第42卷 第11期 电 子 与 信 息 学 报 Vol. 42No. 11
2020年11月 Journal of Electronics & Information Technology Nov. 2020

面向图像识别的测地局部典型相关分析方法

许 欢[①] 苏树智*[①] 颜文婧[②] 邓瀛灏[①] 谢 军[①]

[①](安徽理工大学计算机科学与工程学院 淮南 232001)

[②](北京工商大学计算机与信息工程学院 北京 100037)

摘 要：典型相关分析(CCA)是一种经典的多模态特征学习方法，能够从不同模态同时学习相关性最大的低维特征，然而难以发现隐藏在样本空间中的非线性流形结构。该文提出一种基于测地流形的多模态特征学习方法，即测地局部典型相关分析(GeoLCCA)。该方法利用测地距离构建了低维相关特征的测地散布，并进一步通过最大化模态间的相关性和最小化模态内的测地散布学习更具鉴别力的非线性相关特征。该文不仅在理论上对提出的方法进行了分析，而且在真实的图像数据集上验证了方法的有效性。

关键词：图像识别；典型相关分析；多模态特征学习；流形学习

中图分类号：TN911.73; TP391.4 **文献标识码**：A **文章编号**：1009-5896(2020)11-2813-06

DOI: 10.11999/JEIT200123

A Geodesic Locality Canonical Correlation Analysis Method for Image Recognition

XU Huan[①] SU Shuzhi[①] YAN Wenjing[②] DENG Yinghao[①] XIE Jun[①]

[①](*College of Computer Science and Engineering, Anhui University of Science & Technology, Huainan* 232001, *China*)

[②](*School of Computer and Information Engineering, Beijing Technology and Business University, Beijing* 100037, *China*)

Abstract: Canonical Correlation Analysis (CCA) is a classic multi-modal feature learning method, which can learn low-dimensional features with the maximum correlation from different modalities. However, it is difficult for CCA to find the nonlinear manifold structures hidden in the sample spaces. This paper proposes a multi-modal feature learning method based on geodesic manifolds, namely Geodesic Locality Canonical Correlation Analysis (GeoLCCA).The geodesic distances are used to construct the geodesic scatters of low-dimensional correlation features, and the nonlinear correlation features with better discriminative power are learned by maximizing the between-modal correlation and minimizing the within-modal geodesic scatters. This paper not only analyzes the proposed method in theory, but also verifies the effective of the proposed method on the real-world image datasets.

Key words: Image recognition; Canonical Correlation Analysis (CCA); Multi-modal feature learning; Manifold learning

收稿日期：2020-02-21；改回日期：2020-07-23；网络出版：2020-07-23

*通信作者： 苏树智 sushuzhi@foxmail.com

基金项目：国家自然科学基金(61806006)，安徽省高等学校自然科学研究基金(KJ2018A0083)，中国博士后科学基金(2019M660149)

Foundation Items: The National Natural Science Foundation of China (61806006), The Anhui Province Natural Science Research Foundation of Institutions of Higher Learning (KJ2018A0083), The China Postdoctoral Science Foundation (2019M660149)

1 引言

随着社会信息化的不断深入，在科学研究和现实生活中都积累了大量不同类型的数据信息。每种数据能够称为一种模态，根据同一目标的数据种类数量可以将模态数据分为单模态数据和多模态数据[1]。单模态数据是指对同一目标仅有一种数据表示。目前，特征学习方法主要针对单模态数据，线形判别分析[2]、增量主成分分析[3]、弹性保持投影[4]和局部保持投影[5]是常见的传统单模态特征学习方法，并且已经广泛应用于人脸识别、情感计算、故障诊断、步态分析等领域[6–9]。除了传统单模态特征学习，深度单模态特征学习[10]通过构建多层特征学习网络，从不同的角度学习高维数据的内在鉴别特征，目前已经在目标检测中广泛使用。与单模态数据不同，多模态数据是指同一目标拥有两种或两种以上的数据表示，比如同一个人可以由指纹信息、虹膜数据和面部图像数据等共同来描述。和单模态数据相比，多模态数据能够从不同角度更好地描述同一目标的多种统计信息，并且具有互补性，因此在识别任务中多模态数据比单模态数据更具优势。目前如何从多模态数据中学习有效的特征仍然是一项挑战性的任务。

