



## 基于图嵌入的自适应多视降维方法

尹宝才, 张超辉, 胡永利, 孙艳丰, 王博岳

引用本文:

尹宝才, 张超辉, 胡永利, 等. 基于图嵌入的自适应多视降维方法[J]. 智能系统学报, 2021, 16(5): 963–970.

YIN Baocai, ZHANG Chaohui, HU Yongli, et al. An adaptive multi-view dimensionality reduction method based on graph embedding[J].

CAAI Transactions on Intelligent Systems, 2021, 16(5): 963–970.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202105021>

## 您可能感兴趣的其他文章

### 一种深度自监督聚类集成算法

A deep self-supervised clustering ensemble algorithm

智能系统学报. 2020, 15(6): 1113–1120 <https://dx.doi.org/10.11992/tis.202006050>

### 面向自闭症辅助诊断的无监督模糊特征学习新方法

A novel unsupervised fuzzy feature learning method for computer-aided diagnosis of autism

智能系统学报. 2019, 14(5): 882–888 <https://dx.doi.org/10.11992/tis.201808005>

### 多标记学习自编码网络无监督维数约简

Unsupervised dimensionality reduction of multi-label learning via autoencoder networks

智能系统学报. 2018, 13(5): 808–817 <https://dx.doi.org/10.11992/tis.201804051>

### 多层递阶融合模糊特征映射的模糊C均值聚类算法

Fuzzy C-means clustering algorithm for multilayered hierarchical fusion fuzzy feature mapping

智能系统学报. 2018, 13(4): 594–601 <https://dx.doi.org/10.11992/tis.201703047>

### 目标再确认中的优化扩散距离相似性度量

Target re-identification based on optimized diffusion distance

智能系统学报. 2018, 13(2): 269–280 <https://dx.doi.org/10.11992/tis.201607010>

### 分段聚合近似和数值导数的动态时间弯曲方法

Dynamic time warping based on piecewise aggregate approximation and data derivatives

智能系统学报. 2016, 11(2): 249–256 <https://dx.doi.org/10.11992/tis.201507064>

微信公众平台



关注微信公众号, 获取更多资讯信息

## 吴文俊人工智能科技进步奖一等奖

成果名称：多源高维数据协同表征及应用

获 奖 人：尹宝才、胡永利、刘雪莉、魏运、张可、刘浩、张勇、姚磊、孙艳丰、孙永良、施云惠、孔德慧、田青、王少帆、李敬华

完成单位：北京工业大学、青岛海信网络科技股份有限公司、北京城建设计发展集团股份有限公司、北京市运输管理技术支持中心、北京市交通信息中心、安徽博微广成信息科技有限公司、北方工业大学



### 尹宝才

北京工业大学教授，博士生导师，北京人工智能研究院院长，多媒体与智能软件技术北京市重点实验室主任，中国计算机学会人工智能与模式识别专业委员会委员，全国信息技术标准化技术委员会多媒体分委会委员。ACM北京分会副主席。

长期从事多媒体技术、视频编码、模式识别和智能交通等领域的研究工作，在中国手语合成、3D人脸建模、视频高效编码、高维数据表达、跨媒体智能和交通大数据处理等方面做出创造性贡献。主持包括国基金重大项目课题、国基金重点项目、973计划课题、北京市自然科学基金重点项目在内的10余项国家和省部级科研项目。研究成果获国家科技进步二等奖，高等学校科技进步一等奖、中国标准创新贡献奖和信息产业重大技术发明奖。参与国家音视频编码标准（AVS）研究工作，2项技术提案被AVS标准采纳。在IEEE TPAMI、IEEE TIP、IEEE TNNLS、IEEE TMM、IEEE TCSVT、IEEE TCYB等领域重要刊物和CVPR、ICCV、AAAI、IJCAI、ACM MM、InfoCom、SIGGRAPH等顶级会议上发表论文300余篇。获得国家发明专利60余项。

## 团队简介

项目团队由北京工业大学和北方工业大学两所工科高校，北京市交通信息中心，北京市运输管理技术支持中心等交通行业技术部门，以及青岛海信网络科技股份有限公司、北京城建设计发展集团股份有限公司、安徽博微广成信息科技有限公司等智能交通企业组成，通过各个单位的长期合作，形成了优势互补的产学研用创新团队。在国家973 计划和国家自然科学基金重大等项目的支持下，团队合作研发的高维数据高效表达理论，多源高维数据协同表征方法，以及交通图像视频智能分析技术和客流交通流感知等创新技术，通过技术合作和专利转让等方式在企业产品和工程项目实现技术转化，并应用于北京、青岛、长沙、武汉、合肥和西安等10多个城市的智能交通和安全监控系统。团队合作研发的大规模交通视频互联集成和协同处理技术，解决了交通视频大数据有效利用的行业难题，形成了国家标准和行业规范，相关技术成果支撑建成了接入规模达4万多路的北京交通视频应用平台。团队研发的接续运输系统和客流监测系统在首都机场、北京地铁、苏州地铁每年为数亿旅客提供交通出行服务。团队技术成果两年新增销售额5.1 亿元，新增经济效益 3.2 亿元，年均节省运营成本和能耗支出1.4亿元，取得了良好的经济和社会效益。

DOI: 10.11992/tis.202105021

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20210713.1310.003.html>

## 基于图嵌入的自适应多视降维方法

尹宝才<sup>1,2</sup>, 张超辉<sup>1</sup>, 胡永利<sup>1,2</sup>, 孙艳丰<sup>1,2</sup>, 王博岳<sup>1,2</sup>

(1. 北京工业大学信息学部, 北京 100124; 2. 北京人工智能研究院, 北京 100124)

