SCIENTIA SINICA Informationis



单样本学习下时序约束稀疏表示的物体识别方法

童小宝^{1,2}, 熊鹏文^{2,3*}, 宋爱国⁴, 刘小平^{2,5}

1. 中南大学自动化学院, 长沙 410083, 中国

2. 南昌大学先进制造学院, 南昌 330031, 中国

3. 中国科学技术大学自动化系, 合肥 230026, 中国

4. 东南大学仪器科学与工程学院,南京 210096,中国

5. Department of Systems and Computer Engineering, Carleton University, Ottawa K1S5B6, Canada

* 通信作者. E-mail: steven.xpw@ncu.edu.cn

收稿日期: 2022-06-24; 修回日期: 2022-12-07; 接受日期: 2023-01-16; 网络出版日期: 2024-01-11

国家自然科学基金 (批准号: 62373181, 62163024, 61903175, 61663027)、江西省 "双千计划" (批准号: jxsq2023201097)、江西省 杰出青年基金 (批准号: 20232ACB212002) 和江西省主要学科学术与技术带头人项目 (批准号: 20204BCJ23006) 资助

摘要 非接触式传感器相比接触式触觉传感器可以避免与物体直接接触过程中产生的噪声,因而能够获取更有价值的原始数据表征物体内在属性;然而针对非接触式传感器感知的物体属性数据而言,现有算法难以实现单样本学习下的物体准确识别.为解决这一问题,本文提出一种新颖的单样本学习下时序约束稀疏表示方法 (one-shot learning with temporally constrained sparse representation, OSL-TCSR)用于识别 5 种不同材料下的 50 个物体类别.首先将两种原始数据 (Lumini 光谱和 SCiO 光谱)并行投影至共享子空间,并且使用聚类典型关联分析法 (cluster canonical correlation analysis, C-CCA) 计算两种原始数据的聚类相关性特征;其次通过稀疏表示分别计算得到聚类相关性特征数据以及原始数据的编码向量,并利用原始数据的编码向量对相关性特征数据的编码向量进行二次投影映射;然后将两次映射后的原始数据和相关性特征数据进行重构,以充分耦合化两种光谱数据,解决了单样本学习下的数据稀缺问题;进一步地,设计新颖的时序约束稀疏表示方法计算重构后的原始数据和相关性特征数据,以充分考虑每个光谱序列的时序特征;最后与最新的物体识别方法进行实验对比,结果表明提出的 OSL-TCSR 方法提高了单样本学习情况下的物体识别结果.此外,OSL-TCSR 还可灵活迁移至多种应用场景,比如材料识别或纹理识别等.

关键词 物体识别, 聚类相关性特征, 单样本学习, 时序约束正则化

1 引言

机器人通过触觉传感器获取触觉信号,从而感知到交互对象的多个特性,如纹理^[1,2]、材料^[3]、物体^[4,5]、运动模式^[6]等. GelSight 传感器^[7]利用感知层触觉图像形变的形状与程度信息进行滑动识

引用格式: 童小宝, 熊鹏文, 宋爱国, 等. 单样本学习下时序约束稀疏表示的物体识别方法. 中国科学: 信息科学, 2024, 54: 191-209, doi: 10.1360/SSI-2022-0256
 Tong X B, Xiong P W, Song A G, et al. Object recognition based on one-shot learning with temporally constrained sparse representation (in Chinese). Sci Sin Inform, 2024, 54: 191-209, doi: 10.1360/SSI-2022-0256

© 2024《中国科学》杂志社



图 1 两种光谱数据收集过程 Figure 1 Collection process of two types of spectral data

别. Strese 等^[8,9] 设计 Texplorer 传感器以获取材料的多模态触觉信息 (力、声音、加速度等), 从而进行材料识别. Takahashi 等^[10] 使用 uSkin 传感器同时获得物体的三维力触觉信号 (*X*, *Y* 轴方向摩擦力与 *Z* 轴方向压力), 以识别未知物体. 但上述研究工作主要是利用接触式触觉传感器与交互对象接触而产生触觉信号的, 这无疑在摩擦接触过程中引入冗余噪声, 从而导致传感器获取的数据质量较差.

而与传统接触式测量相比,非接触式测量方式^[3]获取的多光谱数据具有高表征能力、低噪声等特点,对不同物体形状具有较强的适应性,故更适用于物体识别研究.与接触式触觉感知类似,光谱测量也能够感知交互对象的物体和材料等特性,它包含两种光谱序列(Lumini和 SCiO).图 1 为两种光 谱序列收集过程,其中 PR2 机器人左边搭载 SCiO 传感器,右边搭载 Lumini 传感器.两种光谱的波 长不同,Lumini 设定的序列波长范围为 λ = 317 ~ 856 nm,而 SCiO 波长设定的序列波长范围为 λ = 740 ~ 1070 nm.由于光谱数据的高表征能力与低噪声等优点,故适合用于机器人单样本学习研究.而 本文研究单样本学习是源于以下几点问题:(1)如何从光谱数据的单个样本中有效地学习字典特征是 机器人物体识别研究中的一个重要问题;(2)如何在单样本学习情况下充分融合多光谱信息,以计算获 得较高的物体识别准确率;(3)单样本学习相对小样本学习而言,在数据缺乏极端情况下样本收集代价 更小,且算法计算时间更短,更具有应用前景.

为了解决上述问题,本文充分探索两种光谱数据间的耦合关联性,对两种维度的光谱数据进行相关性分析与重构;同时又因光谱数据具有时序特征,即相邻波段的数据存在一定联系,本文进一步探索不同光谱数据各自子数据间的时序约束性,通过约束相邻子编码向量间的时序差异以考虑每个光谱数据的子数据间的连续性.经过两个阶段的探索,使得本文提出的单样本学习下时序约束稀疏表示方法 (one-shot learning with temporally constrained sparse representation, OSL-TCSR) 在单个光谱样本训练下 (即单样本学习)即可得到较高的物体识别结果.综上所述,本文基于非接触式光谱测量设计了一种 OSL-TCSR 算法用于物体识别,其具体价值和贡献概括如下.

(1) 本文首先将两种原始光谱数据并行投影 (数据投影) 至共享子空间, 从而得到相关性特征数据, 然后使用二次投影 (编码向量投影) 再次映射原始数据和相关性特征数据的关系, 最后重构了充分耦

合后的多种光谱数据.通过上述方式,本文提出的 OSL-TCSR 可以充分考虑两种光谱测量的耦合联系,并解决单样本学习下的数据稀缺问题.

(2) 传统稀疏表示方法将序列数据作为整体导入,从而忽略了序列数据相邻子数据间的时序连续性.为解决这一问题,本文首次设计新颖的时序约束稀疏表示方法用以分析重构后的光谱序列,以充分考虑到每个序列的时序特征.

(3) 本文提出的 OSL-TCSR 只需随机选取 1 个训练样本即可达到较高的物体识别结果,从而减 少了算法运行时间和人工数据标记成本.为进一步减少 OSL-TCSR 在实验中的迭代次数,本文还设 计了一种新颖的加速近端梯度下降法 (accelerated proximal gradient descent, APGD) 用于分析提出的 OSL-TCSR,并采用线性搜索方法计算近端梯度方法中的动态步长.

最后,将本文提出的 OSL-TCSR 与最新的物体识别方法进行实验比较,结果表明 OSL-TCSR 提高了单样本学习情况下的物体识别准确率.

