时间消耗呈指数级增长,在时间效率上,GKPCA对KPCA改进很大。由表2可以看出,PCA降维效果比较差,由原来的13维降为11或12

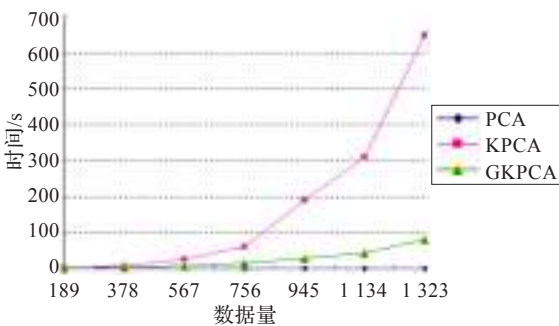


图6 PCA、KPCA与GKPCA时间耗用图

维,KPCA比PCA降维效果好,但是时间消耗大,也不是理想的选择,GKPCA在保证时间消耗可以接受的范围内,有效地实现了数据流降维。

4 结束语

KPCA其特点是特征提取速度快、特征信息保留充分,被广泛应用在信息处理中。实质上是在特征空间中进行PCA分析,在特征空间中,它有着和PCA相同的特性,既保留了PCA的优点,又有处理非线性问题的能力。KPCA又

与PCA有着本质的区别:PCA是基于指标的,KPCA是基于样本的。KPCA不仅适合于解决非线性特征提取问题,而且还能获取比PCA更多的主分量。对于 m 条记录的 n 维样本,PCA能获取 n 个主分量,KPCA能获取 $m-1$ 个主分量。

提出的GKPCA算法是KPCA算法的改进,主要在时间复杂度和空间复杂度上进行优化,提出样本分组和过滤策略,从而简化样本。实验表明,GKPCA算法在保证贡献率不降低的前提下,缩短了高维数据流降维时间,提高了降维效率。但核函数与样本参数都需要人为选择,这就容易对降维造成主观性误差,且随着数据量的增大,GKPCA仍然不会达到PCA的时间标准,KPCA算法仍有待改进。

参考文献:

- [1] De Backer S, Naud A, Scheunders P. Non-linear dimensionality reduction techniques for unsupervised feature extraction[J]. Pattern Recognition Letters, 1998, 19: 711-720.
- [2] 谭璐. 高维数据的降维理论及应用[D]. 长沙: 国防科技大学, 2005: 7-8.
- [3] 吴玲达, 贺玲, 蔡益朝. 高维索引机制中的降维方法综述[J]. 计算机应用研究, 2006, 23(12): 4-7.
- [4] Jolliffe I T. Principal component analysis[M]. 2nd ed. New York: Springer, 2002.
- [5] Saul L K, Roweis S T. An introduction to locally linear embedding[EB/OL]. (2001-03-14). <http://www.cs.nyu.edu/~roweis/lle/publications.html>.
- [6] 梁胜杰, 张志华, 崔立林. 主成分分析法与核主成分分析法在机械噪声数据降维中的应用比较[J]. 中国机械工程, 2011, 22(1): 80-83.
- [7] Tamilnadu S, Tamilnadu C. Enhancing face recognition using improved dimensionality reduction and feature extraction algorithms—an evaluation with orl database[J]. International Journal of Engineering Science and Technology, 2010, 2(6): 2288-2295.
- [8] Teixeira A R, Tomé A M, Stadlthanner K, et al. KPCA denoising and the pre-image problem revisited[J]. Digital Signal Processing, 2008, 18(4): 568-580.
- [9] 王瀛, 郭雷, 梁楠. 基于优选样本的KPCA高光谱图像降维方法[J]. 光子学报, 2011, 40(6): 847-851.
- [10] Burgers K, Fessehatsion Y, Rahmani S, et al. A comparative analysis of dimension reduction algorithms on hyperspectral data[EB/OL]. (2009-07-28). <http://www.math.ucla.edu/~wittman>.
- [11] Rosipal R, Girolami M. An expectation-maximization approach to nonlinear component analysis[J]. Neural Computation, 2001, 13(1): 505-510.
- [12] Baudat G, Anouar F. Generalized discriminant analysis using a kernel function[J]. Neural Computing, 2000, 12(10): 2385-2404.
- [13] 王国胜. 核函数的性质及其构造方法[J]. 计算机科学, 2006: 172-178.
- [14] 高绪伟. 核PCA特征提取方法及其应用研究[D]. 南京: 南京航空航天大学, 2009: 15-17.
- [15] Hammami N, Bedda M. Improved tree model for arabic speech recognition[C]//Proc IEEE ICCSIT10 Conference, 2010.

(上接75页)

从表2和图6可以看出:提出算法的检测正确率优于典型Farid、Q、Shi算法,且随着秘密信息嵌入容量的增加,其检测正确率逐渐增高、检测错误率逐渐降低,适用范围优于典型Farid算法。

5 小结

在对JPEG彩色图像特性、隐写嵌入域及其DCT域相关性等进行全面研究分析的基础上,基于MCM和HVS提出了一种JPEG彩色图像盲隐写分析算法。仿真实验结果表明:该算法在实现对Jsteg、F5、Outguess、MB1、MB2隐写攻击方面,可靠性强、正确率高,并且优于典型的Farid、Q、Shi算法。

参考文献:

- [1] Shi Y Q, Xuan G R, Zou D, et al. Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network[C]//Proceedings of 2005 IEEE International Conference on Multimedia and Expo, Amsterdam, 2005: 269-272.
- [2] Lyu S, Farid H. Steganalysis using higher-order image statistics[J]. IEEE Transactions on Information Forensics and Security, 2006, 1(1): 111-119.
- [3] Fridrich J, Pený T. Multi-class blind steganalysis for JPEG images[C]//Proceedings of SPIE Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII, 2006.
- [4] Shi Y Q, Chen C, Chen W. A Markov process based approach to effective attacking JPEG steganography[C]//Proceedings of the 8th Information Hiding Workshop, 2006.
- [5] Fu D, Shi Y Q, Zou D. JPEG steganalysis using empirical transition matrix in block DCT domain[J]. IEEE, 2006.
- [6] Pevny T, Fridrich J. Merging Markov and DCT features for multi-class JPEG steganalysis[C]//Proc SPIE, 2007.
- [7] 李卓, 陈健, 蒋晓宇, 等. 基于多域特征的JPEG图像盲检测算法[J]. 浙江大学学报:工学版, 2011, 45(9): 1528-1537.
- [8] 盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 3版. 北京: 高等教育出版社, 2001.
- [9] 刘晓芹, 王嘉祯, 徐波, 等. 基于HMT和ICA的主动隐写分析[J]. 武汉大学学报:信息科学版, 2008, 33(10): 1059-1060.
- [10] 冯帆, 王嘉祯, 马懿, 等. 基于HMT的Block-DCT域隐写图像移出攻击[J]. 计算机应用与软件, 2009, 26(11): 274-275.