**摘要:** 随着监控摄像头的普及和数据采集技术的快速发展, 多视数据呈现出规模大、维度高和多源异构的特点, 使得数据存储空间大、传输慢、算法复杂度高, 造成“有数据、难利用”的困境。到目前为止, 国内外在多视降维方面的研究还比较少。针对这一问题, 本文提出一种基于图嵌入的自适应多视降维方法。该方法在考虑视角内降维后数据重构原始高维数据的基础上, 提出自适应学习相似矩阵来探索不同视角之间降维后数据的关联关系, 学习各视数据的正交投影矩阵实现多视降维任务。本文在多个数据集上对降维后的多视数据进行聚类/识别实验验证, 实验结果表明基于图嵌入的自适应多视降维方法优于其他降维方法。

**关键词:** 降维; 多视数据; 图嵌入; 自适应学习; 高维数据; 相似性度量; 无监督学习; 表示学习

**中图分类号:** TP18 **文献标志码:** A **文章编号:** 1673-4785(2021)05-0963-08

中文引用格式: 尹宝才, 张超辉, 胡永利, 等. 基于图嵌入的自适应多视降维方法[J]. 智能系统学报, 2021, 16(5): 963-970.

英文引用格式: YIN Baocai, ZHANG Chaohui, HU Yongli, et al. An adaptive multi-view dimensionality reduction method based on graph embedding[J]. CAAI transactions on intelligent systems, 2021, 16(5): 963-970.

## An adaptive multi-view dimensionality reduction method based on graph embedding

YIN Baocai<sup>1,2</sup>, ZHANG Chaohui<sup>1</sup>, HU Yongli<sup>1,2</sup>, SUN Yanfeng<sup>1,2</sup>, WANG Boyue<sup>1,2</sup>

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. Beijing Artificial Intelligence Institute, Beijing 100124, China)

**Abstract:** With the popularity of surveillance cameras and the rapid development of data acquisition technology, multi-view data shows the traits of large scale, high dimension and multi-source heterogeneity, which cause large data storage, low data transmission speed and high algorithm complexity, resulting in a predicament that “there are plenty of data that are hard to use”. Up to now, few domestic and foreign researches have been done on multi-view dimensionality reduction. In order to solve this problem, this paper proposes an adaptive multi-view dimensionality reduction method based on graph embedding. In consideration of the reconstructed high-dimensional data after the view-angle dimensionality reduction, this method puts forward an adaptive similarity matrix to explore the correlation between dimension-reduced data from different perspectives and learn the orthogonal projection matrix of each perspective to achieve the multi-view dimensionality reduction task. In this paper, a clustering/recognition verification experiment is performed on the dimension-reduced multi-view data from multiple data sets. The experimental results present that the proposed method is better than other dimensionality reduction methods.

**Keywords:** dimensionality reduction; multi-view data; graph embedding; adaptive learning; high-dimensional data; similarity measure; unsupervised learning; representation learning

收稿日期: 2021-05-13. 网络出版日期: 2021-07-14.

基金项目: 国家自然科学基金项目(U19B2039, 61906011); 北京市自然科学基金项目(4204086).

通信作者: 胡永利. E-mail: huyongli@bjut.edu.cn.

随着摄像头、传感器以及互联网技术的快速发展, 人们能够越来越便捷地获取各种各样的多视数据, 例如描述人的脸、笔迹、指纹等特征, 描述图像的颜色、纹理、形状等特征。多视数据

能够更全面的描述目标对象,有益于克服单视数据中常见的光照、遮挡等难题,在识别<sup>[1]</sup>、聚类<sup>[2]</sup>等多视学习任务中表现出显著优势。然而,多视数据一般维度较高<sup>[3]</sup>,对其进行向量化表示时容易引发“维数灾难”问题,给传统多视学习算法带来了极大的困难。降维是数据挖掘的基础和热点问题,在克服“维数灾难”的同时还能挖掘数据中的判别信息,但其很难处理多视数据。因此,如何对多视数据进行有效的降维成为一个亟待解决的问题。

多视降维方法主要分为两类:有监督多视降维方法<sup>[4]</sup>和无监督多视降维方法<sup>[5]</sup>。有监督数据降维通过标签数据学习从高维数据到低维数据的降维模型,并利用该降维模型对新高维数据进行降维。然而,互联网上大部分数据是没有标签的,对其进行人工标注不仅效率低下且需要付出高昂的成本,因此无监督多视降维方法的研究变得更为重要。无监督多视降维方法不使用数据的标签信息,挖掘数据间联系来学习数据的低维表示。截止到目前,针对多视数据的无监督降维方法的研究,国内外都刚刚起步。

典范相关分析 (canonical correlation analysis, CCA) 及其变体<sup>[6-7]</sup>是研究多视图数据之间线性相关的一种有效方法。具体地,判别典型相关分析 (discriminative CCA, DCCA)<sup>[8]</sup>结合了 CCA 和线性判别分析 (linear discriminative analysis, LDA)<sup>[9]</sup>,从不同的角度考虑了同一类别数据之间的相关性,以提高低维子空间的判别能力。张量典型相关分析 (tensor CCA, TCCA)<sup>[10]</sup>通过分析不同视图之间的协方差张量,将 CCA 推广到处理多视数据,但是该方法计算成本过高限制了其应用范围。偏最小二乘 (partial least squares, PLS)<sup>[11]</sup>使用回归方式将不同视的数据映射到公共线性子空间。分布式频谱嵌入 (distributed spectral embedding, DSE)<sup>[12]</sup>将数据平滑地嵌入到低维空间中。基于多核学习 (multiple kernel learning, MKL)<sup>[13]</sup>的方法将多核学习与图嵌入结合实现降维任务。多视联合降维 (multi-view dimensionality co-reduction, MDcR)<sup>[14]</sup>使用核匹配方法挖掘多个视图之间的依存关系,从而获得每视数据的低维投影。耦合块对齐算法 (coupled patch alignment, CPA)<sup>[15]</sup>将样本和其跨视角下的同类近邻以及异类近邻组成局部块,平衡跨视角类内的紧密性与类间的可分离性,并扩展成多维耦合块对齐,解决了任意数量视角的共同学习问题。