2 相关工作

Erickson 等^[11] 使用 CNN 算法对光谱序列和高清光谱图像进行融合以实现材料有效分类. Liu 等^[12] 提出一种基于多手指触觉融合的联合核稀疏编码方法,该方法通过探索每个手指触觉数据的不同编码特征,并利用 L2,1 范数对多个手指的编码向量进行联合稀疏惩罚,使得多个手指触觉数据共享相同的稀疏模式,以保证多个手指触觉数据的空间结构关联,从而有效地对物体进行识别. Roberge 等^[13] 提出基于动态触觉序列的无监督分类方法,该方法通过提取每个动态触觉序列的编码向量为特征集,然后利用支持向量机对编码向量进行分类. Chen 等^[14] 利用多种不同的核函数计算数据间的高维核距离,然后使用 K 核奇异值分解计算特征字典矩阵,并使用核正交匹配追踪计算数据的编码向量,以实现物体的有效感知. Liu 等^[15] 提出一种语义正则化字典学习方法,该方法通过将触觉形容词转化为标签矢量,并融合到监督字典学习中,从而识别不同材料. Tong 等^[16] 提出一种自适应多核字典学习算法,该算法可以在机器人多指抓取物体时同时考虑多个手指间的力关联和单个手指的多个触觉传感器间的力关联. 通过探索这两种力关联信息,该算法有效提高了多指抓取状态识别表现.

然而,上述算法总是通过大量数据训练而进行物体识别研究,这激励学者们开始着重研究单样本 学习、零样本学习等算法. Xiong 等^[2]提出一种基于稀疏表示的单样本学习纹理识别方法,该方法首 先利用字典学习得到多个不同模态(视觉、声音和加速度)的公共编码向量,然后设计类间正则化项, 以增强不同纹理类别的区分程度,使得每个类别只需训练一个样本就可实现较高的纹理识别表现. Liu 等^[17]将视觉模态纳入训练模型以解决零样本学习下的触觉物体识别问题. 为实现视觉 – 触觉跨模态 学习,他们系统地开发了一个新的字典学习框架,该框架通过对视觉和触觉模态之间的编码向量施加 相应的惩罚项,从而可以灵活地探索两种模态间的关联.此外,该框架还将触觉语义属性引入字典学 习中,以评估不同物类别间的相似度和区别度. Ren 等^[18]提出了一种基于渐进式小样本学习的注意 力吸引网络,该网络主要分为预训练阶段与元学习阶段. 在预训练阶段,该网络可以学到一个较好的 基类分类器和一个 CNN 特征提取器,且基类分类器的参数将在预训练后固定;在元学习阶段,通过训 练和测试迭代,使用学到的正则化器训练一个新类别分类器.最后,利用组合在一起的新类别和基类 别分类数据,对元学习的正则化器进行优化,使其也能在基类分类器上取得较好的表现. Xing 等^[19] 提出利用跨模态信息以增强基于度量的少样本学习方法,通过探索视觉图像与文本的语义表示不同的 特点,他们发展了一种自适应模态混合机制,可以根据学习的新图像类别自适应地结合两种模态的信 息. Snell 等^[20]提出一种原型网络结构,该网络将每个类别中的样本数据映射到一个空间中,提取每 个类别样本数据的均值来表示该类的原形. 然后使用欧几里得 (Euclid) 距离作为距离度量, 训练使得 同一类别数据到本类别原形表示的距离为最近, 到其他类别原形表示的距离较远. 测试时, 利用测试 数据到各个类别的原形表示的距离, 从而判断出测试数据的类别标签. Abderrahmane 等^[21]设计了机 器人触觉零样本学习系统, 提高了机器人识别从未遇到物体的能力. 该系统首先利用直接属性预测来 训练基于触觉形容词列表 (其中这些触觉形容词包含形状、纹理、材质等物理属性, 构成了与物体类 别相关的中间层) 的物体语义表示, 然后通过物体的语义表示可实现未知物体与已知物体间的知识传 递, 从而识别到未知物体.

但上述算法没有通过多种不同投影方式 (数据投影和编码向量投影) 去充分探索两种不同传感数据的内在关联,更没有考虑到不同传感数据本身子数据间的时序约束联系.另外,上述研究中大部分工作是基于深度学习算法的,然而传感器序列数据与图像不同,深度学习算法在序列数据取得的物体识别效果难以超越稀疏统计表示方法.为了克服上述相关工作的局限性并充分发挥多光谱序列数据的优势,本文提出一种 OSL-TCSR 方法以提升单样本学习下的物体识别结果,下面进行具体建模与分析.

3 OSL-TCSR 算法原理

为了识别同时给定的 Lumini 和 SCiO 样本属于 50 类不同物体的哪一类,本文提出一种新颖的 OSL-TCSR 算法. 该算法分为两个部分:训练部分和测试部分. 首先训练部分可分为两个阶段,在多 光谱数据相关性分析与重构阶段,本文分析两种原始数据在共享子空间的相关性特征,建立原始数据 和相关性特征数据的投影映射关系,并再将充分耦合后的两种数据进行重构;在时序约束稀疏表示阶段,本文将进行重构后的多种光谱数据导入设计的时序约束稀疏表示模型中,并在此模型中对多种光 谱数据进行联合约束. 然后在测试部分,本文通过训练部分学习到了参数对新的测试样本进行识别.

为了建立问题模型,分别使

$$T_{\rm L} = [t_{\rm L}^1 \quad t_{\rm L}^2 \quad \cdots \quad t_{\rm L}^{N_{\rm L}}] \in \mathbb{R}^{d_{\rm L} \times N_{\rm L}},$$

$$T_{\rm S} = [t_{\rm S}^1 \quad t_{\rm S}^2 \quad \cdots \quad t_{\rm S}^{N_{\rm S}}] \in \mathbb{R}^{d_{\rm S} \times N_{\rm S}},$$

(1)

其中 T_L 和 T_S 分别表示 Lumini 和 SCiO 的原始数据集. d_L 和 d_S 分别代表每个 Lumini 和 SCiO 样本的特征维度. N_L 和 N_S 分别代表 Lumini 和 SCiO 训练样本的数量. 因为本研究 Lumini 和 SCiO 是以配对形式存在的, 故 $N_L = N_S = N = 50$, t_L^i 和 t_S^i 分别是 T_L 和 T_S 中的第 i 个样本, $i \in \{1, 2, ..., N\}$.

3.1 多光谱数据相关性分析与重构

对于给定的 Lumini 和 SCiO 的原始数据,本文首先采用文献 [22] 中的聚类典型关联分析法 (cluster canonical correlation analysis, C-CCA) 先分别求解 Lumini 和 SCiO 投影到共享子空间的投影矩阵,然 后可以通过投影矩阵计算出相应的相关性特征数据.其中 C-CCA 可以表述为

$$\rho = \max_{P_{\rm L}, P_{\rm S}} \frac{P_{\rm L}' V_{xy} P_{\rm S}}{\sqrt{P_{\rm L}' V_{xx} P_{\rm L}} \sqrt{P_{\rm S}' V_{yy} P_{\rm S}}},\tag{2}$$

其中 ρ 代表 Lumini 和 SCiO 的关联度, $P_{L} \in \mathbb{R}^{d_{L} \times d_{L}}$ 和 $P_{S} \in \mathbb{R}^{d_{S} \times d_{S}}$ 分别是计算的投影矩阵, P'_{L} 和 P'_{S} 分别为相应矩阵的转置, 其中 V_{xy} , V_{xx} 和 V_{yy} 分别计算为

$$V_{xy} = \frac{1}{C} \sum_{c=1}^{C} t_{\rm L}^{c} t_{\rm S}^{c'}, \quad V_{xx} = \frac{1}{C} \sum_{c=1}^{C} t_{\rm L}^{c} t_{\rm L}^{c'}, \quad V_{yy} = \frac{1}{C} \sum_{c=1}^{C} t_{\rm S}^{c} t_{\rm S}^{c'}, \tag{3}$$