典型相关分析(Canonical Correlation Analysis, CCA)[11]是一种经典的多模态特征学习方法，旨在为每个模态学习一组相关投影方向，使不同模态投影到相关一致子空间后拥有最大的相关性，其应用已经覆盖多特征融合、过程控制、情绪压力分类、基因分析等多个领域[12–15]。类标签是一种常用的监督信息，利用监督信息可以更好对多模态数据进行分类和特征提取，从而提高目标识别的精度。鉴别多重CCA(Discriminative Multiple CCA, DMCCA)[16]是一种监督多模态特征学习方法，该方法利用类标签信息构建模态间的鉴别相关性，实现了低维特征的鉴别学习，并且在文本分类、人脸识别与手写数字识别实验上都展示出良好的效果。除了鉴别辅助以外，核技术[17]是特征学习中常用的非线性辅助技术，核CCA[18,19]借助核技术将数据映射到更高维度的核空间中，然后在核空间中执行线性的CCA方法，获得非线性低维特征，进而能够在一定程度上解决非线性相关的问题。CCA的线性特性使其难以很好地揭示原始高维数据中的局部几何结构，针对该问题Sun等人[20]提出局部保持CCA方法(Locality Preserving CCA, LPCCA)，利用欧式距离探索了原始高维数据间的局部关系，并在学习的低维特征中尽量保留非线性的局部结构，该方法在姿态估计上效果显著。借助正则化技术，图正则化CCA[21]不仅考虑图诱导的一致结构信息，而且最小化了低维特征间的距离，该方法在手写体分类等实际应用中显示了良好的实验结果。

局部几何结构是原始高维数据中隐含的一类重要信息，在CCA的基础上保持了局部流形结构，能够进一步获得更具鉴别力的低维特征。然而由于原始高维数据包含大量的噪声和冗余信息，现有的局部CCA方法中获取的局部几何结构往往严重偏离真实的局部关系，这会影响低维特征的鉴别力，为此，本文提出测地局部典型相关分析(Geodesic Locality Canonical Correlation Analysis, GeoLCCA)方法，该方法利用测地距离更好地探索高维数据间的局部关系，并构建模态内的测地散布，在多模态相关分析框架中使局部结构更加紧密，从而有效增强非线性相关特征的鉴别力。为了验证该方法的有效性，本文不仅在理论上进行论证并且在两个真实的图像数据集设计针对性实验，良好的实验结果显示了该方法在图像识别上的优越性。

2 典型相关分析

假设 $\boldsymbol{X}=[x_1,x_2,\cdots,x_N]\in R^{d_x\times n}$ 和$\boldsymbol{Y}=[y_1,y_2,\cdots,y_N]\in R^{d_y\times n}$为两个模态样本集，其中$d_x$和$d_y$分别是$\boldsymbol{X}$和$\boldsymbol{Y}$的维度，$n$为样本数量。$\{x_i,y_i\}$ $(i=1,2,\cdots,N)$对应同一目标，并且$\boldsymbol{X}$和$\boldsymbol{Y}$进行了均值化，即模态样本集的均值为零。CCA旨在通过优化相关准则求解一对投影方向$\boldsymbol{w}_x\in R^{d_x\times 1}$和$\boldsymbol{w}_y\in R^{d_x\times 1}$，使得$\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}$和$\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}$具有最大的相关性，具体的优化问题为

$$\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\ \frac{{\boldsymbol{w}_x}^{\mathrm{T}}\boldsymbol{S}_{xy}\boldsymbol{w}_y}{\sqrt{{\boldsymbol{w}_x}^{\mathrm{T}}\boldsymbol{S}_{xx}\boldsymbol{w}_x}\sqrt{{\boldsymbol{w}_y}^{\mathrm{T}}\boldsymbol{S}_{yy}\boldsymbol{w}_y}} \tag{1}$$

