

文章编号: 2095-0411 (2011) 04-0041-04

一种改进的 ISOMAP 算法在图像检索中的应用*

王洪元^{1,2}, 刘爱萍^{1,2}, 冯 燕^{1,2}

(1. 常州大学 信息科学与工程学院, 江苏 常州 213164; 2. 常州市过程感知与互联技术重点实验室, 江苏 常州 213164)

摘要: 传统的核化 ISOMAP (K-ISOMAP) 算法对于多个分散类簇数据集的低维映射不能较好地表现数据集的内在拓扑结构。针对此缺点, 本文将对基于 ISOMAP 的多类多流形算法 (MCMM-ISOMAP) 进行核化, 提出核化的多类多流形 ISOMAP 算法 (K-MCMM-ISOMAP), 该算法不仅使得多类数据集在降维后保持较好的内在拓扑结构, 而且具备了 K-ISOMAP 算法的泛化能力, 可以将测试数据直接映射到低维空间。因此, 该算法可以在多类图像数据集中实现图像检索的功能。实验结果表明该算法与 K-ISOMAP 相比更具有效性。

关键词: 非线性降维; 核等距特征映射; 多类多流形等距特征映射; 图像检索

中图分类号: TP 391.43

文献标识码: A

An Improved Kernel ISOMAP Algorithm with Application to Image Retrieval

WANG Hong-yuan^{1,2}, LIU Ai-ping^{1,2}, FENG Yan^{1,2}

(1. School of Information Science and Engineering, Changzhou University, Changzhou 213164, China; 2. Changzhou Key Laboratory for Process Perception and Interconnected Technology, Changzhou 213164, China)

Abstract: The conventional kernel ISOMAP algorithm (K-ISOMAP) can not work well in keeping the intrinsic topology of datasets from multi-class clusters datasets in the low-dimensional space. In order to avoid this shortcoming, a novel algorithm named kernel multi-class multi-manifold ISOMAP (K-MCMM-ISOMAP) is proposed in this paper, which is the kernel version of MCMM-ISOMAP. The new algorithm doesn't only keep the intrinsic topology of datasets in low-dimensional mapping space, but also has the generalization of K-ISOMAP. It can directly map the test data to low-dimensional space. Therefore it can be applied to the image retrieval system consisting of multi-class image dataset. The experimental results show that the new algorithm is more effective than the K-ISOMAP.

key words: nonlinear dimensionality; kernel ISOMAP; multi-class multi-manifold ISOMAP; image retrieval

随着信息时代的到来, 多媒体、网络技术的迅速发展, 图像信息的应用日益广泛, 对规模越来越

* 收稿日期: 2011-06-13

基金项目: 国家自然科学基金项目 (61070121, 60973094); 江苏省自然科学基金项目 (BK2009538); 江苏省产学研前瞻性联合研究项目 (BY2009117)

作者简介: 王洪元 (1962—), 男, 江苏常熟人, 教授, 博士, 主要从事模式识别与智能系统研究。

越大的图像数据库、可视信息进行有效的管理成为迫切需要解决的问题。灵活、高效、准确的图像检索策略是解决这一问题的关键技术之一。对大型的图像数据库, 高维向量的存储, 以及高维空间中距离的计算, 其空间复杂度和运算复杂度非常高。若能将高维尺度空间的点集映射到低维尺度空间, 并且保持数据原来的分布结构, 这样, 无论是从存储复杂度, 还是运算复杂度上来讲, 都会得到很大地降低。这就需要运用到数据降维算法, 包括线性降维和非线性降维^[1-3], 比较经典的有 PCA、ICA、LLE、ISOMAP 等算法。

非线性降维算法是近年来一个研究热点。2000 年美国《Science》上发表 3 篇论文^[2-4], 从认知上讨论了流形学习, 并使用了 manifold learning 的术语, 强调认知过程的整体性。Choi 等人^[5]针对 ISOMAP 算法提出了 K-ISOMAP 方法, 通过加常数方法使核矩阵满足正定性。这种 K-ISOMAP 算法具有泛化能力, 可以将测试数据投影到相关的低维流形。Cheng 等人对经典的 ISOMAP 算法进行改进, 提出了多类多流形学习算法 MCMM-ISOMAP (Multi-Class Multi-Manifold ISOMAP)^[6], 主要针对多个分散类簇数据集的维数约简问题, 但是该算法与 ISOMAP 算法一样不能显式产生从高维空间到低维空间的映射函数。本文的创新点在于: 结合 MCMM-ISOMAP 适合于多个分散类簇的数据的优点和 K-ISOMAP 算法具有泛化能力的优点, 提出一种核化的多类多流形 ISOMAP 算法 (K-MCMM-ISOMAP), 该算法不仅能够将测试数据直接映射到低维流形上, 而且能适合数据集分布在多个分散的类簇情况。