综上,现有的多视降维方法存在以下主要缺

点: 1) 这些方法仅关注不同视数据之间的关系,而忽略同视数据内部的关系; 2) 这些方法通常仅考虑降维后的数据关系,而忽略原始数据的关联关系; 3) 这些方法通常将不同视数据映射到一个公共空间,当不同视原始数据的维度不同时降维结果受影响较大。针对这些问题,本文提出的基于图嵌入的自适应多视降维 (MVDR-GE) 方法在考虑每视内数据关联性的基础上,自适应地学习相似矩阵来探索不同视之间数据的关联关系,获得各视数据的正交投影矩阵实现多视降维任务。

## 1 相关工作

### 1.1 图嵌入降维

图嵌入降维是使降维后数据保留原始数据局部结构关系的一种降维方法。对给定的训练样本数据  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbf{R}^{V \times N}$ , 其中  $V$  代表数据的维数,  $N$  是样本个数,  $\mathbf{X}$  的低维表示为  $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N] \in \mathbf{R}^{K \times N}$ , 其中  $K \ll V$ ,  $\mathbf{z}_i$  为  $\mathbf{x}_i$  的低维表示。如果两个数据点  $\mathbf{x}_i$  和  $\mathbf{x}_j$  很相似, 那么它们的低维表示  $\mathbf{z}_i$  和  $\mathbf{z}_j$  也应该很相似。定义描述低维数据  $\mathbf{z}_i$  和  $\mathbf{z}_j$  之间相似度关系的值为  $s_{ij}$ , 其构成的相似度矩阵为  $\mathbf{S}$ , 其中  $s_{ij} = -\|\mathbf{x}_i - \mathbf{x}_j\|_F^2$ 。根据相似度矩阵  $\mathbf{S}$  定义对角度矩阵  $\mathbf{D}$ , 其对角线上的元素为相似度矩阵每行元素之和, 即:

$$d_{ii} = \sum_{j=1}^N s_{ij} \quad (1)$$

定义归一化的拉普拉斯矩阵  $\mathbf{L} = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} \mathbf{D}^{-\frac{1}{2}}$ 。则图嵌入降维模型可以表示为<sup>[16]</sup>

$$\min_{\mathbf{P} \in \mathbf{R}^{K \times V}} \|\mathbf{P} \mathbf{x}_i - \mathbf{P} \mathbf{x}_j\|^2 s_{ij} \quad (2)$$

通过式 (2) 可求得高维数据  $\mathbf{X}$  的投影矩阵  $\mathbf{P} \in K \times V$ 。

令  $\mathbf{Z} = \mathbf{P} \mathbf{X}$ , 式 (2) 可以写成图正则化表示:

$$\max_{\mathbf{Z} \in \mathbf{R}^{K \times N}} \text{tr}(\mathbf{Z} \mathbf{L} \mathbf{Z}^T) \quad (3)$$

### 1.2 自适应局部结构学习

自适应局部结构学习旨在通过度量样本之间的欧氏距离来自适应地学习样本间的相似度, 从而获得样本的局部结构信息。对于每个样本  $\mathbf{x}_i$  和样本  $\mathbf{x}_j$ ,  $s_{ij}$  为衡量两个样本之间相似性的值。则自适应学习相似性矩阵  $\mathbf{S}$  可通过下面优化问题求得<sup>[17]</sup>:

$$\min_{\mathbf{S}} \sum_{i,j=1}^N \|\mathbf{x}_i - \mathbf{x}_j\|_F^2 s_{ij} \quad (4)$$

尽管  $s_{ij}$  可以自适应地学习两个样本的相似度, 但是式 (4) 不能避免  $\mathbf{S}$  的对角线为 1 这种平凡



解,故引入正则化约束:

$$\min_S \sum_{i,j=1}^N \|\mathbf{x}_i - \mathbf{x}_j\|_F^2 s_{ij} + \|\mathbf{S}\|_F^2 \quad (5)$$

式(5)是面向单视数据的自适应局部结构学习的公式,将其扩展到多视数据后可得:

$$\min_S \sum_{m=1}^M \sum_{i,j=1}^N \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2 s_{ij} + \|\mathbf{S}\|_F^2 \quad (6)$$

## 2 基于图嵌入的自适应多视降维

对于有  $M$  视的多视数据  $\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(M)}\}$ , 每视的样本个数均为  $N$ , 第  $m$  视的数据集  $\mathbf{X}^{(m)}$ , 其维度为  $V^{(m)}$ 。多视图嵌入降维就是学习各视的投影矩阵  $\mathbf{P}^{(m)} \in \mathbf{R}^{K^{(m)} \times V^{(m)}}$ , 将各视原始数据  $\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(M)}\}$  降维到  $\{\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(M)}\}$ , 降维后的维度为  $K^{(m)}$ 。

基于模型式(3)的思路,通过多视共享相似性矩阵可以有效地探索视与视数据间的关系,则多视图嵌入降维模型为

$$\max_{\mathbf{P}^{(m)}, \mathbf{L}} \sum_{m=1}^M \text{tr}(\mathbf{P}^{(m)} \mathbf{X}^{(m)} \mathbf{L} \mathbf{X}^{(m)\top} \mathbf{P}^{(m)\top}) \quad (7)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

式中  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$  的目的是避免平凡解。

对式(7)进行分解,则:

$$\min_{\mathbf{P}^{(m)}, \mathbf{S}} \sum_{m=1}^M \sum_{i,j=1}^N \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 s_{ij} \quad (8)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

