



(12) 发明专利

(10) 授权公告号 CN 108764351 B

(45) 授权公告日 2021.08.31

(21) 申请号 201810542508.2

CN 106250918 A, 2016.12.21

(22) 申请日 2018.05.30

US 2014207710 A1, 2014.07.24

(65) 同一申请的已公布的文献号

US 2015193910 A1, 2015.07.09

申请公布号 CN 108764351 A

US 2014279755 A1, 2014.09.18

CN 107146219 A, 2017.09.08

(43) 申请公布日 2018.11.06

顾艳春. 一种改进的局部切空间排列算法.

(73) 专利权人 佛山科学技术学院

《计算机应用研究》. 2013, 第30卷 (第3期),

地址 528000 广东省佛山市南海区狮山镇

Mehrtash T. Harandi .etc. Sparse

仙溪水库西路佛山科学技术学院

Coding on Symmetric Positive Definite

(72) 发明人 牛菓 王修才 顾艳春 段志奎

Manifolds Using Bregman Divergences.《IEEE

陈建文 樊耘

TRANSACTIONS ON NEURAL NETWORKS AND

LEARNING SYSTEMS》. 2016, 第27卷 (第6期),

(74) 专利代理机构 广州嘉权专利商标事务所有

王瑞霞等. 基于黎曼流形稀疏编码的图像检

限公司 44205

索算法.《自动化学报》. 2017, 第43卷 (第5期),

代理人 王国标

刘兵. 基于聚类与流形正则化的分类方法研

(51) Int. Cl.

究.《中国博士学位论文全文数据库》. 2014,

G06K 9/62 (2006.01)

仲硕. 基于实对称矩阵判别学习的图像分类

(56) 对比文件

方法研究.《中国优秀硕士学位论文全文数据库

CN 103500450 A, 2014.01.08

库》. 2017,

CN 102298703 A, 2011.12.28

审查员 李娜

CN 107748871 A, 2018.03.02

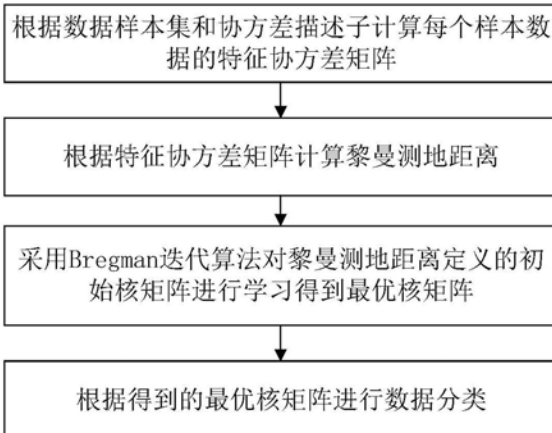
权利要求书2页 说明书12页 附图1页

(54) 发明名称

一种基于测地距离的黎曼流形保持核学习方法及装置

(57) 摘要

本发明公开了一种基于测地距离的黎曼流形保持核学习方法及装置,用于解决数字图像的黎曼流形数据的核学习方法的两个问题:1) 测度保持问题,通过参数化马氏距离学习的方式,使得黎曼流形上的对称正定矩阵变换到欧式空间上的距离与黎曼流形上的测地距离相等,根据Bregman优化算法求得马氏距离矩阵的最优解;2) 固定核方法问题.通过核学习的方式;根据Bregman优化算法,进行核学习得到最优的核矩阵,使得核空间上的距离与测地距离保持一致,最终有效的解决了黎曼流形测度保持以及单一核局限性的问题。



1. 一种基于测地距离的黎曼流形保持核学习方法,其特征在于,所述方法包括以下步骤:

步骤1,根据数据样本集和协方差描述子计算每个样本数据的特征协方差矩阵;

步骤2,根据特征协方差矩阵计算黎曼测地距离;

步骤3,采用Bregman迭代算法对黎曼测地距离定义的初始核矩阵进行学习得到最优核矩阵;

步骤4,根据得到的最优核矩阵进行数据分类;

在步骤1中,所述计算每个样本数据的特征协方差矩阵的方法为:

采用区域协方差算子方法计算每个样本的特征协方差矩阵,设I为一个三维颜色强度的样本图片,F为从I中抽取的 $W \times H \times d$ 维的特征图像: $F(x, y) = \Phi(I, x, y)$,如果I为一个三维图片,F是从I中截取的子图像的特征图像,F的大小为 $W \times H \times d$ 维,x,y的大小范围在W,H内,d为F(x,y)中元素的个数,函数 Φ 是图像强度,颜色,梯度,滤波器响应的映射其中一种,一般特征向量F(x,y):

$$F(x, y) = \left[x, y, R(x, y), G(x, y), B(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T, \quad (x, y) \text{ 表示像}$$

素的位置, $R(x, y), G(x, y), B(x, y)$ 表示像素的颜色值, $\left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|$ 分别表示x,y的一

阶微分的范数, $\left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right|$ 分别表示x,y的二阶微分的范数,对于纹理识别任务,

构造特征向量F(x,y):

$$F(x, y) = \left[I(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T$$

其中,I(x,y)表示图像强度,在构造特征向量时,常用的特征值有图像强度值、相对x,y的一阶和二阶微分的范数,对于给定的矩形区域 $R \subset F$, $\{z_k\}$ 是区域R中d维特征, $k=1 \cdots n$,n

为正整数,用区域R中特征的 $d \times d$ 的协方差矩阵表示这个区域R: $C_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - u)(z_k - u)^T$,u为特征点的均值,对于式中协方差矩阵的第(i,j)个元素表示为, $C_R(i, j) = \frac{1}{n-1} \sum_{k=1}^n (z_k(i) - u(i))(z_k(j) - u(j))$,