本文的主要组织结构是: 首先叙述了传统的 K-ISOMAP 算法, 然后提出 K-MCMM-ISOMAP 算法, 最后将该算法与传统的 K-ISOMAP 方法在图像检索中相比较并得出结论。

1 K-ISOMAP 算法的局限性

传统的 K-ISOMAP 算法是将 ISOMAP 的测地距离矩阵看作一个核矩阵, 主要思想是采用加常数方法, 使得 K 成为一个 Mercer 矩阵 \tilde{K} (即满足正定条件)。K-ISOMAP 所具有的泛化能力 (或者投影性质) 涉及测试数据点在低维空间中的嵌入。也就是说, 泛化能力就是只要给定一个测试数据点 t , 就对应于嵌入在低维空间的点 y 。因此,

K-ISOMAP 要计算测地核 $k(t, x)$, 其中 x 属于训练数据点集, 当测试图像点嵌入到低维空间时, 就可以检索到其周围最近的邻近图像了。

我们可知, K-ISOMAP 是由经典 ISOMAP 核化得到的, 但是 ISOMAP 算法得以成功运用关键在于邻域图构造是否正确, 而构造邻域图依赖于参数 k (本文采用 k 近邻构建邻域图)。对多类图像数据集运用 K-ISOMAP 算法, 当所选择的邻域参数 k 较小时, 各个类内的距离远远小于类间的距离, 虽然类间距离很紧密, 但是图数据点嵌入在二维全局坐标系中时, 有些同图像类的数据点容易映射成为一个点, 不仅不利于类内数据间内在拓扑结构保持, 而且不利于类内图像检索; 相反, 若邻域参数 k 较大时, 虽然能够在整个数据集上产生一个邻域图, 类内数据集在降维后不再出现重复现象, 但是类间数据点在降维后会出现类交叉现象, 同样不利于图像检索。

2 K-MCMM-ISOMAP 算法

由上节可知, 对于多类多流行的图像数据集在使用传统的 K-ISOMAP 算法对图像数据集进行图像检索时存在不足, 本文提出的 K-MCMM-ISOMAP 算法能够有效的解决这一问题。该算法的基本思想是通过求出每个类间最短欧式距离对应的数据点, 并将这些数据点互相设置为邻域点, 从而在整个数据集形成唯一的邻域图, 然后用经典的 Floyd 算法或 Dijkstra 算法计算点与点间的测地线距离, 将得到的测地线距离矩阵作一个核矩阵, 降维后得到数据的内在结构。通过核算法的泛化能力, 对测试的数据点进行低维嵌入, 即可寻找出离该测试数据最近的几个样本数据点。

假设给定 n 维欧式空间的输入数据集为 X , 共有 N 个样本, 分别属于 c 类, 记为 $X_i = \{x_k^{(i)}, k=1, 2, \dots, N_i\}, i=1, 2, \dots, c$, 其中 $\sum_{i=1}^c N_i = N, X_i \cap X_j = \emptyset (1 \leq i, j < c \text{ 且 } i \neq j)$, $x_k^{(i)} \in R^n (1 \leq k \leq N_i)$ 表示第 i 类中的第 k 个样本。相应的低维嵌入输出数据集 $Y_i = \{y_k^{(i)}, k=1, 2, \dots, N_i\}, y_k^{(i)} \in \omega_i \subset R^m (1 \leq k \leq N_i)$ 表示 n 维欧式空间中的样本点 $x_k^{(i)}$ 在 m 维欧式空间中的嵌入, 通常 $m \ll n$ 。

具体步骤如下:

(1) 构造邻域图。选取合适的邻域参数 k 或 ϵ 值, 邻域参数的选取要保证在每个类所对应的流形

内部不产生短路边, 求出任意一个样本点 $x_k^{(i)}$ 对应的邻域点集合 $NE(x_k^{(i)})$ 。

(2) 求出各类数据集间最短的欧式距离分别对应的数据点。即求出 (1) 式的解

$$W_i, W_j = \arg \min_{\substack{w_i \in X_i, 1 \leq i, j \leq c \\ w_j \in X_j, i < j}} (\text{dist}(X_i, X_j)) \quad (1)$$