式中  $\mathbf{P}^{(m)} \in \mathbf{R}^{K^{(m)} \times V^{(m)}}$ 。为避免相似度矩阵  $\mathbf{S}$  出现平凡解,同样引入对于相似度矩阵  $\mathbf{S}$  的正则化约束,则式(8)变为

$$\min_{\mathbf{P}^{(m)}, \mathbf{S}} \sum_{m=1}^M \sum_{i,j=1}^N \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 s_{ij} + \beta \|\mathbf{S}\|_F^2 \quad (9)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

式(9)中的相似度矩阵  $\mathbf{S}$  只关注降维后数据关联关系,忽略了原始数据间的结构关系,从而不能保证降维后的数据保留原始数据的结构。因此,引入对原始数据的约束后,则:

$$\min_{\mathbf{P}^{(m)}, \mathbf{S}} \sum_{m=1}^M \sum_{i,j=1}^N (\lambda_1 \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 + \lambda_2 \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2) s_{ij} + \lambda_3 \|\mathbf{S}\|_F^2 \quad (10)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

式中  $\lambda_1, \lambda_2$  和  $\lambda_3$  是平衡各个模块重要性的参数。此外,为了避免降维后数据丢失信息严重,引入约束视内数据重构误差项:

$$\min_{\mathbf{P}^{(m)}} \sum_{m=1}^M \sum_{i=1}^N \|\mathbf{x}_i^{(m)} - \mathbf{P}^{(m)\top} \mathbf{P}^{(m)} \mathbf{x}_i^{(m)}\|_F^2 \quad (11)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

式中:  $\mathbf{P}^{(m)} \mathbf{x}_i^{(m)}$  表示把原始数据投影到低维空间;  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} \mathbf{x}_i^{(m)}$  表示将投影到低维空间的数据再投影回原始空间。将式(10)与式(11)结合得到最终的目标函数:

$$\min_{\mathbf{P}^{(m)}, \mathbf{S}} \sum_{m=1}^M \sum_{i,j=1}^N \|\mathbf{x}_i^{(m)} - \mathbf{P}^{(m)\top} \mathbf{P}^{(m)} \mathbf{x}_i^{(m)}\|_F^2 + \sum_{i,j=1}^N (\lambda_1 \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 + \lambda_2 \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2) s_{ij} + \lambda_3 \|\mathbf{S}\|_F^2 \quad (12)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

## 3 算法优化

目标函数式(12)是典型非凸优化问题,本文通过迭代更新的方法,求解  $\mathbf{P}^{(m)}$  和  $\mathbf{S}$ 。

### 3.1 固定 $\mathbf{S}$ 更新 $\mathbf{P}^{(m)}$

当固定相似度矩阵  $\mathbf{S}$  不变时,目标函数式(12)中的  $\|\mathbf{S}\|_F^2$  和  $(\|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2) s_{ij}$  固定不变,对剩余项构建拉格朗日函数后得:

$$(\mathbf{X}^{(m)} \mathbf{X}^{(m)\top} + \lambda_1 \mathbf{X}^{(m)} \mathbf{L} \mathbf{X}^{(m)\top}) \mathbf{P}^{(m)} = \mathbf{P}^{(m)} \Lambda \quad (13)$$

通过特征值分解得方法,可以求得投影矩阵  $\mathbf{P}^{(m)}$  并可指定其目标维度。

### 3.2 固定 $\mathbf{P}^{(m)}$ 更新 $\mathbf{S}$

当固定投影矩阵  $\mathbf{P}^{(m)}$  时,重构误差项不参与更新,目标函数式(12)变为

$$\min_S \sum_{m=1}^M \sum_{i,j=1}^N (\lambda_1 \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 + \lambda_2 \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2) s_{ij} + \lambda_3 \|\mathbf{S}\|_F^2 \quad (14)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

将式(14)中的低维数据和原始数据之间的距离设为  $d_{ij}$ :

$$d_{ij} = \sum_{m=1}^M \lambda_1 \|\mathbf{P}^{(m)} \mathbf{x}_i^{(m)} - \mathbf{P}^{(m)} \mathbf{x}_j^{(m)}\|_F^2 + \lambda_2 \|\mathbf{x}_i^{(m)} - \mathbf{x}_j^{(m)}\|_F^2 \quad (15)$$

则式(14)可简化为

$$\min_S \sum_{m=1}^M \left( \sum_{i,j=1}^N d_{ij} s_{ij} + \lambda_3 \|\mathbf{S}\|_F^2 \right) \quad (16)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

将式(16)变换为最小二乘形式:

$$\min_S \sum_{m=1}^M \sum_{i,j=1}^N \left\| \mathbf{s}_i + \frac{\mathbf{d}_i}{2\lambda_3} \right\|_2^2 \quad (17)$$

s.t.  $\mathbf{P}^{(m)} \mathbf{P}^{(m)\top} = \mathbf{I}$

从而获得闭合解  $\mathbf{s}_i = (-\frac{\mathbf{d}_i}{2\beta})$ 。

综上所述,通过不断的迭代更新,最终得到各视的投影矩阵  $\mathbf{P}^{(m)}$ , 实现多视降维任务。

该模型的算法如下:

**算法1** 基于图嵌入的自适应多视降维方法 (MVDR-GE)。

**输入** 多视数据  $\{X^{(1)}, X^{(2)}, \dots, X^{(M)}\}$ ; 每视维度降至  $\{K^{(1)}, K^{(2)}, \dots, K^{(M)}\}$ ; 参数  $\lambda_1, \lambda_2$  和  $\lambda_3$ 。

1) 初始化  $S = \sum_{m=1}^M X^{(m)T} X^{(m)}$ ;

2) 当不满足收敛条件时:

For  $m=1:M$

根据式(13)计算各式投影矩阵  $P^{(m)}$ ;

End for

根据式(17)计算相似度矩阵  $S$ ;