展开均值,将上式改为, $C_R(i, j) = \frac{1}{n-1} \left[\sum_{k=1}^n z_k(i) z_k(j) - \frac{1}{n} \sum_{k=1}^n z_k(i) \sum_{k=1}^n z_k(j) \right]$,计算各维特征

的和 $\sum_{k=1}^n z_k(i)$,以及任意两个特征的乘积的和 $z(i) z(j)$, $i=1 \cdots n, j=1 \cdots n$,n为正整数,得到

区域协方差 C_R ;

在步骤2中,所述根据特征协方差矩阵计算黎曼测地距离方法为:

假设给定的数据集为 $X = \{x_1 \cdots x_n\}$,其中, x_i 表示数据样本, $i=1, \cdots, n$,n为正整数;

对任意两个对称正定矩阵 X_i, X_j 进行对数操作得到 $Z_i = \log(X_i), Z_j = \log(X_j)$,经过对数

计算后, Z_i, Z_j 依旧是一个矩阵, 进而, 把 Z_i, Z_j 矩阵处理成一个列向量 $\hat{X}_i, \hat{X}_j$, 新得到的列向量 $\hat{X}_i, \hat{X}_j$ 作为欧式空间中的点, 利用参数化马氏距离形式得到:

$$d_A^2(\hat{X}_i, \hat{X}_j) = (\hat{X}_i - \hat{X}_j)^T A (\hat{X}_i - \hat{X}_j), \text{ 对上式矩阵 } A \text{ 的学习利用 LogDet 散度来解决, 即, 其}$$

中, 对于 $n \times n$ 矩阵 A 和 A_0 , A 和 A_0 的 LogDet 散度为: $D_{ld}(A, A_0) = \text{tr}(AA_0^{-1}) - \log \det(AA_0^{-1}) - n$, 为了预防最小化 LogDet 散度没有可行的解决方案, 在表述形式中加入松弛变量 γ 来保证有最优解, 对称正定矩阵经变换之后的点在欧式空间上进行马氏距离的学习之后, 得到

$$d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij}, \text{ 加上松弛变量 } \gamma \text{ 后, 矩阵 } A \text{ 的学习问题表示为: } \min_{A>0} D_{ld}(A, A_0) + \gamma \|\zeta\|_2^2,$$

$$\text{s.t. } d_A^2(\hat{X}_i, \hat{X}_j) \approx d_{ges}^2(X_i, X_j), \text{ 其中, } \zeta \text{ 是变量 } \zeta_{ij} \text{ 的列向量, } d_A^2(\hat{X}_i, \hat{X}_j) = (\hat{X}_i - \hat{X}_j)^T A (\hat{X}_i - \hat{X}_j)$$

是欧式空间里两点的向量之间马氏距离的定义, d_A 为马氏距离, 黎曼流形上的数据之间的测地距离, 把流形数据转换到欧式空间, 用马氏距离近似等于测地距离, 即用

$$d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij} \text{ 来保证, } d_{ges} \text{ 为测地距离, 所以矩阵 } A \text{ 就包含有测距的距离信息。}$$

2. 根据权利要求1所述的一种基于测地距离的黎曼流形保持核学习方法, 其特征在于, 在步骤3中, 初始核矩阵根据马氏距离的矩阵 A , 定义初始核矩阵 $K_0 = \hat{X}^T A^* \hat{X}$ 。

3. 根据权利要求1所述的一种基于测地距离的黎曼流形保持核学习方法, 其特征在于, 在步骤3中, 采用 Bregman 迭代算法对核矩阵进行学习, 使得到的最优核矩阵 K^* 逼近 K_0 , 最终得到的核矩阵满足核空间上的距离等于黎曼流形上的测地距离, 对于 $n \times n$ 矩阵 K 和 K_0 ,

K 和 K_0 的 LogDet 散度为: $D_{ld}(K, K_0) = \text{tr}(KK_0^{-1}) - \log \det(KK_0^{-1}) - n$, 其中 n 为输入空间的维度, 令 $K_{ii} + K_{ii} - K_{ij} - d_{ges}^2(X_i, X_j) = \zeta_{ij}$, 加上松弛变量 γ 后, 表述形式为:

$$\min_{K>0} D_{ld}(K, K_0) + \gamma \|\zeta\|_2^2, \text{ s.t. } K_{ii} + K_{ii} - K_{ij} \approx d_{ges}^2(X_i, X_j), X_i, X_j \text{ 为黎曼流形正定矩阵, 其中 } \zeta \text{ 是变量 } \zeta_{ij} \text{ 的列向量。}$$

4. 一种基于测地距离的黎曼流形保持核学习装置, 其特征在于, 所述装置包括: 存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序, 所述处理器执行所述计算机程序时实现权利要求1-3任一项所述一种基于测地距离的黎曼流形保持核学习方法的步骤。