其中$\boldsymbol{S}_{xy}=\dfrac{1}{n}\sum\nolimits_{i=1}^{n}(\boldsymbol{x}_{\mathrm{i}}-\overline{\boldsymbol{x}})(\boldsymbol{y}_i-\overline{\boldsymbol{y}})^{\mathrm{T}}$是样本集$\boldsymbol{X}$和样本集$\boldsymbol{Y}$的协方差矩阵，$\boldsymbol{S}_{xx}=\dfrac{1}{n}\sum\nolimits_{i=1}^{n}(\boldsymbol{x}_i-\overline{\boldsymbol{x}})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^{\mathrm{T}}$和$\boldsymbol{S}_{yy}=\dfrac{1}{n}\sum\nolimits_{i=1}^{n}(\boldsymbol{y}_i-\overline{\boldsymbol{y}})(\boldsymbol{y}_i-\overline{\boldsymbol{y}})^{\mathrm{T}}$分别是$\boldsymbol{X}$和$\boldsymbol{Y}$的方差。$\boldsymbol{S}_{xy}$揭示了模态间的相关性，而$\boldsymbol{S}_{xx}$和$\boldsymbol{S}_{yy}$反映了模态内数据的总体散布信息。

为避免$\boldsymbol{w}_x$和$\boldsymbol{w}_y$的无穷等价解问题，利用尺度不变性[11]，能够将式(1)中的优化问题等价转变为

$$\left.\begin{array}{ll}\max\limits_{\boldsymbol{w}_x,\boldsymbol{w}_y} & \boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y\\ \mathrm{s.t.} & \boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x=1\,,\ \boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y=1\end{array}\right\} \tag{2}$$

3 测地局部典型相关分析(GeoLCCA)

CCA通过最大化相关特征$\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}$和$\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}$的相关性，获得每个模态的投影方向，然而忽略了隐藏在样本空间中的局部几何结构，并且目前在已有的局

部CCA方法中使用的局部几何结构往往严重偏离真实的局部关系，这将影响低维特征的鉴别力。为此，本文利用测地距离更好地探索高维数据间的局部关系，并构建模态内的测地散布，进而在多模态相关分析框架下提出了GeoLCCA方法。

首先，对CCA方法中揭示全局散布信息的$\boldsymbol{S}_{xx}$进行等价推导

$$\begin{aligned}\boldsymbol{S}_{xx}&=\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\bar{\boldsymbol{x}})^{\mathrm{T}}\\&=\frac{1}{n}\left[\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}-\sum_{i=1}^{n}\boldsymbol{x}_i\bar{\boldsymbol{x}}_i^{\mathrm{T}}-\bar{\boldsymbol{x}}_i\sum_{i=1}^{n}\boldsymbol{x}_i^{\mathrm{T}}+\bar{\boldsymbol{x}}_i\bar{\boldsymbol{x}}_i^{\mathrm{T}}\right]\\&=\frac{1}{2n^2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}+\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{x}_j\boldsymbol{x}_j^{\mathrm{T}}\right.\\&\qquad\left.-\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_j^{\mathrm{T}}-\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{x}_j\boldsymbol{x}_i^{\mathrm{T}}\right]\\&=\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\boldsymbol{x}_i-\boldsymbol{x}_j)(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\\&=\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2\end{aligned}\tag{3}$$

利用同样的推导，$\boldsymbol{S}_{yy}$能够表述为

$$\boldsymbol{S}_{yy}=\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\|\boldsymbol{y}_i-\boldsymbol{y}_j\|^2\tag{4}$$

为此，CCA的优化问题能够等价地转化为

$$\left.\begin{aligned}&\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\quad\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_i-\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_i\right\|^2\\&\text{s.t.}\quad\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_i-\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_j\right\|^2=1\\&\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_i-\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_j\right\|^2=1\end{aligned}\right\}\tag{5}$$

式(5)是基于两个样本间距离的形式，该等价形式更利于局部关系的约束。为了获得更真实的局部关系，本文借助ISOMAP[22]的思想，利用Dijkstra方法计算模态内样本间的测地距离，每个样本的前k个样本为其近邻样本，从而确定样本间的近邻关系，这种近邻关系，称之为测地近邻。在测地近邻的基础上，进一步构建测地指示矩阵$\boldsymbol{H}_X$和$\boldsymbol{H}_Y$。以样本集$\boldsymbol{X}$为例，指示矩阵$\boldsymbol{H}_X$的构建方式为

$$h_x^{ij}=\begin{cases}1, & x_j\in\mathrm{Nei}_k(x_i)\\0, & \text{其它}\end{cases}\tag{6}$$