其中 V_{xy} , V_{xx} 和 V_{yy} 为协方差矩阵, t_{L}^{c} 和 t_{S}^{c} 分别代表 T_{L} 和 T_{S} 的第 c 类样本, $c \in \{1, 2, ..., C\}$, C 为 所有物体类别. $t_{L}^{c'}$ 和 $t_{S}^{c'}$ 分别为 t_{L}^{c} 和 t_{S}^{c} 的转置. 然后, 在共享子空间中 Lumini 和 SCiO 相关性特征 数据的联合字典学习模型为

$$\min_{D_{\mathrm{L}}^{(1)}, D_{\mathrm{S}}^{(1)}, X} H_{1} = \left\| \begin{bmatrix} P_{\mathrm{L}}' T_{\mathrm{L}} \\ P_{\mathrm{S}}' T_{\mathrm{S}} \end{bmatrix} - \begin{bmatrix} D_{\mathrm{L}}^{(1)} \\ D_{\mathrm{S}}^{(1)} \end{bmatrix} X \right\|_{\mathrm{F}}^{2} + \alpha \|X\|_{1}, \tag{4}$$

其中 α 为惩罚权重, $D_{L}^{(1)}$ 和 $D_{S}^{(1)}$ 分别为 Lumini 和 SCiO 的相关性特征数据的字典, X 为公共编码向 量. $\|X\|_1$ 为 X 的 L1 范数, 其作用是促进 X 稀疏化. 当计算得到 X 后, 则 Lumini 和 SCiO 原始数 据的联合字典学习模型为

$$\min_{D_{\mathrm{L}}^{(2)}, D_{\mathrm{S}}^{(2)}, Q_{\mathrm{L}}, Q_{\mathrm{S}}, \Lambda_{\mathrm{L}}, \Lambda_{\mathrm{S}}} H_{2} = \left\| \begin{bmatrix} T_{\mathrm{L}} \\ T_{\mathrm{S}} \end{bmatrix} - \begin{bmatrix} D_{\mathrm{L}}^{(2)} \Lambda_{\mathrm{L}} \\ D_{\mathrm{S}}^{(2)} \Lambda_{\mathrm{S}} \end{bmatrix} \right\|_{\mathrm{F}}^{2} + \beta \left\| \begin{bmatrix} X \\ X \end{bmatrix} - \begin{bmatrix} Q_{\mathrm{L}} \Lambda_{\mathrm{L}} \\ Q_{\mathrm{S}} \Lambda_{\mathrm{S}} \end{bmatrix} \right\|_{\mathrm{F}}^{2} + \gamma \left\| [Q_{\mathrm{L}} \ Q_{\mathrm{S}}] \right\|_{\mathrm{F}}^{2} + \delta \left\| [\Lambda_{\mathrm{L}} \ \Lambda_{\mathrm{S}}] \right\|_{1,1}, \quad (5)$$

其中 β , γ 和 δ 为惩罚参数, $D_{\rm L}^{(2)}$ 和 $D_{\rm S}^{(2)}$ 分别为 Lumini 和 SCiO 原始数据的字典, Λ_L 和 Λ_S 为原始数据的编码向量. $Q_{\rm L}$ 和 $Q_{\rm S}$ 分别为二次投影矩阵, 与 $P_{\rm L}$ 和 $P_{\rm S}$ 不同的是, 它们是对编码向量进行投影. $\|[Q_{\rm L} Q_{\rm S}]\|_{\rm F}^2$ 代表 $[Q_{\rm L} Q_{\rm S}]$ 的 F 范数平方, 其作用是防止 H_2 在优化过程中出现过拟合. $\|[\Lambda_{\rm L} \Lambda_{\rm S}]\|_{1,1}$ 代表 $[\Lambda_{\rm L} \Lambda_{\rm S}]$ 的 L1,1 范数, 其作用是促进 $\Lambda_{\rm L}$ 和 $\Lambda_{\rm S}$ 联合稀疏. 综合 H_1 和 H_2 , 则多光谱数据相关性 分析与重构阶段的联合优化模型为

$$\min_{\substack{D_{\rm L}^{(1)}, D_{\rm S}^{(1)}, D_{\rm L}^{(2)}, D_{\rm S}^{(2)}, X, \\ Q_{\rm L}, Q_{\rm S}, \Lambda_{\rm L}, \Lambda_{\rm S}}} H_1 + H_2,$$
(6)

在计算得到 $D_{\rm L}^{(1)}, D_{\rm S}^{(1)}, D_{\rm L}^{(2)}, D_{\rm S}^{(2)}, \Lambda_{\rm L}, \Lambda_{\rm S}$ 和 X 后,本文计算具有充分耦合联系的 Lumini 和 SCiO 原 始重构数据和相关性特征重构数据,如下所示:

$$T_{\rm L}^{(1)^*} = D_{\rm L}^{(1)} X \quad T_{\rm S}^{(1)^*} = D_{\rm S}^{(1)} X, \tag{7}$$

$$T_{\rm L}^{(2)^*} = D_{\rm L}^{(2)} \Lambda_{\rm L} \quad T_{\rm S}^{(2)^*} = D_{\rm S}^{(2)} \Lambda_{\rm S}, \tag{8}$$

其中, $T_{\rm L}^{(1)^*}$ 和 $T_{\rm S}^{(1)^*}$ 分别代表两种光谱在共享子空间的相关性特征重构数据, $T_{\rm L}^{(2)^*}$ 和 $T_{\rm S}^{(2)^*}$ 分别代表两种光谱的原始重构数据. 通过上述设计的多光谱数据相关性分析与重构方法, 不但增强了 Lumini 和 SCiO 光谱间的耦合联系, 而且增加了单样本学习下的样本数据量.

3.2 时序约束稀疏表示

在计算获得相关性特征重构数据 $T_{\rm L}^{(1)*}$, $T_{\rm S}^{(1)*}$ 和原始重构数据 $T_{\rm L}^{(2)*}$, $T_{\rm S}^{(2)*}$ 后, 我们设计稀疏表示 方法对它们进行二次融合时, 同时考虑重构后的每个数据的时序特征. 则多光谱时序约束稀疏表示模 型设计为

$$\min_{W} H_{3} = \frac{1}{2} \sum_{m=1}^{2} \left\| \begin{bmatrix} T_{\mathrm{L}}^{(m)^{*}} \\ T_{\mathrm{S}}^{(m)^{*}} \end{bmatrix} - \sum_{t=1}^{T} \begin{bmatrix} d_{\mathrm{L}}^{(m)} \\ d_{\mathrm{S}}^{(m)} \end{bmatrix} w^{(m)^{(t)}} \right\|_{\mathrm{F}}^{2} + \lambda_{1} \sum_{m=1}^{2} \left\| W^{(m)} \right\|_{2,1}$$

$$+ \lambda_2 \sum_{m=1}^{2} \left\| W^{(m)} \right\|_{1,1} + \lambda_3 \sum_{m=1}^{2} \sum_{t=1}^{T} \left\| w^{(m)^{(t+1)}} - w^{(m)^{(t)}} \right\|_1, \tag{9}$$