一种基于测地距离的黎曼流形保持核学习方法及装置

技术领域

[0001] 本公开涉及机器学习及图像处理领域,具体涉及一种基于测地距离的黎曼流形保持核学习方法及装置。

背景技术

[0002] 随着多媒体技术的快速发展以及互联网技术的迅猛发展,数字图像信息的采集,存储,传播以及访问呈现出爆炸性的增长。数字图像不仅充斥在互联网上,而且在民用,商业,军事,医疗,生物等领域中都会产生大量的图像信息,包括日常生活中拍摄的照片,新闻图片,医学图片,生物图片,遥感图片等等。上述的种种情况,使得图像的处理变得越来越复杂。由于图像处理所面对的图像数据,数量多且维数高,如果采用传统方法对图像进行处理,会导致高昂的时间成本以及识别率不理想的情况发生。比如图像处理常用的方法里,将一张图像变换成一个行向量或者列向量,即使用一个行向量或列向量来表示一张图像,如果图像维数相对小一些或者维数不太高,这样变换后,运算量不会太大。但是随着互联网技术的迅速发展,图像越来越清晰,维数也越来越大,不再局限在几百维以内,而是几万,几十万维,甚至更大。如果此时再将图像变换成向量,进行后续的分类、识别等应用时,计算量剧增,导致高昂的时间成本产生。每个图像都会有自己的特征,一般情况下,对一个图像而言,常用特征提取方法有协方差描述子(文献1,Tuzel O.,Porikli F.,Meer F.,Region Covariance:A Fast Descriptor for Detection and Classification,European Conference on Computer Vision(ECCV),2006)、扩散张量描述子(文献2,Pennec X.,Fillard P.,Ayache N.,A Riemannian Framework for Tensor Computing,International Journal of Computer Vision(IJCV),2006)、结构张量描述子(文献3,Goh A.,Vidal R.,Clustering and dimensionality reduction on Riemannian manifolds.Computer Vision and Pattern Recognition(CVPR),2008)等。使用协方差矩阵作为区域描述子不但可以融合图像里多个可能相关的特征,还可以从一个区域提取的单个协方差矩阵,来实现在不同的视角和姿势下该区域的匹配问题。

[0003] 测度学习与核学习在机器学习中应用广泛。普遍认为测度学习来源于2002年Xing等人的工作(文献4,Xing E.P.,Jordan M.I.,Russell S.,Distance metric learning with application to clustering with side-formation.Proceeding of the Advances in Neural Information Proceeding Systems,2002)。Davis等人(文献5,Davis J.V.,Kulis B.,Jani P.,Information-theoretic metric learning,Proceedings of the International Conference on Machine Learning,2007)将测度学习问题转化为Bregman问题,并从信息论角度提出了信息理论度量学习。这样在距离函数约束条件下,测度学习变为通过最小化多变量高斯之间的微分相对熵的问题。Davis等人通过使用Bregman算法,将测度学习问题表述为一种特殊的Bregman优化问题,最后通过最小化LogDet散度,来学习数据之间的测度。现有的测度学习方法,大多依赖于马氏距离,且局限于低维数据。核学习(文献6Shawe-Taylor J.and Cristianini N.,Kernel methods for pattern

analysis.Cambridge University Press,2004) 在机器学习中也有广泛的应用。但由于核方法大多是直推式的,不能有效地拓展到新来的数据应用中。然而把测度学习方法引入到核学习,可以有效的改进现有的核方法。有些研究者在把测度学习作为线性转换学习问题时,联合核学习来改进算法。具体做法是将学习问题表示为满足一定线性条件下,最小化LogDet散度的问题。但是一般来讲,线性约束含有分类信息,需要在全监督或者半监督情形下进行。

发明内容

[0004] 本公开的目的是针对现有技术的不足,提出一种基于测地距离的黎曼流形保持核学习方法及装置,通过把黎曼流形上的任意两点之间的测地距离引入测度学习,来优化初始核矩阵,进一步将核学习的问题表示为满足一定线性条件下,最小化LogDet散度的问题,最终根据得到的核矩阵进行数据分类。这样做的优点是:提出的约束条件是保持流形上的测地距离,即核空间上两点之间的距离等于其测地距离,并不含有分类信息,是一种无监督学习方式。同时根据得到的最优马氏距离,带入到线性变换 $K_0 = X^T A^* X$,将 K_0 作为初始矩阵,再进行核学习,这样使得最后得到的核矩阵满足核空间上两点之间的距离等于其测地距离,根据给定数据集,计算每个样本数据特征协方差矩阵,应用黎曼测地距离到测度学习,再根据得到最优测度来计算核学习问题的初始核矩阵,最后根据Bregman迭代方法求最优核矩阵,本发明提高流形学习分类方法的效率。

[0005] 为了实现上述目的,本公开提供一种基于测地距离的黎曼流形保持核学习方法及装置,其中,所述方法包括以下步骤:

[0006] 步骤1,根据数据样本集和协方差描述子计算每个样本数据的特征协方差矩阵;

[0007] 步骤2,根据特征协方差矩阵计算黎曼测地距离;

[0008] 步骤3,采用Bregman迭代算法对黎曼测地距离定义的初始核矩阵进行学习得到最优核矩阵;

[0009] 步骤4,根据得到的最优核矩阵进行数据分类。

[0010] 进一步地,在步骤1中,所述计算每个样本数据的特征协方差矩阵的方法为:

[0011] 采用区域协方差算子方法计算每个样本的特征协方差矩阵,设 I 为一个三维颜色强度的样本图片, F 为从 I 中抽取的 $W \times H \times d$ 维的特征图像: $F(x, y) = \Phi(I, x, y)$,如果 I 为一个三维图片, F 是从 I 中截取的子图像的特征图像, F 的大小为 $W \times H \times d$ 维, x, y 的大小范围在 W, H 内, d 为 $F(x, y)$ 中元素的个数,函数 Φ 是图像强度,颜色,梯度,滤波器响应的映射其中一种,一般特征向量 $F(x, y)$:

$$[0012] \quad F(x, y) = \left[x, y, R(x, y), G(x, y), B(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T, \quad (x, y) \text{ 表}$$