其中 $w_i \in X_i, w_j \in X_j$ 。(1) 式的含义是求出第 i 类与第 j 类之间最短欧式距离对应的数据点, 并将该数据点分别记为 w_i 和 w_j 。

(3) 对步骤 (2) 中得到的样本点 w_i, w_j 所对应的邻域集合 $NE(w_i), NE(w_j)$ 进行修正: 将 w_i 设置为 w_j 的邻域点, 将 w_j 设置为 w_i 的邻域点, 即使得 $w_j \in NE(w_i), w_i \in NE(w_j)$ 。

(4) 求出所有数据点间的最短路径: 在步骤 (3) 结束后, 在整个数据集上形成了唯一邻域图, 在一个以欧式距离加权的无向邻域图中, 数据点间的最短路径可以使用经典的 Floyd 算法或 Dijkstra 算法求得, 两点间的测地线距离为 D_{ij} , 定义 $D^2 = [D_{ij}^2] \in \mathbf{R}^{N \times N}$;

(5) 在测地线距离矩阵 D^2 的基础上构建核矩阵

$$K(D^2) = -HD^2H/2 \quad (2)$$

其中, H 是中心化矩阵;

(6) 计算下面矩阵的最大特征值 c^*

$$M = \begin{pmatrix} 0 & 2K(D^2) \\ -I & -4K(D) \end{pmatrix} \quad (3)$$

然后构建一个 Mercer 核矩阵 $\tilde{K} = \tilde{K}(\tilde{D}^2)$, 表示如下:

$$\tilde{K} = K(D^2) + 2cK(D) + \frac{1}{2}c^2H \quad (4)$$

当 $c \geq c^* > 0$ 时, \tilde{K} 是正定的。其中, \tilde{D} 具有平移不变性, 满足:

$$\tilde{D}_{ij} = D_{ij} + c(1 - \delta_{ij}) \quad (5)$$

δ_{ij} 为克罗内克符号 (Kronecker Delta)。

(7) 计算 \tilde{K} 中的前 n 个特征向量, 生成特征向量矩阵 $V \in \mathbf{R}^{N \times n}$ 和特征值矩阵 $\Lambda \in \mathbf{R}^{n \times n}$;

(8) n 维 Euclidean 空间中的 N 个点的坐标由 $Y = \Lambda^{-\frac{1}{2}}V^T$ 的列向量给出。

由上述步骤可知, K-ISOMAP 算法与所提出的算法主要区别在于构造多个分散类簇的邻域图上, 即第 (1) 步到第 (4) 步。

3 实验

3.1 实验数据

为了验证本文所提出的算法的性能, 在实际数据集上进行实验, 编程环境为 Matlab R2008a。本实验使用 4 类 (鸭子、积木、猫、人脸) 图像作实际数据集, 每类 72 幅图像, 其中鸭子、积木和猫这 3 类数据集来源于哥伦比亚大学计算机视觉实验室^[7], 人脸图像集是从经典 ISOMAP 实验所使用的人脸数据集中挑选出一幅正面图像, 然后将其每隔 5 度旋转得到的。所有图像的分辨率图被压缩到 32×32 。

3.2 实验结果

图 1 给出了 $k=10$ 时运行 K-ISOMAP 算法后的效果图。不同的形状表示不同的类, 加号表示鸭子, 实心表示人脸, 点号表示积木, 星号表示猫, 由图可见同类图像的内在结构没有很好的表现出来, 且不同类图像之间出现了类交叉现象。图 2 为图 1 的部分放大, 在这部分用 K-ISOMAP 算



图 1 K-ISOMAP 算法的实验效果 ($k=10$)

Fig. 1 The result of K-ISOMAP ($k=10$)

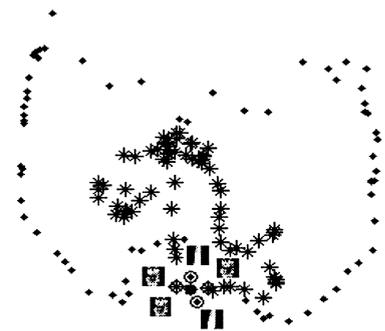


图 2 K-ISOMAP 算法得到的检索效果图

Fig. 2 The retrieval result of K-ISOMAP

法做图像检索, 用猫图片作为测试数据点 (较大的实心点为测试点), 让其检索相邻的 4 个图像数据点 (分别为 2 个点号和 2 个星号, 图中已经用空心圈圈出来了), 可见结果并不理想, 因为用猫的图