End

**输出** 各视降维矩阵  $\{P^{(1)}, P^{(2)}, \dots, P^{(M)}\}$ 。

## 4 实验及结果分析

为了验证本文所提出的多视降维方法的有效性,在多个公开数据集上进行了验证实验。

### 4.1 数据集描述

IXMAS<sup>[18]</sup>数据集是由5个不同视角拍摄的视频组成的多视视频数据集,包含10个人的11种日常行为:看手表、抱胳膊、抓头、坐下、起来、转身、走、挥手、拳击、踢、捡。本文为每个视角提取177维视频特征。

ORL数据集由40个人、每人10幅不同的人脸照片构成。每个人拍摄照片的时间不同,光照和面部表情不同,本文调整图像大小为  $64 \times 64$ ,并提取图像的光照强度(intensity)、局部二值模式(local binary pattern, LBP)<sup>[19]</sup>和Gabor<sup>[19]</sup>特征构造多视数据,维度分别是4 096、3 304和6 750。

PIE数据集包含从68个人中采集的共1 428张人脸图像,该数据集提取特征的方法跟ORL数据集相同。

MSRCV1<sup>[21]</sup>数据集包含240张图片,这些图片属于9个类别,本文实验选取其中的树、建筑物、飞机、牛、人脸、汽车和自行车等7个类别共210张图片。分别提取每张图片的CENT(1 302维)、COLOR(48维)、GIST(512维)、LBP(256维)4种特征作为此数据集的多视数据。

Notting-Hill<sup>[22]</sup>诺丁山数据集来源于电影《诺丁山》,收集了5个角色的多张人脸照片。本文调整图像大小为  $64 \times 64$ 并提取跟ORL数据集相同的特征,维度分别为2 000、3 304和6 750。

### 4.2 实验及结果

为了验证方法的有效性,与以下几个经典降维方法做了实验结果对比。

主成分分析(principal component analysis, PCA)<sup>[23]</sup>:该方法是经典的无监督线性降维方法,通过将原始数据投影到线性子空间中对数据进行降维。

朴素多视降维(naive multi-view dimension reduction, NaMDR)<sup>[14]</sup>:此方法独立地减小了每个视图的维度,而没有约束不同的视图。

多视联合降维(multi-view dimensionality co-reduction, MDcR)<sup>[14]</sup>:该方法利用核匹配方法挖掘多个视图之间的依存关系,并获得每视数据的低维投影。

对于不同的数据集,特征的维数也不相同,本文统一将数据的目标特征维度设置为10维。为了进一步验证多视学习的有效性,本文对比了不同单视数据的最好性能和多视联合的性能,其中不同单视数据的最好性能在表格中用single表示,多视联合的性能在表格中用multiple表示。具体地说,在得到降维数据后,本文进行了聚类/识别实验从而间接对降维性能进行了评估。全部实验是在Window 10、64位操作系统、Inter Corei5-6500 3.20 GHz和20 G RAM的工作站上实现的,软件环境为Matlab 2018a。

#### 4.2.1 聚类实验

在聚类实验中,采用3个评价指标度量聚类性能:准确性(accuracy, ACC)、归一化互信息(normalized mutual information, NMI)和纯度(Purity)。这些指标越高代表聚类表现越好。在获得不同降维方法的降维数据后,本文使用自动加权多图(auto-weighted multi-view learning, AMGL)<sup>[23]</sup>聚类方法对其进行聚类。由于此方法最后一步是执行K-means操作,所以本文进行30次实验取均值和标准差作为最终的实验结果。最后的实验结果为平均值±标准差的形式。本文在IXMAS、PIE、Notting-Hill 3个数据集上进行了聚类实验,实验结果如表1~3所示。

本文用加粗的字体表示最好的实验结果。从表1~3中可以看出,本文提出的多视降维方法无论在单视特征还是多视特征上都取得了最好的聚类表现。其他降维方法在ACC、NMI和Purity指标上,都不能取得一个令人满意的结果。由于本文提出的多视降维方法利用一个融合多视数据相似度的矩阵去辅助生成投影矩阵,使得用投影矩阵和原始数据运算得到的每视的低维数据能够融合其他视数据的信息,并且本文不同视数据的相似度矩阵  $S$  是自适应学习得到的,所以在拥有很

强的灵活性的同时又能保证多视数据相互融合之后的低维数据具有代表高维数据的判别性信息, 最终对多视低维数据进行聚类时获得一个令人满

意的表现。同时, 对比单视低维数据聚类也能获得最好的结果, 表明本文提出的多视降维方法已经很好地融合了多视数据。

表 1 IXMAS 数据集聚类实验结果

Table 1 Clustering experiment results of IXMAS dataset

特征	方法	ACC	NMI	Purity
single	PCA <sup>[23]</sup>	0.354 3±0.021 7	0.399 2±0.027 7	0.383 6±0.021 4
	NaMDR <sup>[14]</sup>	0.351 5±0.022 3	0.390 1±0.021 4	0.384 7±0.022 5
	McDR <sup>[14]</sup>	0.386 5±0.017 1	0.407 3±0.019 0	0.404 5±0.015 6
	MVDR-GE	<b>0.413 5±0.022 4</b>	<b>0.412 0±0.022 9</b>	<b>0.423 4±0.016 3</b>
multiple	PCA <sup>[23]</sup>	0.365 3±0.020 3	0.432 2±0.019 0	0.379 1±0.019 4
	NaMDR <sup>[14]</sup>	0.347 5±0.018 3	0.431 4±0.024 6	0.379 4±0.015 2
	McDR <sup>[14]</sup>	0.387 8±0.018 2	0.462 0±0.019 8	0.411 9±0.015 9
	MVDR-GE	<b>0.397 7±0.023 1</b>	<b>0.484 5±0.023 4</b>	<b>0.433 4±0.018 7</b>