其中 $W = [W^{(1)} W^{(2)}]$ 为相关性特征重构数据和原始重构数据的联合编码矩阵. m = 1 指代相关性 特征重构数据, m = 2 指代原始重构数据, $m \in \{1,2\}$. $W^{(m)} = [w^{(m)^{(1)}} w^{(m)^{(2)}} \cdots w^{(m)^{(T)}}]$ 为各个时 刻子编码向量的联合编码矩阵. 通过 min_W H_3 计算重构数据的每个时刻的子编码向量 $w^{(m)^{(t)}}$ (本质上 $w^{(m)^{(t)}}$ 为 $W^{(m)}$ 的第 t 行), $\begin{bmatrix} d_1^{(m)} \\ d_3^{(m)} \end{bmatrix}^t$ 为 $\begin{bmatrix} D_1^{(m)} \\ D_2^{(m)} \end{bmatrix}$ 的第 t 列, 正则化项 $\sum_{m=1}^2 \sum_{t=1}^T \|w^{(m)^{(t+1)}} - w^{(m)^{(t)}}\|_1$ 保证了相邻两个子编码向量间的连续性 (本质上保证了每个具有时序特征的光谱数据内任意相邻两个 子数据间的连续性). $\sum_{m=1}^2 \|W^{(m)}\|_{2,1}$ 与 $\sum_{m=1}^2 \|W^{(m)}\|_{1,1}$ 分别为 W 的 L2,1 范数与 L1,1 范数. λ_1 , λ_2 和 λ_3 为惩罚权重.

上述时序约束稀疏表示模型比较复杂, 我们使用 APGD 去求解它. 首先 f(W) 函数设定为

$$f(W) = \frac{1}{2} \sum_{m=1}^{2} \left\| \begin{bmatrix} T_{\rm L}^{(m)^*} \\ T_{\rm S}^{(m)^*} \end{bmatrix} - \sum_{t=1}^{T} \begin{bmatrix} d_{\rm L}^{(m)} \\ d_{\rm S}^{(m)} \end{bmatrix} w^{(m)^{(t)}} \right\|_{\rm F}^2, \tag{10}$$

g(W)函数设定为

$$g(W) = \lambda_1 \sum_{m=1}^{2} \left\| W^{(m)} \right\|_{2,1} + \lambda_2 \sum_{m=1}^{2} \left\| W^{(m)} \right\|_{1,1} + \lambda_3 \sum_{m=1}^{2} \sum_{t=1}^{T} \left\| w^{(m)^{(t+1)}} - w^{(m)^{(t)}} \right\|_1.$$
(11)

为了近似复合函数 l(W) = f(W) + g(W),本文利用近似二阶泰勒展开 l(W, W(k)) 为

$$f(W(k)) + \sum_{m=1}^{2} \left\langle W^{(m)} - W^{(m)}(k), \nabla f\left(W^{(m)}(k)\right) \right\rangle + \frac{\mu}{2} \|W - W(k)\|_{\rm F}^{2} + g(W), \qquad (12)$$

其中 $\langle \cdot, \cdot \rangle$ 为内积运算, μ 为步长, 可使用线性搜索方法确定 μ , $\nabla f(W^{(m)}(k))$ 为 $f(W^{(m)})$ 在第 k 次 迭代时的导数. 然后使用近端梯度下降法 (proximal gradient descent, PGD) 求解第 k+1 次迭代时的 W(k+1):

$$W(k+1) = \arg\min\frac{1}{2} \|W - V(k)\|_{\rm F}^2 + \frac{1}{\mu}g(W), \qquad (13)$$

其中 $V^{(m)}(k) = W^{(m)}(k) - \frac{1}{\mu} \nabla f(W^{(m)}(k)), V(k) = [V^{(1)}(k) V^{(2)}(k)], W(k)$ 为第 k 次迭代计算的编码 向量. 由于 W 的每行是相互独立的, 为便于求解, 故将 W(k+1) 分解为多个独立行单独计算, 其中 W 的第 d 行计算为

$$W_{d}(k+1) = \arg\min\frac{1}{2} \|W_{d} - V_{d}(k)\|_{\mathrm{F}}^{2} + \frac{1}{\mu}g(W_{d})$$

$$= \arg\min\frac{1}{2} \|W_{d} - V_{d}(k)\|_{\mathrm{F}}^{2} + \frac{\lambda_{1}}{\mu}\sum_{m=1}^{2} \left\|W_{d}^{(m)}\right\|_{2} + \frac{\lambda_{2}}{\mu}\sum_{m=1}^{2} \left\|W_{d}^{(m)}\right\|_{1}$$

$$+ \frac{\lambda_{3}}{\mu}\sum_{m=1}^{2}\sum_{t=1}^{T} \left\|w_{d}^{(m)^{(t+1)}} - w_{d}^{(m)^{(t)}}\right\|_{1}.$$
 (14)

按照行独立分解后, $\sum_{m=1}^{2} \|W^{(m)}\|_{2,1}$ 转化为多个行 $\sum_{m=1}^{2} \|W_d^{(m)}\|_2$ 计算, 使得计算更加方便. 对于式 (14), 本文将其分解为组 LASSO 和融合 LASSO 问题, 并分段使用 PGD 计算, 其中组 LASSO 为

$$G_d(k+1) = \arg\min\frac{1}{2} \|W_d - V_d(k)\|_{\rm F}^2 + \frac{\lambda_1}{\mu} \sum_{m=1}^2 \|W_d^{(m)}\|_2.$$
(15)



图 2 (网络版彩图) APGD 与 PGD 收敛情况 Figure 2 (Color online) Convergence of APGD and PGD

在计算得到 $G_d(k+1)$ 后, 融合 LASSO 问题为

$$W_{d}(k+1) = \arg\min\frac{1}{2} \|W_{d} - G_{d}(k+1)\|_{\rm F}^{2} + \frac{\lambda_{2}}{\mu} \sum_{m=1}^{2} \|W_{d}^{(m)}\|_{1} + \frac{\lambda_{3}}{\mu} \sum_{m=1}^{2} \sum_{t=1}^{T} \|w_{d}^{(m)^{(t+1)}} - w_{d}^{(m)^{(t)}}\|_{1}.$$
(16)

经过一定次数迭代后, 通过 PGD 算法可以计算获得收敛的 $W_d(k+1)$. 为加快收敛速度, 最终使用改进的 APGD 算法计算 $W_d(k+1)$, 首先使用 P(k) 替代 W(k):

$$P(k) = W(k) + \frac{[\alpha(k-1)-1]}{\alpha(k)} (W(k) - W(k-1)), \qquad (17)$$

其中 $\alpha(k) = (1 + \sqrt{1 + 4\alpha(k-1)^2})/2$, 设定初始值 $\alpha(0) = 1$, $P(k) = [P^{(1)}(k) P^{(2)}(k)]$. 然后 $V^{(m)}(k)$ 被替换为

$$V^{(m)}(k) = P^{(m)}(k) - \frac{1}{\mu} \nabla f\left(P^{(m)}(k)\right).$$
(18)

则最终经过 APGD 加速后的 W(k+1) 计算方式为

$$W(k+1) = \arg\min\frac{1}{2} \|W - V(k)\|_{\rm F}^2 + \frac{1}{\mu}g(W), \qquad (19)$$

其中 $V(k) = [V^{(1)}(k) V^{(2)}(k)]$, APGD 与 PGD 计算目标函数 W(k+1) 的误差收敛情况如图 2 所示, 经过加速后, APGD 大幅度减少了迭代次数.

4 算法实现

第 3 节已阐述 OSL-TCSR 训练部分的两个阶段. 第 1 阶段为多光谱数据相关性分析与重构, 第 2 阶段为时序约束稀疏表示. 接下来本节继续阐述如何优化两阶段目标函数的各参数, 直至目标函数 收敛.