其中h_x^{ij}是$\boldsymbol{H}_X$的第(i,j)个元素，$i,j=1,2,\cdots,n$，$\mathrm{Nei}_k(x_i)$表示x_{i}的前k个测地近邻集合。

通过测地指示矩阵，能够样本集$\boldsymbol{X}$构建测地散布ρ_X

$$\rho_X=\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}h_x^{ij}\left\|\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_i-\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_j\right\|^2\tag{7}$$

在CCA中仅仅考了模态内的整体散布，从式(1)能够看出，CCA可以视为最大化模态间的相关性，同时最小化模态内的散布信息。式(7)中的测地散布通过进一步约束测地近邻，有效地嵌入了局部关系。最小化测地散布，能够使近邻样本间更加紧密，从而使低维特征更具局部性。同样，能用样本集$\boldsymbol{Y}$构建测地散布ρ_Y

$$\rho_Y=\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}h_y^{ij}\left\|\boldsymbol{w}_{\mathrm{y}}^{\mathrm{T}}\boldsymbol{y}_i-\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_j\right\|^2\tag{8}$$

CCA相关准则是整体上最大化了模态间的相关性，同时减弱了模态内的散布情况，但是CCA不能真实地反应样本空间中的局部信息，为此，GeoLCCA在最大化模态间相关性的同时，最小化了模态内测地散布，借助测地散布嵌入了更真实的局部几何结构。

因此，GeoLCCA的优化问题可以描述为

$$\left.\begin{aligned}&\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\quad\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_i-\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_i\right\|^2\\&\text{s.t.}\quad\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}h_x^{ij}\left\|\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_i-\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{x}_j\right\|^2=1\\&\frac{1}{2n^2}\sum_{i=1}^{\mathrm{n}}\sum_{j=1}^{n}h_y^{ij}\left\|\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_i-\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{y}_j\right\|^2=1\end{aligned}\right\}\tag{9}$$

$\boldsymbol{X}$和$\boldsymbol{Y}$为均值归一化矩阵，并且目标函数和约束中的常数不影响式(9)优化求解，因此式(9)中的常数能够去掉，进一步对上述优化问题进行推导，式(9)能够等价地转化为

$$\left.\begin{aligned}&\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\quad\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y\\&\text{s.t.}\quad\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{H}_{\mathrm{X}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x=1\,,\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{H}_{\mathrm{Y}}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y=1\end{aligned}\right\}\tag{10}$$

为了对式(10)进行优化求解，首先构建式(10)的拉格朗日乘子函数$L(\boldsymbol{w}_x,\boldsymbol{w}_y)$

$$\begin{aligned}L(\boldsymbol{w}_x,\boldsymbol{w}_y)=&\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y-\frac{\lambda}{2}\left(\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{H}_X\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x-1\right)\\&-\frac{\theta}{2}(\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{H}_Y\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y-1)\end{aligned}\tag{11}$$

其中λ和θ为拉格朗日乘子，分别对$\boldsymbol{w}_x$和$\boldsymbol{w}_y$求偏导，并将其设为零，可得

$$\frac{\partial L}{\partial\boldsymbol{w}_x}=\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y-\lambda\boldsymbol{X}\boldsymbol{H}_X\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x=0\tag{12}$$

$$\frac{\partial L}{\partial \boldsymbol{w}_y}=\boldsymbol{Y}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x-\theta\boldsymbol{Y}\boldsymbol{H}_Y\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y=0 \tag{13}$$

式(12)和式(13)可以等价地转化为式(14)的多元特征值问题

$$\begin{bmatrix} & \boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}} \\ \boldsymbol{Y}\boldsymbol{X}^{\mathrm{T}} & \end{bmatrix}\begin{bmatrix}\boldsymbol{w}_x\\ \boldsymbol{w}_y\end{bmatrix}=\begin{bmatrix}\lambda\boldsymbol{X}\boldsymbol{H}_X\boldsymbol{X}^{\mathrm{T}} & \\ & \theta\boldsymbol{Y}\boldsymbol{H}_Y\boldsymbol{Y}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{w}_x\\ \boldsymbol{w}_y\end{bmatrix} \tag{14}$$