示像素的位置, $R(x, y), G(x, y), B(x, y)$ 表示像素的颜色值, $\left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|$ 分别表示相

对 x, y 的一阶微分的范数, $\left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right|$ 分别表示相对 x, y 的二阶微分的范数,对于

纹理识别任务,构造特征向量 $F(x, y)$:

$$[0013] \quad F(x, y) = \left[I(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T$$

[0014] 其中, $I(x, y)$ 表示图像强度, 在构造特征向量时, 常用的特征值有图像强度值、相对 x, y 的一阶和二阶微分的范数, 对于给定的矩形区域 $R \subset F$, $\{z_k\}$ 是区域 R 中 d 维特征, $k = 1 \cdots n$, n 为正整数, 用区域 R 中特征的 $d \times d$ 的协方差矩阵表示这个区域 R : $C_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - u)(z_k - u)^T$, u 为特征点的均值, 对于式中协方差矩阵的第 (i, j) 个元素表示为,

$$C_R(i, j) = \frac{1}{n-1} \sum_{k=1}^n (Z_k(i) - u(i))(Z_k(j) - u(j))$$
 , 展开均值, 上式可以改为,

$$C_R(i, j) = \frac{1}{n-1} \left[\sum_{k=1}^n Z_k(i) Z_k(j) - \frac{1}{n} \sum_{k=1}^n Z_k(i) \sum_{k=1}^n Z_k(j) \right],$$

[0015] 计算各维特征的和 $\sum_{k=1}^n Z_k(i)$, 以及任意两个特征的乘积的和 $Z(i) Z(j)$, $Z(i) Z(j)$, $i = 1 \cdots n, j = 1 \cdots n$, n 为正整数, 就可以得到区域协方差 C_R 。

[0016] 假设给定的数据集为 $X = \{x_1 \cdots x_n\}$, 其中, x_i 表示数据样本, $x_i, i = 1, \cdots, n$, n 为正整数;

[0017] 进一步地, 根据特征协方差矩阵, 计算黎曼测地距离。由步骤1得到特征协方差矩阵均为对称正定矩阵, 而所有的对称矩阵就是一个黎曼流形。对任意两个对称正定矩阵 X_i, X_j 进行对数操作, 可以得到 $Z_i = \log(X_i), Z_j = \log(X_j)$ 。经过对数计算后, Z_i, Z_j 依旧是一个矩阵。注意, 这里的对数运算是黎曼流形特殊运算符, 它是依赖数据的。进而, 把 Z_i, Z_j 矩阵处理成一个列向量 $\hat{X}_i, \hat{X}_j$ 。新得到的列向量 $\hat{X}_i, \hat{X}_j$ 可以看作是欧式空间中的点, 利用参数化马氏距离形式, 可以得到:

$$[0018] \quad d_A^2(\hat{X}_i, \hat{X}_j) = \left(\hat{X}_i - \hat{X}_j \right)^T A \left(\hat{X}_i - \hat{X}_j \right).$$

[0019] 对上式矩阵 A 的学习, 可以利用LogDet散度来解决。即, 其中对于 $n \times n$ 矩阵 A 和 A_0 , A 和 A_0 的LogDet散度为:

$$[0020] \quad D_{ld}(A, A_0) = \text{tr}(AA_0^{-1}) - \log \det(AA_0^{-1}) - n.$$

[0021] 为了预防最小化LogDet散度没有可行的解决方案, 在表述形式中加入松弛变量 γ 来保证有最优解。在本发明里, 为达到保持测地距离的目的, 对称正定矩阵经变换之后的点在欧式空间上进行马氏距离的学习之后, 得到 $d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 。加上松弛变量 γ 后, 矩阵 A 的学习问题可以表示为:

$$[0022] \quad \min_{A>0} D_{ld}(A, A_0) + \gamma \|\zeta\|_2^2,$$

$$[0023] \quad s.t. \quad d_A^2(\hat{X}_i, \hat{X}_j) \approx d_{ges}^2(X_i, X_j)。$$

[0024] 进一步地,在步骤2中,所述根据特征协方差矩阵计算黎曼测地距离方法为:

[0025] 对任意两个对称正定矩阵 X_i, X_j 进行对数操作,可以得到 $Z_i = \log(X_i), Z_j = \log(X_j)$ 。经过对数计算后, Z_i, Z_j 依旧是一个矩阵。注意,这里的对数运算是黎曼流形特殊运算符,它是依赖数据的。进而,把 Z_i, Z_j 矩阵处理成一个列向量 $\hat{X}_i, \hat{X}_j$ 。新得到的列向量 $\hat{X}_i, \hat{X}_j$ 可以看作是欧式空间中的点,利用参数化马氏距离形式,可以得到:

$$[0026] \quad d_A^2(\hat{X}_i, \hat{X}_j) = (\hat{X}_i - \hat{X}_j)^T A (\hat{X}_i - \hat{X}_j), \text{对上式矩阵A的学习,可以利用LogDet散度来}$$

解决。即,其中对于 $n \times n$ 矩阵A和 A_0 ,A和 A_0 的LogDet散度为: $D_{ld}(A, A_0) = \text{tr}(AA_0^{-1}) - \log \det(AA_0^{-1}) - n$,为了预防最小化LogDet散度没有可行的解决方案,在表述形式中加入松弛变量 γ 来保证有最优解。在本发明里,为达到保持测地距离的目的,对称正定矩阵经变换之后的

点在欧式空间上进行马氏距离的学习之后,得到 $d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 。加上松弛变