4.1 第1阶段参数更新

式 (6) 显而易见是难以直接求解的,因此本文采用交替优化的方式分别更新各参数,直至达到预 设定最大迭代次数,主要分为下面几步.

4.1.1 更新相关性特征数据的公共编码向量

在通过 C-CCA 法计算得到投影矩阵 $P_{\rm L}$ 和 $P_{\rm S}$ 后,本文分别固定 Lumini 和 SCiO 相关性特征数据的字典 $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$ 以及 Lumini 和 SCiO 原始数据的二次投影矩阵 $Q_{\rm L}$, $Q_{\rm S}$,编码向量 $\Lambda_{\rm L}$, $\Lambda_{\rm S}$,则两种相关性特征数据的公共编码向量 X 更新方法如下:

$$\min_{X} \left\| \begin{bmatrix} P_{\mathrm{L}}^{\prime} T_{\mathrm{L}} \\ P_{\mathrm{S}}^{\prime} T_{\mathrm{S}} \\ \sqrt{\beta} Q_{\mathrm{L}} \Lambda_{\mathrm{L}} \\ \sqrt{\beta} Q_{\mathrm{S}} \Lambda_{\mathrm{S}} \end{bmatrix} - \begin{bmatrix} D_{\mathrm{L}}^{(1)} \\ D_{\mathrm{S}}^{(1)} \\ \sqrt{\beta} I_{\mathrm{L}} \\ \sqrt{\beta} I_{\mathrm{S}} \end{bmatrix} X \right\|_{\mathrm{F}}^{2} + \alpha \|X\|_{1},$$
(20)

其中 I_L 和 I_S 为单位矩阵. 上式为 LASSO 估计问题, 可通过快速软阈值迭代算法 (fast soft threshold iterative algorithm, FISTA)^[23] 解决, 这里使用稀疏模型软件包 SPAMS^[24] 求解.

4.1.2 更新相关性特征数据的字典

在通过 C-CCA 法计算得到投影矩阵 $P_{\rm L}$ 和 $P_{\rm S}$ 后,本文固定 Lumini 和 SCiO 相关性特征数据的 公共编码向量 X, 然后 Lumini 和 SCiO 相关性特征数据的字典 $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$ 更新方法如下:

$$\min_{D_{\mathrm{L}}^{(1)}, D_{\mathrm{S}}^{(1)}} \left\| P_{\mathrm{L}}^{\prime} T_{\mathrm{L}} - D_{\mathrm{L}}^{(1)} X \right\|_{\mathrm{F}}^{2} + \left\| P_{\mathrm{S}}^{\prime} T_{\mathrm{S}} - D_{\mathrm{S}}^{(1)} X \right\|_{\mathrm{F}}^{2}$$
s.t. $\left\| d_{\mathrm{L}}^{(1)} \right\|_{2} \leq 1, \left\| d_{\mathrm{S}}^{(1)} \right\|_{2} \leq 1$ for $k = 1, 2, \dots, K.$
(21)

这是经典的字典学习问题,这里使用块坐标下降法 (BCD)^[25] 求解.

4.1.3 更新原始数据的编码向量

分别固定 Lumini 和 SCiO 原始数据的字典 $D_{\rm L}^{(2)}$, $D_{\rm S}^{(2)}$, 二次投影矩阵 $Q_{\rm L}$, $Q_{\rm S}$, Lumini 和 SCiO 相关性特征数据的公共编码向量 X, 然后 Lumini 和 SCiO 原始数据的编码向量 $\Lambda_{\rm L}$ 和 $\Lambda_{\rm S}$ 更新方法 如下:

$$\min_{\Lambda_{\rm L},\Lambda_{\rm S}} \left\| \begin{bmatrix} T_{\rm L} \\ \sqrt{\beta}X \end{bmatrix} - \begin{bmatrix} D_{\rm L}^{(2)} \\ \sqrt{\beta}Q_{\rm L} \end{bmatrix} \Lambda_{\rm L} \right\|_{\rm F}^{2} + \left\| \begin{bmatrix} T_{\rm S} \\ \sqrt{\beta}X \end{bmatrix} - \begin{bmatrix} D_{\rm S}^{(2)} \\ \sqrt{\beta}Q_{\rm S} \end{bmatrix} \Lambda_{\rm S} \right\|_{\rm F}^{2} + \delta \| [\Lambda_{\rm L} \ \Lambda_{\rm S}] \|_{1,1}.$$
(22)

类似地,上式本质上为 L1 范数优化问题,这里通过 FISTA^[23] 或最小角回归 (least angle regression, LAR)^[26] 求解.

4.1.4 更新原始数据的字典

分别固定 Lumini 和 SCiO 原始数据的编码向量 Λ_L 和 Λ_S ,则原始数据的字典 $D_L^{(2)}$ 与 $D_S^{(2)}$ 更新 如下:

$$\min_{D_{\mathrm{L}}^{(1)}, D_{\mathrm{S}}^{(1)}} \left\| T_{\mathrm{L}} - D_{\mathrm{L}}^{(2)} \Lambda_{\mathrm{L}} \right\|_{\mathrm{F}}^{2} + \left\| T_{\mathrm{S}} - D_{\mathrm{S}}^{(2)} \Lambda_{\mathrm{S}} \right\|_{\mathrm{F}}^{2} \\
\text{s.t.} \quad \left\| d_{\mathrm{L}}^{(2)}{}_{k} \right\|_{2} \leq 1, \left\| d_{\mathrm{S}}^{(2)}{}_{k} \right\|_{2} \leq 1 \text{ for } k = 1, 2, \dots, K.$$
(23)

类似地,这里使用 BCD [25] 求解.

算法 1 Multispectral data correlation analysis and reconstruction

Input: $T_{\rm L}$ and $T_{\rm S}$. Initialization: $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$, $D_{\rm L}^{(2)}$, $D_{\rm S}^{(2)}$, $Q_{\rm L}$, $Q_{\rm S}$, and Max-iterations = 20. 1: Use C-CCA method to calculate $P_{\rm L}$ and $P_{\rm S}$. 2: for step = 1: Max-iterations do 3: Fix $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$, $Q_{\rm L}$, $Q_{\rm S}$, $\Lambda_{\rm L}$, and $\Lambda_{\rm S}$, update X using (20); 4: Fix X, update $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$ using (21); 5: Fix $D_{\rm L}^{(2)}$, $D_{\rm S}^{(2)}$, $Q_{\rm L}$, $Q_{\rm S}$, and X, update $\Lambda_{\rm L}$ and $\Lambda_{\rm S}$ using (22); 6: Fix $\Lambda_{\rm L}$ and $\Lambda_{\rm S}$, update $D_{\rm L}^{(2)}$, $D_{\rm S}^{(2)}$ using (23);

7: Fix $\Lambda_{\rm L}$, $\Lambda_{\rm S}$, and X, update $Q_{\rm L}$ and $Q_{\rm S}$ using (24);

8: end for

Output: $D_{\rm L}^{(1)}, D_{\rm S}^{(1)}, D_{\rm L}^{(2)}, D_{\rm S}^{(2)}, \Lambda_{\rm L}, \Lambda_{\rm S}$, and X.