多元特征值问题无法获得解析解，为此，对式(12)和式(13)分别左乘以$\boldsymbol{w}_x^{\mathrm{T}}$和$\boldsymbol{w}_y^{\mathrm{T}}$，可得

$$\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y-\lambda\boldsymbol{w}_x^{\mathrm{T}}\boldsymbol{X}\boldsymbol{H}_X\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x=0 \tag{15}$$

$$\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{w}_x-\theta\boldsymbol{w}_y^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{H}_Y\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{w}_y=0 \tag{16}$$

由式(15)和式(16)能够推导出$\lambda=\theta$，因此，式(14)中的多元特征值问题能够等价地转化为式(17)的广义特征值问题

$$\begin{bmatrix} & \boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}} \\ \boldsymbol{Y}\boldsymbol{X}^{\mathrm{T}} & \end{bmatrix}\begin{bmatrix}\boldsymbol{w}_x\\ \boldsymbol{w}_y\end{bmatrix}=\lambda\begin{bmatrix}\boldsymbol{X}\boldsymbol{H}_X\boldsymbol{X}^{\mathrm{T}} & \\ & \boldsymbol{Y}\boldsymbol{H}_Y\boldsymbol{Y}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{w}_x\\ \boldsymbol{w}_y\end{bmatrix} \tag{17}$$

上述广义特征值问题能够快速地获得解析解，其中λ是特征值。通过求解式(17)，能够获得前d个最大特征值对应的特征向量$\{(\boldsymbol{w}_{xi}^{\mathrm{T}},\boldsymbol{w}_{yi}^{\mathrm{T}})\}_{i=1}^{d}$，进而可以构建模态样本集$\boldsymbol{X}$和$\boldsymbol{Y}$对应的投影矩阵$\boldsymbol{W}_x=[\boldsymbol{w}_{x1},\boldsymbol{w}_{x2},\cdots,\boldsymbol{w}_{xd}]^{\mathrm{T}}\in R^{\mathrm{d}x\times d}$和$\boldsymbol{W}_y=[\boldsymbol{w}_{y1},\boldsymbol{w}_{y2},\cdots,\boldsymbol{w}_{yd}]^{\mathrm{T}}\in R^{\mathrm{d}y\times d}$。通过$\boldsymbol{W}_x$和$\boldsymbol{W}_y$的投影，能够快速获得$X$和$\boldsymbol{Y}$的低维相关特征$\widetilde{X}=\boldsymbol{W}_x^{\mathrm{T}}X\in R^{d\times n}$和$\widetilde{\boldsymbol{Y}}=\boldsymbol{W}_y^{\mathrm{T}}\boldsymbol{Y}\in R^{d\times n}$。

4 实验的结果和分析

为了验证提出方法的有效性，在GT图像数据集[23]和ORL图像数据集[24]上设计了一些实验来评估本文所提出方法的图像识别性能。通过模态策略[25]获得每个图像不同的模态数据。具体而言，利用Coiflets和Daubechies正交小波变换能够获得每幅图像的两幅低频子图像，并且为了减轻小样本问题，采用主成分分析将子图像的维数约减到100维，进而获得每幅图像的两个模态数据。在该实验部分，GeoLCCA方法与图多视角典型相关分析(Graph Multiview Canonical Correlation Analysis, GMCCA)[12]，LPCCA, DMCCA, CCA进行对比分析，并且对于所有方法的最终识别率均是利用基于欧式距离的最近邻分类器来获得。

4.1 在GT图像数据集上的实验

GT图像数据集采集共50人的图像，该数据集包含每个人15幅背景杂乱的图像。在实验部分，分别从每类中随机选择q (q=5, 6, 7, 8)样本作为训练样本，其余为测试样本，独立运行10次样本随机实验，在表1中展示了所有方法的平均识别率。