量 γ 后,矩阵A的学习问题可以表示为: $\min_{A>0} D_{ld}(A, A_0) + \gamma \|\zeta\|_2^2, \quad s.t. \quad d_A^2(\hat{X}_i, \hat{X}_j) \approx d_{ges}^2(X_i, X_j),$

其中, ζ 是变量 ζ_{ij} 的列向量, $d_A^2(\hat{X}_i, \hat{X}_j) = (\hat{X}_i - \hat{X}_j)^T A (\hat{X}_i - \hat{X}_j)$ 是欧式空间里两点(向量)之

间马氏距离的定义, d_A 为马氏距离,黎曼流形上的数据之间的测地距离,不能直接采用欧式距离计算,所以这里我们先把流形数据转换到欧式空间,用马氏距离近似等于测地距离,即

用 $d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 来保证,其中, d_{ges} 为测地距离,所以矩阵A就包含有测距的

距离信息。

[0027] 进一步地,在步骤3中,初始核矩阵根据马氏距离的矩阵A,定义初始核矩阵

$$K_0 = \hat{X}^T A^* \hat{X}。$$

[0028] 进一步地,在步骤3中,采用Bregman迭代算法对核矩阵进行学习,使得到的最优核矩阵 K^* 逼近 K_0 ,最终得到的核矩阵满足核空间上的距离等于黎曼流形上的测地距离,对于 $n \times n$ 矩阵K和 K_0 ,

[0029] K和 K_0 的LogDet散度为: $D_{ld}(K, K_0) = \text{tr}(KK_0^{-1}) - \log \det(KK_0^{-1}) - n$,其中n为输入空间的维度,令 $K_{ii} + K_{ii} - K_{ij} - d_{ges}^2(X_i, X_j) = \zeta_{ij}$,加上松弛变量 γ 后,表述形式为:

$$\min_{K>0} D_{ld}(K, K_0) + \gamma \|\zeta\|_2^2, \quad s.t. \quad K_{ii} + K_{ii} - K_{ij} \approx d_{ges}^2(X_i, X_j), \text{其中}\zeta\text{是变量}\zeta_{ij}\text{的列向量}, X_i, X_j\text{为黎曼流形正定矩阵。}$$

[0030] 进一步地,在步骤3中,初始核矩阵根据马氏距离的矩阵A,定义初始核矩阵

$$K_0 = \hat{X}^T A^* \hat{X}, \quad \hat{X}_i, \hat{X}_j \text{为} X_i, X_j \text{的列向量,采用Bregman迭代算法对核矩阵进行学习,使得}$$

到的最优核矩阵 K^* 逼近 K_0 ,使最终得到的核矩阵满足核空间上的距离等于黎曼流形上的测地距离,从而达到测地距离保持的目的。根据测度学习与核学习之间的联系 $K_0 = \hat{X}^T A^* \hat{X}$,可以根据测度学习得到适合问题最优解 A^* ,然后直接得到 K^* 。本发明采用将得到的最优马氏矩阵带入线性变换 $K_0 = \hat{X}^T A^* \hat{X}$,在LogDet散度下,进行核学习,得到最优的核矩阵,使得核空间上的距离与黎曼流形上的测地距离保持一致。

[0031] 对于 $n \times n$ 矩阵 K 和 K_0 , K 和 K_0 的LogDet散度为:

[0032] $D_{ld}(K, K_0) = \text{tr}(KK_0^{-1}) - \log \det(KK_0^{-1}) - n$,其中 n 为输入空间的维度。

[0033] 为了预防最小化LogDet散度没有可行的解决方案,在表述形式中加入松弛变量 γ 来保证可行方案的存在。本发明的约束条件是核空间上两点间的距离等于黎曼流形上两点间的测地距离,即令 $K_{ii} + K_{ii} - K_{ij} - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 。

[0034] 加上松弛变量 γ 后,表述形式为:

[0035] $\min_{K>0} D_{ld}(K, K_0) + \gamma \|\zeta\|_2^2$

[0036] s.t. $K_{ii} + K_{ii} - K_{ij} \approx d_{ges}^2(X_i, X_j)$

[0037] 其中 ζ 是变量 $\zeta_{i,j}$ 的列向量。

[0038] 进一步地,根据得到的最优核矩阵 K^* ,进行数据分类。

[0039] 数据分类说明:采用线性分类,即可以直接计算两点之间的相似性,然后分类;采用核方法分类就是把数据影射到核空间,再分类,这是一种非线性分类方法。在这个过程中需要计算核矩阵(一般有直接采用核函数比如高斯核函数,或学习一个核函数或者核矩阵,我们这里属于第二种)。

[0040] 本发明还提供了一种基于测地距离的黎曼流形保持核学习装置,所述装置包括:存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序运行在以下装置的单元中:

[0041] 黎曼流形计算单元,用于根据数据样本集和协方差描述子计算每个样本数据的特征协方差矩阵;

[0042] 黎曼测地距离计算单元,用于根据特征协方差矩阵计算黎曼测地距离;

[0043] 迭代核优化单元,用于采用Bregman迭代算法对黎曼测地距离定义的初始核矩阵进行学习得到最优核矩阵;

[0044] 最优核分类单元,用于根据得到的最优核矩阵进行数据分类。

[0045] 本公开的有益效果为:本公开提出了一种新的在数字图像的黎曼流形数据上基于测地距离保持的核学习方法,解决核学习方法的两个问题:1)测度保持问题。根据黎曼流形上的Log-Euclidean测地距离,本文将对数对称正定矩阵处理成一个列向量,通过参数化马氏距离学习的方式,使得黎曼流形上的对称正定矩阵变换到欧式空间上的距离与黎曼流形上的测地距离相等。同时将马氏距离学习问题表述为最小化LogDet散度,根据Bregman优化算法求得马氏距离矩阵的最优解。2)固定核方法问题。本方法通过核学习的方式,找到适合问题最优的核。同时将核学习问题表述为最小化LogDet散度。将得到的最优马氏距离矩阵,