4.1.5 更新二次投影矩阵

分别固定 Lumini 和 SCiO 原始数据的编码向量 Λ_L 和 Λ_S , Lumini 和 SCiO 相关性特征数据的公 共编码向量 X, 然后原始数据到相关性特征数据的二次投影矩阵 Q_L 和 Q_S 更新方法如下:

$$\min_{Q_{\rm L},Q_{\rm S}} \|X - Q_{\rm L}\Lambda_{\rm L}\|_{\rm F}^2 + \|X - Q_{\rm S}\Lambda_{\rm S}\|_{\rm F}^2 + \frac{\gamma}{\beta} \|[Q_{\rm L}Q_{\rm S}]\|_{\rm F}^2.$$
(24)

显而易见, 上式为岭回归问题, 可以使用梯度下降法解决.

4.1.6 第 1 阶段算法总结

上述更新程序交替进行,直至达到设定的最大迭代次数,具体如算法1所示.最终计算得到上述 各参数后,本文再进行重构具有充分耦合联系的相关性特征数据和原始数据.

4.2 第2阶段参数更新

在重构具有充分耦合联系的相关性特征数据 $T_{\rm L}^{(1)^*}$, $T_{\rm S}^{(1)^*}$ 和原始数据 $T_{\rm L}^{(2)^*}$, $T_{\rm S}^{(2)^*}$ 后, 将其输入到 式 (9), 经过设计的 APGD 加速后得到 W(k+1), 可计算为

$$W(k+1) = \arg\min\frac{1}{2} \|W - V(k)\|_{\rm F}^2 + \frac{1}{\mu}g(W).$$
(25)

为了求解 W(k+1), 这里将其转化为组 LASSO 和融合 LASSO 问题分段计算. 类似地, 依然按照多个 行单独计算, 第1段为

$$G_d(k+1) = \arg\min\frac{1}{2} \|W_d - V_d(k)\|_{\rm F}^2 + \frac{\lambda_1}{\mu} \sum_{m=1}^2 \|W_d^{(m)}\|_2.$$
(26)

根据文献 [27], 求解 G_d(k+1) 为

$$G_{d}(k+1) = \begin{cases} \left(1 - \frac{\lambda_{1}}{\mu \|V_{d}(k)\|_{2}^{2}}\right) V_{d}(k), \text{ if } \|V_{d}(k)\|_{2}^{2} > \frac{\lambda_{1}}{\mu}, \\ 0, \qquad \text{otherwise.} \end{cases}$$
(27)

在计算得到 G_d(k + 1) 后, 在第 2 段中, 根据文献 [28], 这里用简单的软阈值估计求解 W_d(k + 1), 首先 计算

$$\pi_{\frac{\lambda_3}{\mu}}^0 \left(G_d \left(k+1 \right) \right) = \arg\min\frac{1}{2} \left\| W_d - G_d \left(k+1 \right) \right\|_{\mathrm{F}}^2 + \frac{\lambda_3}{\mu} \sum_{m=1}^2 \sum_{t=1}^T \left\| w_d^{(m)^{(t+1)}} - w_d^{(m)^{(t)}} \right\|_1, \quad (28)$$

算法 2 Temporally constrained sparse representation

Input: $T_{\rm L}^{(1)^*}$ and $T_{\rm S}^{(1)^*}$, $T_{\rm L}^{(2)^*}$ and $T_{\rm S}^{(2)^*}$, $D_{\rm L}^{(1)}$, $D_{\rm S}^{(1)}$, $D_{\rm L}^{(2)}$, and $D_{\rm S}^{(2)}$. **Initialization:** $\mu = 1$, $\alpha (0) = 1$, and Max-iterations = 500. 1: Use SPAMS to calculate W(0), and make W(1) = W(0). 2: while 1 do 3: trigger-flag = 0;W(k-1) = W(0), W(k) = W(k-1);4: 5: for step = 1: Max-iterations do $\alpha(k) = (1 + \sqrt{1 + 4\alpha(k-1)^2})/2;$ 6: $P(k) = W(k) + \frac{[\alpha(k-1)-1]}{\alpha(k)} (W(k) - W(k-1));$ $V^{(m)}(k) = P^{(m)}(k) - \frac{1}{\mu} \nabla f (P^{(m)}(k));$ 7: 8: for d = 1 : N do 9: $G_{d}(k+1) = \arg\min \frac{1}{2} \|W_{d} - V_{d}(k)\|_{\mathrm{F}}^{2} + \frac{\lambda_{1}}{\mu} \sum_{m=1}^{2} \|W_{d}^{(m)}\|_{2};$ 10: $W_{d}\left(k+1\right) = \arg\min\frac{1}{2} \left\|W_{d} - G_{d}\left(k+1\right)\right\|_{\mathrm{F}}^{2} + \frac{\lambda_{2}}{\mu} \sum_{m=1}^{2} \left\|W_{d}^{(m)}\right\|_{1} + \frac{\lambda_{3}}{\mu} \sum_{m=1}^{2} \sum_{t=1}^{T} \left\|w_{d}^{(m)(t+1)} - w_{d}^{(m)(t)}\right\|_{1};$ 11: 12:end for 13: Use (30) to determine μ whether meeting the condition: if meeting, trigger-flag = 0; else, trigger-flag = 1; end for 14: If trigger-flag = 0, break; 15:16: $\mu = 2\mu;$ 17: end while Output: W(k+1).

然后计算

$$W_{d}(k+1) = \begin{cases} \pi_{\frac{\lambda_{3}}{\mu}}^{0} \left(G_{d}(k+1)\right) - \frac{\lambda_{2}}{\mu}, & \text{if } \pi_{\frac{\lambda_{3}}{\mu}}^{0} \left(G_{d}(k+1)\right) \geqslant \frac{\lambda_{2}}{\mu}, \\ 0, & \text{if } \pi_{\frac{\lambda_{3}}{\mu}}^{0} \left(G_{d}(k+1)\right) \leqslant \frac{\lambda_{2}}{\mu}, \\ \pi_{\frac{\lambda_{3}}{\mu}}^{0} \left(G_{d}(k+1)\right) + \frac{\lambda_{2}}{\mu}, & \text{if } \pi_{\frac{\lambda_{3}}{\mu}}^{0} \left(G_{d}(k+1)\right) \leqslant -\frac{\lambda_{2}}{\mu}. \end{cases}$$
(29)

上述 $G_d(k+1)$ 和 $W_d(k+1)$ 的计算可通过高效投影稀疏学习包 SLEP ^[28] 实现. 对于动态步长 μ , 需 满足下面条件:

$$\sum_{m=1}^{2} f\left(W^{(m)}\left(k+1\right)\right) \leqslant \sum_{m=1}^{2} f\left(W^{(m)}\left(k\right)\right) + \nabla f\left(W^{(m)}\left(k\right)\right)' \\ \times \left(W^{(m)}\left(k+1\right) - W^{(m)}\left(k\right)\right) \\ + \frac{\mu}{2} \left\|W^{(m)}\left(k+1\right) - W^{(m)}\left(k\right)\right\|_{\mathrm{F}}^{2}.$$
(30)

上述 W(k+1) 的更新分两步交替进行, 直至达到设定的最大迭代次数, 具体如算法 2 所示.