表 2 PIE 数据集聚类实验结果

Table 2 Clustering experiment results of PIE dataset

特征	方法	ACC	NMI	Purity
single	PCA <sup>[23]</sup>	0.362 1±0.011 4	0.621 7±0.012 7	0.431 1±0.007 5
	NaMDR <sup>[14]</sup>	0.288 6±0.011 0	0.595 9±0.011 6	0.343 2±0.008 4
	McDR <sup>[14]</sup>	0.290 2±0.007 5	0.584 7±0.010 1	0.344 8±0.008 0
	MVDR-GE	<b>0.546 6±0.021 3</b>	<b>0.735 1±0.013 8</b>	<b>0.602 1±0.017 1</b>
multiple	PCA <sup>[23]</sup>	0.363 6±0.011 2	0.624 8±0.009 5	0.432 4±0.008 7
	NaMDR <sup>[14]</sup>	0.354 9±0.011 2	0.621 6±0.011 1	0.419 0±0.009 4
	McDR <sup>[14]</sup>	0.324 3±0.011 1	0.597 2±0.015 2	0.389 3±0.015 7
	MVDR-GE	<b>0.637 7±0.018 4</b>	<b>0.801 6±0.016 8</b>	<b>0.706 9±0.013 9</b>

表 3 Notting-Hill 数据集聚类实验结果

Table 3 Clustering experiment results of Notting-Hill dataset

特征	方法	ACC	NMI	Purity
single	PCA <sup>[23]</sup>	0.612 2±0.069 9	0.521 9±0.076 9	0.633 5±0.067 4
	NaMDR <sup>[14]</sup>	0.601 7±0.058 8	0.508 7±0.063 8	0.625 3±0.055 7
	McDR <sup>[14]</sup>	0.663 6±0.041 7	0.587 4±0.042 6	0.697 9±0.021 9
	MVDR-GE	<b>0.693 7±0.010 4</b>	<b>0.641 2±0.010 9</b>	<b>0.718 6±0.005 4</b>
multiple	PCA <sup>[23]</sup>	0.717 5±0.114 9	0.683 0±0.136 4	0.751 2±0.107 5
	NaMDR <sup>[14]</sup>	0.790 0±0.112 8	0.742 6±0.115 9	0.805 4±0.103 0
	McDR <sup>[14]</sup>	0.841 5±0.057 7	0.792 6±0.058 4	0.849 6±0.051 1
	MVDR-GE	<b>0.849 3±0.052 0</b>	<b>0.827 5±0.043 3</b>	<b>0.873 9±0.038 7</b>

#### 4.2.2 识别实验

本文在 ORL、Notting-Hill 和 MSRCV1 数据集上进行了识别实验, 采用最近邻分类器作为识别方法, 采用识别准确度作为评价指标。由于在 Notting-Hill 和 MSRCV1 数据集上每类的样本不相同, 本文从每类所有的样本中选取了  $M$  个样本进行训练, 记作  $G_m$ , 其余的样本进行测试。本文

实验设置为随机选取 30 次训练集进行训练和测试, 并报告平均识别结果。最后的实验结果为平均值±标准差的形式。

实验结果如表 4~6 所示, 本文提出的方法在绝大多数情况下获得了最好的识别效果。由于本文提出的多视降维方法能够很好地融合多视数据的特征, 使得降维后的低维数据能充分表示高维

数据且具有更强的判别性,因此本文的方法能在识别实验中得到一个比较高的准确度,这也验证了本文的方法相较于其他对比方法,能够更好地对不同视角数据进行融合,使降维后的低维数据更有利于进行识别。在 ORL 人脸数据集和 Notting-Hill 数据集上,本文提出的方法虽然在单视

数据和多视数据上均获得最好的实验结果,单视实验结果整体略优于多视实验结果,其他方法也同样出现了该问题。这一现象的原因在于两个数据集提取的 LBP 特征能够很好地反应图像的特征,在拼接多视数据进行识别时,本文选用的识别方法不能很好地融合多视数据进行识别。

表 4 ORL 数据集识别实验结果

Table 4 Recognition experiment results of ORL dataset

特征	方法	$G_m=60$	$G_m=50$	$G_m=40$	$G_m=30$
single	PCA <sup>[22]</sup>	0.845±0.024 7	0.8173±0.027 4	0.783 6±0.027 8	0.743 4±0.031 1
	NaMDR <sup>[14]</sup>	0.822 0±0.021 1	0.797 1±0.025 5	0.762 5±0.027 2	0.725 3±0.024 5
	McDR <sup>[14]</sup>	0.839 5±0.027 4	0.812 6±0.024 3	0.788 1±0.023 1	0.735 3±0.027 7
	MVDR-GE	<b>0.845 1±0.027 0</b>	<b>0.822 5±0.188 8</b>	<b>0.794 7±0.023 6</b>	<b>0.750 9±0.018 6</b>
multiple	PCA <sup>[22]</sup>	0.858 3±0.025 9	0.805 6±0.031 4	0.775 0±0.038 64	0.712 6±0.033 6
	NaMDR <sup>[14]</sup>	0.853 1±0.025 9	0.819 8±0.030 5	0.779 1±0.034 8	0.705 9±0.035 8
	McDR <sup>[14]</sup>	0.850 8±0.027 0	0.820 0±0.026 2	0.775 8±0.037 3	0.715 ±0.039 7
	MVDR-GE	<b>0.861 0±0.029 7</b>	<b>0.824 5±0.032 3</b>	<b>0.786 8±0.024 9</b>	<b>0.727 6±0.032 3</b>