根据马氏距离矩阵与核矩阵之间的线性变换关系,赋予核矩阵初值。然后根据Bregman优化算法,进行核学习得到最优的核矩阵,使得核空间上的距离与测地距离保持一致。最终有效的解决了黎曼流形测度保持以及单一核局限性的问题。

附图说明

[0046] 通过对结合附图所示出的实施方式进行详细说明,本公开的上述以及其他特征将更加明显,本公开附图中相同的参考标号表示相同或相似的元素,显而易见地,下面描述中的附图仅仅是本公开的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图,在附图中:

[0047] 图1所示为本公开的一种基于测地距离的黎曼流形保持核学习方法的流程图;

[0048] 图2所示为本公开的一种基于测地距离的黎曼流形保持核学习装置图。

具体实施方式

[0049] 以下将结合实施例和附图对本公开的构思、具体结构及产生的技术效果进行清楚、完整的描述,以充分地理解本公开的目的、方案和效果。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0050] 如图1所示为根据本公开的一种基于测地距离的黎曼流形保持核学习方法的流程图,下面结合图1来阐述根据本公开的实施方式的一种基于测地距离的黎曼流形保持核学习方法。

[0051] 本发明主要提供一种基于测地距离的黎曼流形保持核学习方法,所述方法的具体实施方式如下:

[0052] 假设给定的数据集为 $X = \{x_1 \cdots x_n\}$,其中 x_i 表示数据样本, $i = 1, \cdots, n$,注意的样本数据是矩阵,不需要转换为向量。

[0053] 计算每个样本的特征协方差矩阵。这里采用区域协方差算子方法。

[0054] 设 I 为一维或者三维颜色强度的样本图片。 F 为从 I 中抽取的 $W \times H \times d$ 维的特征图像:

[0055] $F(x, y) = \Phi(I, x, y)$,其中 W, H, d 为空间的维度,

[0056] 函数 Φ 是任意的映射比如图像强度,颜色,梯度,滤波器响应等。

[0057] 一般特征向量 $F(x, y)$:

[0058] $F(x, y) = \left[x, y, R(x, y), G(x, y), B(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T$, (x, y) 表

示像素的位置, $R(x, y), G(x, y), B(x, y)$ 表示像素的颜色值, $\left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|$ 分别表示相

对 x, y 的一阶微分的范数, $\left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right|$ 分别表示相对 x, y 的二阶微分的范数。对于

纹理识别任务,一般构造特征向量 $F(x, y)$:

$$[0059] \quad F(x, y) = \left[I(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right| \right]^T,$$

[0060] $I(x, y)$ 表示图像强度, 其他同上。在构造特征向量时, 常用的特征值有图像强度值, 相对 x, y 的一阶, 二阶微分的范数。

[0061] 对于给定的矩形区域 $R \subset F$, $\{z_k\}_{k=1 \dots n}$ 是区域 R 中 d 维特征。用区域 R 中特征的 $d \times d$ 的协方差矩阵表示这个区域 R :

$$[0062] \quad C_R = \frac{1}{n-1} \sum_{k=1}^n (Z_k - u)(Z_k - u)^T,$$

[0063] u 为特征点的均值。对于式中协方差矩阵的第 (i, j) 个元素表示为:

$$[0064] \quad C_R(i, j) = \frac{1}{n-1} \sum_{k=1}^n (Z_k(i) - u(i))(Z_k(j) - u(j)),$$

[0065] 展开均值, 上式可以改为:

$$[0066] \quad C_R(i, j) = \frac{1}{n-1} \left[\sum_{k=1}^n Z_k(i) Z_k(j) - \frac{1}{n} \sum_{k=1}^n Z_k(i) \sum_{k=1}^n Z_k(j) \right],$$

[0067] 计算各维特征的和 $\sum_{k=1}^n Z_k(i)_{i=1 \dots n}$, 以及任意两个特征的乘积的和 $Z(i) Z(j)_{i, j=1 \dots n}$,

就可以得到区域协方差。

[0068] 根据计算的特征协方差矩阵, 在保持测地距离下, 学习马氏距离。

[0069] 对任意两个对称正定矩阵 X_i, X_j 进行对数操作, 可以得到 $Z_i = \log(X_i), Z_j = \log(X_j)$ 。经过对数计算后, Z_i, Z_j 依旧是一个矩阵。进而, 把 Z_i, Z_j 矩阵处理成一个列向量 $\hat{X}_i, \hat{X}_j$ 。

新得到的列向量 $\hat{X}_i, \hat{X}_j$ 可以看作是欧式空间中的点, 利用参数化马氏距离形式, 可以得到:

$$[0070] \quad d_A^2(\hat{X}_i, \hat{X}_j) = (\hat{X}_i - \hat{X}_j)^T A (\hat{X}_i - \hat{X}_j)$$

[0071] 对上式矩阵 A 的学习, 可以利用 LogDet 散度来解决。即, 其中对于 $n \times n$ 矩阵 A 和 A_0 , A 和 A_0 的 LogDet 散度为:

[0072] $D_{ld}(A, A_0) = \text{tr}(AA_0^{-1}) - \log \det(AA_0^{-1}) - n$, 其中 n 为输入空间的维度。

[0073] 为了预防最小化 LogDet 散度没有可行的解决方案, 在表述形式中加入松弛变量 γ 来保证有最优解。在本发明里, 为达到保持测地距离的目的, 对称正定矩阵经变换之后的点