CCA仅仅利用了全局的结构信息，忽略了局部几何结构，无法获取强鉴别力的非线性特征，在表1中也表现出了最差的识别性能。GeoLCCA, GMCCA, LPCCA均在一定程度上考虑了局部关系，并且在学习的低维特征中尽量保留局部关系。原始高维数据包含大量的噪声和冗余信息，直接基于欧式距离的局部关系难以很好地反映真实的局部几何结构，LPCCA在低维特征学习中嵌入了这种失真的局部关系，这是在一定程度上影响了非线性相关特征的鉴别力，这是LPCCA的识别率低于GeoLCCA的重要原因。不同于LPCCA, GMCCA借助图正则化技术从不同的角度探索局部信息的嵌入，DMCCA在相关分析框架下约束原始高维数据的鉴别结构，GMCCA和DMCCA对局部失真的鲁棒性更强，在表1中拥有更高的识别率。GeoLCCA利用测地距离更好地探索高维数据间的测地局部，并借助模态内的测地散布有效增强了相关特征的局部紧密性，从而学习了更具鉴别力的低维相关特征，在表1中，GeoLCCA也展现了最好的识别率。

4.2 在ORL图像数据集上的实验

ORL图像数据集从40个人采集了400幅，这些图像具有不同的面部表情、照明条件和遮挡等。图像数据通常是超高维的非线性数据，CCA属于线性方法，仅考虑了全局的结构信息，难以从非线性数据中学习强鉴别力的低维相关特征。尽管LP-CCA利用原始高维数据间的局部信息直接约束模态间的相关性和模态内的散布结构，但是由于原始高维的原始中包含大量的噪声和冗余信息，原始高

表 1 在GT图像数据集上的识别率(%)及标准差

	训练样本数5	训练样本数6	训练样本数7	训练样本数8
GeoLCCA	67.26±2.01	71.36±1.83	76.10±1.28	78.20±1.31
GMCCA	65.22±1.64	66.64±1.56	69.70±1.75	72.06±1.66
LPCCA	44.84±1.73	50.09±3.79	54.15±1.74	57.46±2.56
DMCCA	63.56±2.77	67.80±1.29	73.67±1.71	75.80±1.99
CCA	59.08±1.81	61.78±1.35	66.22±1.66	68.14±2.01

$A\pm B$: A表示平均识别率(%)，B表示对应的识别率标准差

维数据间的局部关系经常严重失真，这种失真的局部关系导致在图像识别中LPCCA拥有比CCA更低的识别率。GMCCA同样利用了原始高维数据间的局部信息，但是GMCCA利用局部信息约束了融合后的低维相关特征，并且利用图正则化技术进一步增强了失真局部信息的鲁棒性。DMCCA对原始高维数据的鉴别结构进行约束，从不同的角度解决失真局部信息影响类分离性的问题。在表2中，GMCCA和DMCCA同样也展现出比LPCCA和CCA更高的识别率。GeoLCCA使用测地距离更好地探索原始高维数据间的测地局部，获得更加真实的局部关系，有效弱化了上述方法中的局部失真问题，并且通过约束模态内的测地散布，有效增强了低维特征的局部紧密性，进而学习了更具鉴别力的非线性相关特征，这是GeoLCCA在表2中拥有最高识别率的重要原因。此外，在大多数情况下，GeoLCCA的识别率标准差比其他方法更小，这表明GeoLCCA对于样本随机的鲁棒性更好。在GT和ORL图像数据集上的实验结果可以给出一个合理的观察，GeoLCCA是一个有效的图像识别方法。

表 2　在ORL图像数据集上的识别率(%)及标准差

	训练样本数5	训练样本数6	训练样本数7	训练样本数8
GeoLCCA	95.15±1.58	97.19±1.33	98.25±0.83	99.50±0.65
GMCCA	93.90±2.04	95.19±0.89	97.00±1.53	98.50±1.42
LPCCA	84.70±3.00	87.81±2.40	89.17±2.00	94.25±2.58
DMCCA	93.80±1.53	95.50±1.74	96.75±1.49	99.38±0.66
CCA	90.35±1.94	93.19±1.94	93.83±1.68	97.25±1.15