像检索到了两幅猫的图像的同时也检索到了两幅积木的图像。

图 3 为当 $k=4$ 时用改进的算法降维后的效果图, 可以看出改进的算法能够成功地将这 4 类图像数据集嵌入在二维全局坐标系内, 可以很清晰的看出四类数据集在二维空间的效果。将测试数据点嵌入到低维特征空间中, 图 4 是用改进算法做的图像检索效果图, 星号表示的点为测试的数据点, 检索到它周围最接近的 4 个图像数据点 (用空心圈圈出来), 可见能够很好的达到图像检索的效果。

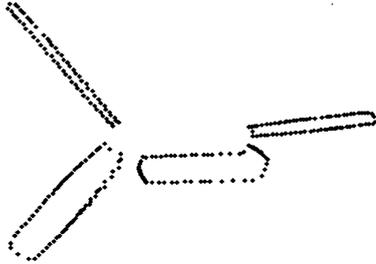


图 3 用改进算法后降维得到的二维效果 ($k=4$)
Fig. 3 The result of improved algorithm ($k=4$)

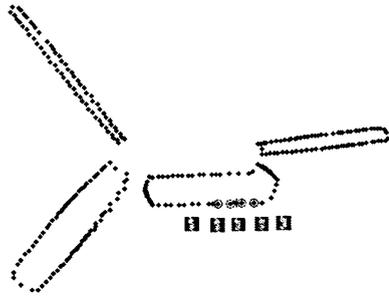


图 4 星号表示的是测试的数据点, 它检索到 4 个数据点
Fig. 4 The retrieval result of K-MCMM-ISMOAP, asterisk indicates the test point, it retrieves four nearest image data points

3.3 实验分析

用 K-ISOMAP 算法, k 值太大和太小都不利于数据集的降维。本实验中, 当 $k=10$ 时, 数据集的降维效果最佳。由图 1 可见, 4 类图像数据集没有全部很好的分开, 也没有很好的保持类内拓扑结构。图 2 中, 由于传统的 K-ISOMAP 算法在降维时没有很好的保持类内和类间结构, 因此导致了类交叉现象, 在做图像检索时, 用猫的图像检索到了两幅猫的图像和两幅积木的图像, 从而影响

了检索效果。

用改进的算法能够成功地将这四类图像嵌入在二维全局坐标系内。二维空间的效果很清晰, 不但不同的类数据集得以很好的分开, 没有出现类交叉现象, 而且各个类内数据集也保持了很好的拓扑结构。因此, 由图 4 中可见, 在应该新的算法做图像检索时, 也达到较佳的效果。

4 总 结

本文主要针对传统 K-ISOMAP 算法对多类图像数据集进行图像检索时存在的缺陷, 提出了 K-MCMM-ISOMAP 算法, 该算法对于多类图像数据集先进行预处理, 在保持图像数据集类内结构的基础上, 再求出图像数据集类间的距离, 并对图像数据集类间距离给予适当的设置, 使得整体多类图像数据集能够构造出很好的结构。该算法能够将测试点直接从高维空间映射到低维空间, 从而在多类图像数据集中能实现图像检索的功能。本文在图像数据集上做了实验, 表明新算法能够保持多类数据集的内在结构, 证实其有效性和实用性。

参考文献:

- [1] Jolliffe I T. Principal Component Analysis [M]. New York: Springer-Verlag, 1986.
- [2] Tenenbaum J B, Silva V de, Langford J C. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290 (12): 2 319-2 323.
- [3] Roweis S, Saul L. Nonlinear dimensionality analysis by locally linear embedding [J]. Science, 2000, 290 (5500): 2 323-2 326.
- [4] Seung H S, Lee D D. The manifold ways of perception [J]. Science, 2000, 290 (5500): 2 268-2 269.
- [5] Choi H, Choi S. Kernel Isomap [J]. Electronics Letters, 2005, 40 (25): 1 612-1 613.
- [6] 程起才, 王洪元, 刘爱萍, 等. 基于 ISOMAP 的一种多流行学习算法 [J]. 微电子学与计算机, 2009, 26 (10), 115-117.
- [7] Nene S A, Nayar S K, Murase H. Columbia automated vision environment; columbia university image library (coil-20) [EB/OL]. (2010-09-11) [2011-08-04]. <http://www1.cs.columbia.edu/CAVE/>. 1996.