在欧式空间上进行马氏距离的学习之后, 得到 $d_A^2(\hat{X}_i, \hat{X}_j) - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 加上松弛变

量 γ 后, 矩阵 A 的学习问题可以表示为:

$$[0074] \quad \min_{A > 0} \quad D_{ld}(A, A_0) + \gamma \|\zeta\|_2^2,$$

$$[0075] \quad s.t. \quad d_A^2(\hat{X}_i, \hat{X}_j) \approx d_{ges}^2(X_i, X_j),$$

[0076] 其中 ζ 是变量 $\zeta_{i,j}$ 的列向量。

[0077] 最优化矩阵A的过程是一个迭代过程,步骤如下

输入: $\hat{X}$: 输入 $n^2 \times 1$ 列向量, v_{ij} : 相关对偶变量

A_0 : 输入的 $n^2 \times n^2$ 单位矩阵, γ : 松弛变量;

输出: A : 输出马氏矩阵;

1 $A \leftarrow A_0, v_{ij} \leftarrow 0 \quad \forall i, j$;

2 *repeat*;

[0078]

2.1 选择 $\hat{X}_i, \hat{X}_j \in R^{n \times n}$;

2.2 $p \leftarrow \left(\hat{X}_i - \hat{X}_j \right)^T A \left(\hat{X}_i - \hat{X}_j \right)$;

2.3 $\alpha \leftarrow \min \left(v_{ij}, \frac{1}{2} \left(\frac{1}{p} - \frac{\gamma}{\xi_{ij}} \right) \right)$;

2.4 $\beta \leftarrow \alpha / (1 - \alpha p)$;

2.5 $\xi_{ij} \leftarrow \gamma \xi_{ij} / (\gamma + \alpha \xi_{ij})$;

2.6 $v_{ij} \leftarrow v_{ij} - \alpha$;

[0079]

2.7 $A = A + \beta A \left(\hat{X}_i - \hat{X}_j \right) \left(\hat{X}_i - \hat{X}_j \right)^T A$;

3 直至收敛;

返回 $A^* = A$;

[0080] 根据得到最优的A,可以得到初始核矩阵,可以定义初始核矩阵 $K_0 = \hat{X}^T A^* \hat{X}$,有了初始核矩阵,采用Bregman迭代算法对核矩阵进行学习,使得到的最优核矩阵 K^* 逼近 K_0 ,最终得到的核矩阵满足核空间上的距离等于黎曼流形上的测地距离,从而达到测地距离保持的目的。

[0081] Bregman迭代算法一般的迭代过程为:

[0082] $J: X \rightarrow R$, 凸函数, 非负。 $u \in X$ 。

- [0083] $H: X \rightarrow R$, 凸函数, 非负。 $u \in X$, f 是已知常数 (图像数据, 是矩阵或向量)。
- [0084] X : 作用域, 是凸集也是闭集。
- [0085] 注意: 上述泛函会根据具体问题的不同具有不同的具体表达式。 J 、 H 、 X 是数据项。
- [0086] 首先, 初始化相关的参数为零;
- [0087] 然后, 再迭代公式 u 。直到 u_k 满足收敛条件。
- [0088] u 是泛函 J 的 Bregman 距离。
- [0089] p 是泛函 H 的梯度。
- [0090] Bregman 迭代算法经过多次的迭代, 就能够收敛到真实的最优解。
- [0091] 对于 $n \times n$ 矩阵 K 和 K_0 , K 和 K_0 的 LogDet 散度为:
- [0092] $D_{ld}(K, K_0) = \text{tr}(KK_0^{-1}) - \log \det(KK_0^{-1}) - n$,
- [0093] 其中 n 为输入空间的维度。
- [0094] 为了预防最小化 LogDet 散度没有可行的解决方案, 在表述形式中加入松弛变量 γ 来保证可行方案的存在。本发明的约束条件是核空间上两点间的距离等于黎曼流形上两点间的测地距离, 即令 $K_{ii} + K_{ii} - K_{ij} - d_{ges}^2(X_i, X_j) = \zeta_{ij}$ 。
- [0095] 加上松弛变量 γ 后, 表述形式为:
- [0096]
$$\min_{K > 0} D_{ld}(K, K_0) + \gamma \|\zeta\|_2^2$$
- [0097]
$$s.t. \quad K_{ii} + K_{ii} - K_{ij} \approx d_{ges}^2(X_i, X_j)$$
- [0098] 其中 ζ 是变量 ζ_{ij} 的列向量。
- [0099] 最优解 K^* 的求解过程也是一个迭代过程, 如下:

输入: K_0 : 输入的 $n^2 \times n^2$ 单位矩阵, v_{ij} : 相关对偶变量

γ : 松弛变量

输出: K : 输出核矩阵

$$1 \quad A \leftarrow K_0 = \hat{X}^T A^* \hat{X}, \quad v_{ij} \leftarrow 0 \quad \forall i, j$$