$A\pm B$: A表示平均识别率(%)，B表示对应的识别率标准差

5　结论

CCA作为一种经典的多模态特征学习方法，能够同时从多模态数据集中学习低维相关特征，但是CCA是一种线性方法，难以发现隐藏在样本空间中的非线性流行结构，为此，本文提出一种GeoLCCA方法，该方法利用测地距离更好地探索了高维数据间的测地局部，并以此为基础，在多模态相关分析框架中构建了每个模态的测地散布，使嵌入的局部几何结构更加紧密，从而有效增强了非线性相关特征的鉴别力。在两个真实的图像数据集上，良好的实验结果显示了GeoLCCA在图像识别任务中的有效性。

参 考 文 献

[1] 刘政怡, 段群涛, 石松, 等. 基于多模态特征融合监督的RGB-D图像显著性检测[J]. 电子与信息学报, 2020, 42(4): 997–1004. doi: 10.11999/JEIT190297.
LIU Zhengyi, DUAN Quntao, SHI Song, *et al*. RGB-D image saliency detection based on multi-modal feature-fused supervision[J]. *Journal of Electronics & Information Technology*, 2020, 42(4): 997–1004. doi: 10.11999/JEIT190297.

[2] YE Qiaolin, FU Liyong, ZHANG Zhao, *et al*. Lp- and Ls-norm distance based robust linear discriminant analysis[J]. *Neural Networks*, 2018, 105: 393–404. doi: 10.1016/j.neunet.2018.05.020.

[3] 王肖锋, 孙明月, 葛为民. 基于图像协方差无关的增量特征提取方法研究[J]. 电子与信息学报, 2019, 41(11): 2768–2776. doi: 10.11999/JEIT181138.
WANG Xiaofeng, SUN Mingyue, and GE Weimin. An incremental feature extraction method without estimating image covariance matrix[J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2768–2776. doi: 10.11999/JEIT181138.

[4] YUAN Sen and MAO Xia. Exponential elastic preserving projections for facial expression recognition[J]. *Neurocomputing*, 2018, 275: 711–724. doi: 10.1016/j.neucom.2017.08.067.

[5] WANG Rong, NIE Feiping, HONG Richang, *et al*. Fast and orthogonal locality preserving projections for dimensionality reduction[J]. *IEEE Transactions on Image Processing*, 2017, 26(10): 5019–5030. doi: 10.1109/TIP.2017.2726188.

[6] ZHU Yani, ZHU Chaoyang, and LI Xiaoxin. Improved principal component analysis and linear regression classification for face recognition[J]. *Signal Processing*, 2018, 145: 175–182. doi: 10.1016/j.sigpro.2017.11.018.

[7] KUMAR S, BHUYAN M K, LOVELL B C, *et al*. Hierarchical uncorrelated multiview discriminant locality preserving projection for multiview facial expression recognition[J]. *Journal of Visual Communication and Image Representation*, 2018, 54: 171–181. doi: 10.1016/j.jvcir.2018.04.013.

[8] GAJJAR S, KULAHCI M, and PALAZOGLU A. Real-time fault detection and diagnosis using sparse principal component analysis[J]. *Journal of Process Control*, 2018, 67: 112–128. doi: 10.1016/j.jprocont.2017.03.005.

[9] WANG Hao, FAN Yuanyuan, FANG Baofu, *et al*. Generalized linear discriminant analysis based on Euclidean

norm for gait recognition[J]. *International Journal of Machine Learning and Cybernetics*, 2018, 9(4): 569–576. doi: 10.1007/s13042-016-0540-0.

[10] 董书琴, 张斌. 基于深度特征学习的网络流量异常检测方法[J]. 电子与信息学报, 2020, 42(3): 695–703. doi: 10.11999/JEIT190266.
DONG Shuqin and ZHANG Bin. Network traffic anomaly detection method based on deep features learning[J]. *Journal of Electronics & Information Technology*, 2020, 42(3): 695–703. doi: 10.11999/JEIT190266.

[11] SUN Quansen, ZENG Shenggen, LIU Yan, *et al.* A new method of feature fusion and its application in image recognition[J]. *Pattern Recognition*, 2005, 38(12): 2437–2448. doi: 10.1016/j.patcog.2004.12.013.