表 5 Notting-Hill 数据集识别实验结果

Table 5 Recognition experiment results of Notting-Hill dataset

特征	方法	$G_m=66$	$G_m=55$	$G_m=44$	$G_m=33$
single	PCA <sup>[22]</sup>	0.986 3±0.007 5	0.980 7±0.007 8	0.976 1±0.011 4	0.973 1±0.008 6
	NaMDR <sup>[14]</sup>	0.985 6±0.007 9	0.981 8±0.007 5	0.978 4±0.007 9	0.970 9±0.012 1
	McDR <sup>[14]</sup>	0.993 4±0.004 0	0.993 2±0.007 0	0.988 2±0.007 0	0.982 1±0.010 8
	MVDR-GE	<b>0.995 1±0.005 8</b>	<b>0.993 6±0.004 9</b>	<b>0.990 4±0.007 6</b>	<b>0.985 4±0.008 0</b>
multiple	PCA <sup>[22]</sup>	0.989 6±0.005 6	0.989 8±0.008 5	0.987 8±0.007 5	0.975 2±0.013 5
	NaMDR <sup>[14]</sup>	0.992 5±0.007 6	0.988 0±0.008 9	0.982 1±0.010 4	0.975 9±0.010 2
	McDR <sup>[14]</sup>	0.993 7±0.006 9	0.990 9±0.005 8	0.986 0±0.004 7	0.977 7±0.010 4
	MVDR-GE	<b>0.994 0±0.004 5</b>	<b>0.993 7±0.005 5</b>	<b>0.988 4±0.009 1</b>	<b>0.981 1±0.010 0</b>

表 6 MSRCV1 数据集识别实验结果

Table 6 Recognition experiment results of MSRCV1 dataset

特征	方法	$G_m=18$	$G_m=15$	$G_m=12$	$G_m=9$
single	PCA <sup>[22]</sup>	0.818 2±0.030 6	0.805 0±0.035 7	0.793 6±0.026 9	0.782 0±0.022 1
	NaMDR <sup>[14]</sup>	0.822 6±0.042 1	0.810 4±0.030 1	0.804 4±0.027 2	0.790 0±0.029 3
	McDR <sup>[14]</sup>	<b>0.862 6±0.033 1</b>	<b>0.854 2±0.021 6</b>	<b>0.844 7±0.031 9</b>	<b>0.833 2±0.024 5</b>
	MVDR-GE	0.834 9±0.033 1	0.822 8±0.028 0	0.810 0±0.030 1	0.793 8±0.033 9
multiple	PCA <sup>[22]</sup>	0.893 2±0.023 4	0.870 1±0.026 8	0.861 1±0.031 6	0.843 0±0.025 0
	NaMDR <sup>[14]</sup>	0.885 7±0.029 0	0.871 1±0.029 4	0.862 9±0.025 1	0.851 4±0.026 7
	McDR <sup>[14]</sup>	0.894 0±0.035 0	0.879 6±0.023 7	0.866 1±0.027 9	0.849 2±0.027 5
	MVDR-GE	<b>0.898 4±0.031 2</b>	<b>0.885 8±0.028 4</b>	<b>0.871 7±0.026 5</b>	<b>0.856 2±0.021 6</b>

## 5 结束语

本文提出了一种基于图嵌入的自适应多视降维方法。通过图嵌入的方法能够将高维数据嵌入到一个低维的子空间,并使得到的低维数据能够

具有高维数据的特点。通过自适应学习一个多视数据共享的相似性矩阵,能够对不同视的数据在降维的过程中进行融合,使得降维后的每个视之间的数据都包含其它视特征的信息。实验结果



证明,本文提出的方法在多视数据降维的同时能够促进多视数据的融合,并且能够提高后续聚类/识别任务的实验效果。虽然本文提出的方法获得了令人满意的多视降维表现,但降维过程比较耗时,今后的研究将致力于如何降低算法的复杂度。