2 repeat

2.1 选择 $\hat{X}_i, \hat{X}_j \in R^{n \times n}$

$$2.2 \quad p \leftarrow \left(\hat{X}_i - \hat{X}_j \right)^T A \left(\hat{X}_i - \hat{X}_j \right)$$

$$2.3 \quad \alpha \leftarrow \min \left(v_{ij}, \frac{\gamma}{\gamma+1} \left(\frac{1}{p} - \frac{1}{\xi_{ij}} \right) \right)$$

$$2.4 \quad \beta \leftarrow \alpha / (1 - \alpha p)$$

$$2.5 \quad \xi_{ij} \leftarrow \gamma \xi_{ij} / (\gamma + \alpha \xi_{ij})$$

$$2.6 \quad \lambda_{ij} \leftarrow \lambda_{ij} - \alpha$$

$$2.7 \quad K = K + \beta K (e_i - e_j)(e_i - e_j)^T K$$

3 直至收敛

返回 $K^* = K$

[0102] 根据测度学习与核学习之间的联系 $K_0 = \hat{X}^T A^* \hat{X}$, 可以根据测度学习得到适合问题最优解 A^* , 然后直接得到 K^* 。本发明采用将得到的最优马氏矩阵带入线性变换 $K_0 = \hat{X}^T A^* \hat{X}$, 在LogDet散度下, 进行核学习, 得到最优的核矩阵, 使得核空间上的距离与黎曼流形上的测地距离保持一致。

[0103] 进一步地, 根据得到的最优核矩阵 K^* , 进行数据分类。

[0104] 数据分类说明: 采用线性分类, 即可以直接计算两点之间的相似性, 然后分类; 采用核方法分类就是把数据影射到核空间, 再分类, 这是一种非线性分类方法。在这个过程中

需要计算核矩阵(一般有直接采用核函数比如高斯核函数,或学习一个核函数或者核矩阵,我们这里属于第二种)。

[0105] 本公开的实施例提供的一种基于测地距离的黎曼流形保持核学习装置,如图2所示为本公开的一种基于测地距离的黎曼流形保持核学习装置图,该实施例的一种基于测地距离的黎曼流形保持核学习装置包括:处理器、存储器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,例如读写程序。所述处理器执行所述计算机程序时实现上述各个读写方法实施例中的步骤。

[0106] 所述装置包括:存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序运行在以下装置的单元中:

[0107] 黎曼流形计算单元,用于根据数据样本集和协方差描述子计算每个样本数据的特征协方差矩阵;

[0108] 黎曼测地距离计算单元,用于根据特征协方差矩阵计算黎曼测地距离;

[0109] 迭代核优化单元,用于采用Bregman迭代算法对黎曼测地距离定义的初始核矩阵进行学习得到最优核矩阵;

[0110] 最优核分类单元,用于根据得到的最优核矩阵进行数据分类。

[0111] 所述一种基于测地距离的黎曼流形保持核学习装置可以运行于桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备中。所述一种基于测地距离的黎曼流形保持核学习装置,可运行的装置可包括,但不仅限于,处理器、存储器。本领域技术人员可以理解,所述例子仅仅是一种基于测地距离的黎曼流形保持核学习装置的示例,并不构成对一种基于测地距离的黎曼流形保持核学习装置的限定,可以包括比例子更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述一种基于测地距离的黎曼流形保持核学习装置还可以包括输入输出设备、网络接入设备、总线等。

[0112] 所称处理器可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述一种基于测地距离的黎曼流形保持核学习装置运行装置的控制中心,利用各种接口和线路连接整个一种基于测地距离的黎曼流形保持核学习装置可运行装置的各个部分。

[0113] 所述存储器可用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述一种基于测地距离的黎曼流形保持核学习装置的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0114] 尽管本公开的描述已经相当详尽且特别对几个所述实施例进行了描述,但其并非旨在局限于任何这些细节或实施例或任何特殊实施例,而是应当将其视作是通过参考所附权利要求考虑到现有技术为这些权利要求提供广义的可能性解释,从而有效地涵盖本公开的预定范围。此外,上文以发明人可预见的实施例对本公开进行描述,其目的是为了提供有用的描述,而那些目前尚未预见的对本公开的非实质性改动仍可代表本公开的等效改动。

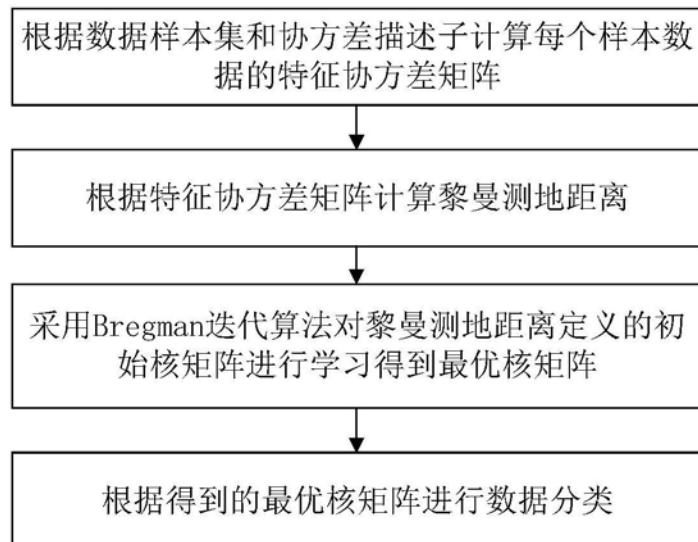


图1

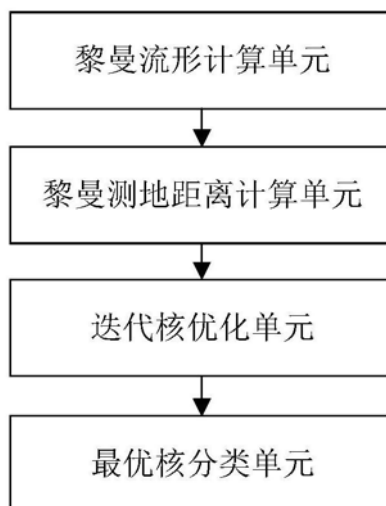


图2