5 分类器设计

在第3节的目标函数收敛之后,可计算获得关于训练样本的字典 $D_{\rm L}^{(1)^*}$, $D_{\rm S}^{(2)^*}$,



图 3 本文提出的 OSL-TCSR 算法模型框架 Figure 3 The framework of the proposed OSL-TCSR in this paper

特征数据 $\hat{t_L}^{(2)}$ 和 $\hat{t_S}^{(2)}$. 首先将两对测试样本输入时序约束稀疏表示算法中,

$$\min_{\hat{W}} H_{3} = \frac{1}{2} \sum_{m=1}^{2} \left\| \begin{bmatrix} \hat{t}_{\mathrm{L}}^{(m)} \\ \hat{t}_{\mathrm{S}}^{(m)} \end{bmatrix} - \sum_{t=1}^{T} \begin{bmatrix} d_{\mathrm{L}}^{(m)^{*}} \\ d_{\mathrm{S}}^{(m)^{*}} \end{bmatrix} \hat{w}^{(m)^{(t)}} \right\|_{\mathrm{F}}^{2} + \lambda_{1} \sum_{m=1}^{2} \left\| \hat{W}^{(m)} \right\|_{2,1} + \lambda_{2} \sum_{m=1}^{2} \left\| \hat{W}^{(m)} \right\|_{1,1} + \lambda_{3} \sum_{m=1}^{2} \sum_{t=1}^{T} \left\| \hat{w}^{(m)^{(t+1)}} - \hat{w}^{(m)^{(t)}} \right\|_{1},$$
(31)

其中 $\begin{bmatrix} a_{L}^{(m)^*} \\ a_{S}^{(m)^*} \end{bmatrix}^t$ 为 $\begin{bmatrix} D_{L}^{(m)^*} \\ D_{S}^{(m)^*} \end{bmatrix}^t$ 的第 t 列, $\widehat{W} = [\widehat{W}^{(1)} \ \widehat{W}^{(2)}]$, $\widehat{W}^{(m)} = [\widehat{w}^{(m)^{(1)}} \ \widehat{w}^{(m)^{(2)}} \ \cdots \ \widehat{w}^{(m)^{(T)}}]$, 然后通过 APGD 算法计算获得测试样本的编码向量 $\widehat{W}^{(1)}$, $\widehat{W}^{(2)}$. 由于训练样本的标签是已知的, 所以最后通 过训练样本的编码向量与测试样本的编码向量间的余弦距离估计测试样本的标签. 为便于计算, 这里 使用两种池化模式对原始数据编码向量和相关性特征数据编码向量进行融合. 其中训练样本的编码向 量为

$$W^* = \left(W^{(1)^*} + W^{(2)^*}\right)/2,\tag{32}$$

而测试样本的编码向量为

$$\widehat{W} = \left(\widehat{W}^{(1)} + \widehat{W}^{(2)}\right)/2.$$
(33)

最后利用余弦距离计算测试样本的物体标签属性:

$$l^* = \arg\max_{l \in \{1,2,\dots,N\}} \frac{\widehat{W}W^{*(l)'}}{\|\widehat{W}\|_2 \|W^{*(l)'}\|_2}.$$
(34)

*l** 代表我们计算的测试样本物体标签, *W**(*l*)['] 代表 *W** 第*l* 行的转置. 因为本文从事单样本学习研究, 所以 *W**(*l*) 也代表第*l* 个物体类别样本的编码向量. 结合上述的多光谱数据相关性分析与重构、时序 约束稀疏表示和分类器设计, 本文提出的 OSL-TCSR 主要框架如图 3 所示.

6 实验结果与分析

传统的接触式传感器在获取数据时由于与物体摩擦会产生噪声、延时等问题,导致获取的数据存 在着多样性噪声,需要经过一定的降噪、特征提取等方法才能导入算法中分类,并且分类表现较差.而 与传统接触式测量相比,非接触式传感器获取的光谱数据具有低噪声、高表征能力等特点,对不同物 体形状具有比较强的适应性.并且我们发现光谱数据相邻波段之间具有时序连续性,这种时序特性对 物体识别研究是有效的.故本文选择一种由非接触式传感器(光谱仪)所收集的数据集(SMM50数据 集^[3])用于验证提出的 OSL-TCSR 方法,该数据集由 5 种不同材料下的 50 个物体类别组成(每种材 料细分为 10 种不同物体),本文工作是识别 5 种不同材料下的 50 个物体类别.对于每次物体识别实 验,我们首先使用 Z-score 数据标准化方法对光谱数据进行预处理以消除两种光谱样本不同量级的影 响. 然后,从每个物体类别中(每个物体类别包含 100 个样本)随机抽取一个样本用于训练 OSL-TCSR, 并将剩余的样本用于测试 OSL-TCSR,这样保证训练样本和测试样本互不相关.最后,利用余弦距离 (式 (34))计算训练样本与测试样本间的相似度,并通过相似度计算得到未知测试样本的物体标签,从 而检验 OSL-TCSR 模型的识别表现.实验在设备 LAPTOP-0VKE2AD9 上进行,此电脑处理器为 Intel(R) Core(TM) i7-8750 CPU @ 2.20 GHz,机带 16.0 GB RAM.

6.1 相关性分析方法选择

(1) 典型关联分析法 (canonical correlation analysis, CCA): 分析两组数据线性相关性的方法, 通过 余弦距离计算两组数据的投影矩阵, 然后通过投影矩阵将两种数据投影到共享子空间, 从而得到两组 不同数据的相关性特征, 缺点是在子空间中不能分离两个数据集合.

(2) 聚类典型关联分析法 (C-CCA)/均值典型关联分析法 (mean canonical correlation analysis, M-CCA): 针对单样本学习情况进行物体识别研究, C-CCA 和 M-CCA 计算方式相同, 它们与 CCA 不同 之处是引入物体类别为导向进行协方差计算. 它们通过余弦距离计算两组数据的投影矩阵, 然后通过 投影矩阵将两种数据投影到共享子空间, 在子空间中能将数据集较好地分离.

(3) 改进的聚类典型关联分析法 (C-CCA-1, C-CCA-2): 两种方法均来自文献 [29], 在 C-CCA-1 中, 通过修改两组数据间的协方差计算方式, 使具有相似度较大的不同类别的数据质心尽可能远离. 对于 C-CCA-2, 在构造数据间的协方差时, 不是使数据质心远离, 而是尝试使不同类中的相似数据尽可能地 远离. 上述提到的相关性分析方法用在多光谱数据相关性分析与重构阶段, 且各个相关性分析方法得 到的相关性特征数据不同, 从而导致物体识别表现不同. 为了选取最佳的相关性分析方法, 本文对上述 方法进行了识别准确率比较, 如图 4 所示.

由图 4 可知,改进的 C-CCA-2 识别效果明显优于其他相关性分析方法,原因是这种改进数据间协方差方式使属于不同物体类别的所有相似性数据尽可能远离,加大了不同类别区分性特征,从而减小误分类.而 C-CCA-1 仅仅使不同类别的数据质心进行远离,只能整体地分离不同类别的相似性数据,所以分类结果相对于 C-CCA-2 较差.并且从图 4 中发现,C-CCA 和 M-CCA 在单样本学习情况下的识别效果相同,原因是 C-CCA 在计算数据间和数据内协方差矩阵时考虑了每个类别的所有数据特征,而 M-CCA 考虑的是每个类别所有数据的均值特征,所以在单样本学习下,C-CCA 与 M-CCA 识别结果相同.最后由于 CCA 没引入物体类别信息,因此不能在共享子空间中有效分离两种数据集,导致计算的相关性特征数据融入 OSL-TCSR 算法中识别效果较差.









6.2 相关分类算法分析

在 OSL-TCSR 的多光谱数据相关性分析与重构阶段,本文选定最优的相关性分析方法 C-CCA-2. 然后,将最新的相关分类算法与 OSL-TCSR 进行比较,以验证 OSL-TCSR 的有效性.