[12] CHEN Jia, WANG Gang, and GIANNAKIS G B. Graph multiview canonical correlation analysis[J]. *IEEE Transactions on Signal Processing*, 2019, 67(11): 2826–2838. doi: 10.1109/TSP.2019.2910475.

[13] LIU Yiqi, LIU Bin, ZHAO Xiujie, *et al.* A mixture of variational canonical correlation analysis for nonlinear and quality-relevant process monitoring[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(8): 6478–6486. doi: 10.1109/TIE.2017.2786253.

[14] HONG Kan, LIU Guodong, CHEN Wentao, *et al.* Classification of the emotional stress and physical stress using signal magnification and canonical correlation analysis[J]. *Pattern Recognition*, 2018, 77: 140–149. doi: 10.1016/j.patcog.2017.12.013.

[15] SAFO S E, AHN J, JEON Y, *et al.* Sparse generalized eigenvalue problem with application to canonical correlation analysis for integrative analysis of methylation and gene expression data[J]. *Biometrics*, 2018, 74(4): 1362–1371. doi: 10.1111/biom.12886.

[16] GAO Lei, QI Lin, CHEN Enqing, *et al.* Discriminative multiple canonical correlation analysis for information fusion[J]. *IEEE Transactions on Image Processing*, 2018, 27(4): 1951–1965. doi: 10.1109/TIP.2017.2765820.

[17] GENG Fazhan and QIAN Suping. An optimal reproducing kernel method for linear nonlocal boundary value problems[J]. *Applied Mathematics Letters*, 2018, 77: 49–56. doi: 10.1016/j.aml.2017.10.002.

[18] MELZER T, REITER M, and BISCHOF H. Appearance models based on kernel canonical correlation analysis[J]. *Pattern Recognition*, 2003, 36(9): 1961–1971. doi: 10.1016/s0031-3203(03)00058-x.

[19] ALAM M A, FUKUMIZU K, and WANG Yuping. Influence function and robust variant of kernel canonical correlation analysis[J]. *Neurocomputing*, 2018, 304: 12–29. doi: 10.1016/j.neucom.2018.04.008.

[20] SUN Tingkai and CHEN Songcan. Locality preserving CCA with applications to data visualization and pose estimation[J]. *Image and Vision Computing*, 2007, 25(5): 531–543. doi: 10.1016/j.imavis.2006.04.014.

[21] CHEN Jia, WANG Gang, SHEN Yanning, *et al.* Canonical correlation analysis of datasets with a common source graph[J]. *IEEE Transactions on Signal Processing*, 2018, 66(16): 4398–4408. doi: 10.1109/TSP.2018.2853130.

[22] BALASUBRAMANIAN M, SCHWARTZ E L, TENENBAUM J B, *et al.* The Isomap algorithm and topological stability[J]. *Science*, 2002, 295(5552): 7. doi: 10.1126/science.295.5552.7a.

[23] ZHANG Guiying, ZOU Wenbin, ZHANG Xianjie, *et al.* Singular value decomposition based virtual representation for face recognition[J]. *Multimedia Tools and Applications*, 2018, 77(6): 7171–7186. doi: 10.1007/s11042-017-4627-8.

[24] SU Shuzhi, GE Hongwei, YUAN Yunhao, *et al.* A label embedding kernel method for multi-view canonical correlation analysis[J]. *Multimedia Tools and Applications*, 2017, 76(12): 13785–13803. doi: 10.1007/s11042-016-3786-3.

[25] SU Shuzhi, FANG Xianjin, YANG Gaoming, *et al.* Self-balanced multi-view orthogonality correlation analysis for image feature learning[J]. *Infrared Physics & Technology*, 2019, 100: 44–51. doi: 10.1016/j.infrared.2019.05.008.

许　欢：女，1982年生，助教，研究方向为机器学习、图像处理、模式识别.

苏树智：男，1987年生，副教授，研究方向为多模态模式识别、特征学习、子空间融合、图像处理.

颜文婧：女，1984年生，讲师，研究方向为机器学习、模式识别、信号处理.

责任编辑：余　蓉