## 参考文献:

- [1] SHARIF M, MOHSIN S, JAVED M Y. A survey: face recognition techniques[J]. Research journal of applied sciences, engineering and technology, 2012, 4(23): 4979–4990.
- [2] CHALLA A, DANDA S, SAGAR B S D, et al. Power spectral clustering[J]. Journal of mathematical imaging and vision, 2020, 62(9): 1195–1213.
- [3] DONOHO D L. High-dimensional data analysis: the curses and blessings of dimensionality[J]. AMS math challenges lecture, 2000, 1: 32.
- [4] YAN Yan, RICCI E, SUBRAMANIAN R, et al. Multitask linear discriminant analysis for view invariant action recognition[J]. IEEE transactions on image processing, 2014, 23(12): 5599–5611.
- [5] SUN Yaoqi, LI Liang, ZHENG Liang, et al. Image classification base on PCA of multi-view deep representation [J]. Journal of visual communication and image representation, 2019, 62: 253–258.
- [6] BLASCHKO M B, LAMPERT C H. Correlational spectral clustering[C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA, 2008: 1–8.
- [7] ZHANG Yanyan, ZHANG Jianchun, Pan Zhisong, et al. Multi-view dimensionality reduction via canonical random correlation analysis[J]. Frontiers of computer science, 2016, 10(5): 856–869.
- [8] SUN Tingkai, CHEN Songcan, Yang Jingyu, et al. A novel method of combined feature extraction for recognition[C]//2008 Eighth IEEE International Conference on Data Mining. Pisa, Italy, 2008: 1043–1048.
- [9] 杨健, 杨静宇, 叶晖. Fisher 线性鉴别分析的理论研究及其应用 [J]. 自动化学报, 2003, 29(4): 481–493.  
YANG Jian, YANG Jingyu, YE Hui. Theory of Fisher linear discriminant analysis and its application[J]. Acta automatica sinica, 2003, 29(4): 481–493.
- [10] LUO Yong, TAO Dacheng, RAMAMOHANARAO K, et al. Tensor canonical correlation analysis for multi-view dimension reduction[J]. IEEE transactions on knowledge and data engineering, 2015, 27(11): 3111–3124.
- [11] SHARMA A, JACOBS D W. Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch[C]//CVPR 2011. Colorado Springs, USA, 2011: 593–600.
- [12] XIA Tian, TAO Dacheng, MEI Tao, et al. Multiview spectral embedding[J]. IEEE transactions on systems, man, and cybernetics, part B (cybernetics), 2010, 40(6): 1438–1446.
- [13] LIN Y Y, LIU T L, FUH C S. Multiple kernel learning for dimensionality reduction[J]. IEEE transactions on pattern analysis and machine intelligence, 2011, 33(6): 1147–1160.
- [14] ZHANG Changqing, FU Huazhu, HU Qinghua, et al. Flexible multi-view dimensionality co-reduction[J]. IEEE transactions on image processing, 2017, 26(2): 648–659.
- [15] BEN X, GONG C, ZHANG P, et al. Coupled patch alignment for matching cross-view gait[J]. IEEE transactions on image processing, 2019, 28(6): 3142–3157.
- [16] NIE Feiping, CAI Guohao, LI Xuelong. Multi-view clustering and semi-supervised classification with adaptive neighbours[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. San Francisco, USA, 2017: 2408–2414.
- [17] WEINLAND D, RONFARD R, BOYER E. Free view-point action recognition using motion history volumes[J]. Computer vision and image understanding, 2006, 104(2/3): 249–257.
- [18] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE transactions on pattern analysis and machine intelligence, 2002, 24(7): 971–987.
- [19] LADES M, VORBRUGGEN J C, BUHMANN J, et al. Distortion invariant object recognition in the dynamic link architecture[J]. IEEE transactions on computers, 1993, 42(3): 300–311.
- [20] WINN J, JOJIC N. Locus: learning object classes with unsupervised segmentation[C]//Tenth IEEE International Conference on Computer Vision. Beijing, China, 2005: 756–763.
- [21] WU Baoyuan, ZHANG Yifan, HU Baogang, et al. Constrained clustering and its application to face clustering in videos[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 3507–3514.
- [22] ABDI H, WILLIAMS L J. Principal component analysis[J]. WIREs: computational statistics, 2010, 2(4): 433–459.
- [23] NIE Feiping, CAI Guohao, LI Jing, et al. Auto-weighted multi-view learning for image clustering and semi-supervised classification[J]. IEEE transactions on image processing, 2018, 27(3): 1501–1511.

- [24] YAN Shuicheng, XU Dong, ZHANG Benyu, et al. Graph embedding and extensions: a general framework for dimensionality reduction[J]. IEEE transactions on pattern analysis and machine intelligence, 2007, 29(1): 40–51.

#### 作者简介:



尹宝才, 教授, 博士生导师, 多媒体与智能软件技术北京市重点实验室主任, 北京人工智能研究院院长, 中国计算机学会人工智能与模式识别专业委员会委员, ACM 北京分会副主席。主要研究方向为多媒体技术、跨媒体智能、视频编码。主持国家 973 项目、国家自然科学基金重大项目、国家自然科学基金重点项目、北京市自然科学基金重点项目等 10 余项。发表学术论文 60 余篇。



张超辉, 硕士研究生, 主要研究方向为机器学习、数据挖掘。



胡永利, 教授, 博士生导师, 主要研究方向为模式识别、计算机视觉、跨媒体智能和智能交通。北京市高等学校创新团队-交通大数据处理学术创新团队。主持国家自然科学基金联合基金重点项目、面上项目等 10 余项。发表学术论文 30 余篇。

## 中国人工智能学会推荐的 35 场会议入选中国 科协《重要学术会议指南(2021)》

### The 35 conferences recommended by Chinese Association for Artificial Intelligence were selected into 《2021 Top Academic Conference》 of the China Association for Science and Technology

《重要学术会议指南(2021)》(以下简称《指南》)正式上线发布, 同时, “重要学术会议指南”(Top Academic Conference, TAC)标识正式启用, 入选的 707 场会议成为首批授权使用该标识的学术会议。此次, 由中国人工智能学会推荐的 35 场会议入选《指南》, 其中, 由中国人工智能学会主办的会议有 10 场。

#### 2021 第六年全球人工智能技术大会

举办时间: 2021 年 6 月 5—6 日

举办地点: 浙江杭州

会议官网: <https://gaitc.caii.cn/>

#### 2021 第七届中国人工智能大会

举办时间: 2021 年 8 月 7—8 日

举办地点: 四川成都

会议官网: <https://ccai.caii.cn/>

#### 2021 第十一届中国智能产业高峰论坛

举办时间: 2021 年 11 月 6—7 日

举办地点: 福建厦门

#### 2021 首届国际人工智能会议

举办时间: 2021 年 6 月 6 日

举办地点: 浙江杭州

会议官网: <https://cicai.caii.cn/#/>

#### 第七届 IEEE 云计算与智能系统国际会议

举办时间: 2021 年 11 月 6—7 日

举办地点: 陕西西安

会议官网: <http://2021ccis.csp.escience.cn/dct/page/1>

#### 中国粒计算与知识发现学术会议

举办时间: 2021 年 8 月 20—22 日

举办地点: 江西南

昌会议官网: <http://cgckd2021.ecjtu.edu.cn/>

#### 大数据智能高峰论坛

举办时间: 2021 年 10 月—11 月

举办地点: 重庆

会议官网: <http://bdaiid.cqupt.edu.cn/#/introduction>

#### 亚洲人工智能技术大会

举办时间: 2021 年 11 月 19—21 日

举办地点: 海南海口

会议官网: <http://www.acait.cn/>

#### 第十八届中国机器学习会议

举办时间: 2021 年 8 月 6—8 日

举办地点: 湖南长沙

会议官网: <https://2021ccml.scimeeting.cn/>

#### 第六届认知系统和信息处理国际会议

举办时间: 2021 年 8 月 20—23 日

举办地点: 江苏苏州