(1) 改进的深度神经网络 (improved deep neural network, IDNN): 参考文献 [3] 神经网络结构, 该 算法采用两个平行深度网络结构 (每个平行深度网络包含两个 256 节点层和两个 128 节点层, 并在每 层之后进行了 Dropout 和 Relu 操作) 分别对两种光谱序列进行处理, 然后在全连接层对两种光谱序 列的深层特征进行融合.

(2) 广义耦合式字典学习 (generalized coupled dictionary learning, GCDL): 该算法在文献 [29] 中提出,利用聚类规范相关分析 C-CCA 探索不同模态数据间相同类的相关性特征,且可通过稀疏编码特点进行模式识别,也可用于跨模态检索.它可以处理配对的多模态数据,被发展为 GCDL-1 和 GCDL-2.

(3) 结构与非凸约束限制的核组稀疏表示 (kernel group sparse representation via structural and non-convex constraints, KGSRSN): 该算法在文献 [30] 中提出, 通过将训练样本映射到核空间, 缓解了范数归一化问题.此外, 在不牺牲解的唯一性和凸优化的鲁棒性前提下, 提供惩罚函数参数一定的值 区间以促进稀疏性. 但该算法没有探索两种光谱间的耦合特征.

(4) 分离式核稀疏编码 (separate kernel sparse coding, SKSC): 该算法在文献 [18] 中提出, 将两种 光谱数据分别进行核稀疏编码计算, 同时对两种光谱数据的编码向量施加不同的稀疏模式约束. 这里 选用的核函数为高斯 RBF 核.

(5) 多模态单样本字典学习 (dictionary learning with one-shot, DL-O): 该算法在文献 [2] 中提出, 它首先利用稀疏表示得到多个不同模态 (视觉、声音和加速度) 的公共编码向量, 然后设计类间正则化 项, 以增强不同类别的区分程度, 使得每个类别只需训练单个样本就可实现较高的识别表现.

(6) 单样本学习下稀疏表示方法 (one-shot learning with sparse representation, OSL-SR): 利用 OSL-TCSR 的第一阶段获得编码向量, 然后通过余弦距离计算分类结果, 没有考虑到数据的时序特征.

(7) 单样本学习下时序约束稀疏表示方法 (OSL-TCSR): 经过多光谱数据相关性分析与重构后, 将获得的具有耦合特性的多种光谱数据导入设计的时序约束稀疏表示方法中, 使得在多光谱融合过程中考虑到数据的时序特征, 以提高物体识别结果.

各算法在单样本学习下的物体识别结果如图 5 所示. 由该图可知, OSL-TCSR 通过考虑多光谱序 列间的相关特征和光谱序列内的时序性, 相对于其他算法在单样本学习情况下物体识别准确率更高, 而 OSL-SR 由于忽略了光谱数据的时序特征, 导致物体识别结果略低于 KGSRSN. 原因是 KGSRSN

OSL-SR				
Plastic	Wood	Fabric	Metal	Paper
PVC/1.00	Red $Elm/0.62$	Satin/1.00	Lead/0.98	Cardboard/1.00
Polyethylene Green $/1.00$	Cherry/1.00	Cotton Towel/ 1.00	Iron/0.73	Textbook Paper/1.00
Polyethylene Yellow/1.00 $$	Hard Maple/ 1.00	Wool/0.46	Titanium/1.00	Constr. Paper Orange/1.00
Polyethylene $\operatorname{Red}/1.00$	Red Oak $/0.81$	Denim/1.00	Copper/1.00	Constr. Paper $\operatorname{Red}/1.00$
ThermoTeflon/1.00	Walnut/1.00	Cotton Canvas/ 1.00	Brass/0.62	Newspaper $/0.42$
Polyethylene Blue/1.00	Red Cedar/ 0.98	Gauze/1.00	Steel/0.60	Receipt Paper $/0.99$
PET/1.00	White $Oak/0.95$	Flannel/1.00	Aluminum Foil $/0.62$	Constr. Paper Green $/1.00$
PP/1.00	Ash/1.00	Linen/1.00	$\operatorname{Zinc}/1.00$	Printer Paper/1.00
ThermoPolypropylene/0.25	Hickory/0.99	Cotton Sweater/1.00	Aluminum/1.00	Magazine Paper $/0.80$
HDPE/1.00	Curly Maple/ 0.93	Felt/1.00	Magnesium/0.88	Notebook Paper/1.00

表 1 OSL-SR 对不同材料下每个物体的识别表现

Table 1 OSL-SR's recognition performance of each object with different materials

考虑了模态内组稀疏限制、结构局部性、非凸约束和非线性映射等优点,但迁移至多光谱融合分类问 题中, 忽略了光谱间的相关性, 所以相对于 OSL-TCSR 效果较差. SKSC 使用高斯 RBF 核函数将原 始低维空间数据的非线性特征映射到高维空间数据的线性特征,并通过核距离计算任意两个样本间的 相似性,最后将两种光谱的残差累加后进行分类,故相对于 GCDL, GCDL-1 和 GCDL-2 等方法取得 了较好分类表现,但由于没探索两种光谱数据其编码向量间的共性特征,故准确率低于 OSL-TCSR. GCDL, GCDL-1 和 GCDL-2 由于只在数据的编码向量上建立不同光谱相同类之间的联系, 而并非类 似于 OSL-SR 通过对原始数据进行相关性分析并生成相似性特征数据以进行数据增强,也没考虑光谱 序列的时序特性,所以更适用于跨模态应用场景,而在物体识别场景下准确率较低,最新基于稀疏表 示的单样本学习算法 DL-O 设计公共编码向量以探索单样本学习下多模态间的公共特征,并且设计类 间正则化项以增强不同类别的区分程度,但这种设计没有充分考虑到不同模态的数据特点,也没有充 分发挥模态间的相关性. 故只取得了 90.9% 的准确率, 比 OSL-TCSR 的物体识别结果低 1.6%. 深度 网络结构一般在图像分类任务中能达到较好的表现,但迁移至光谱序列,其卷积结构很难发挥较好作 用. 故处理序列数据时, 深度网络一般采用类似于多层感知机的神经元全连接方式. 由于 IDNN 其结 构限制忽略掉了序列的时序联系,且只能在全连接层对两种光谱特征进行融合,故其物体识别结果低 于 OSL-TCSR, OSL-SR, KGSRSN, SKSC, DL-O 等稀疏表示方法. 通过比较结果也能进一步表明稀疏 表示方法在序列信号的压缩感知方面有着优越的表现,通过学习到的字典和编码向量的线性组合,能 有效重构复杂的序列数据.为进一步直观体现 OSL-SR 和 OSL-TCSR 对每个物体的具体识别结果,本 文使用混淆矩阵分别呈现 OSL-SR 和 OSL-TCSR 在单样本学习下不同材料下各个物体类别的识别情 况,如表1和2所示.从中我们做出以下分析:

(1) OSL-TCSR 在同一材料不同物体的识别结果更接近 (如织物材料物体的识别结果), 而在不同 材料下不同物体的识别结果更加远离 (如织物材料物体与金属材料物体的识别结果), 说明物体的识别 表现与其材料属性相关, 相同材料属性的物体其测量的光谱数据更相似.

(2) 金属材料物体相比其他 4 种材料物体的识别表现较差, 这是因为金属材料物体可以反射红外光 谱信号, 从而影响了光谱传感器测量性能; 而织物材料的物体由于能够吸收光谱信号, 使得 OSL-TCSR 算法能够完全识别出这种材料的物体, 准确率达到 100%. 故材料反射、吸收光线等因素会影响算法实

OSL-TCSR					
Plastic Wood		Fabric Metal		Paper	
PVC/1.00	Red $Elm/0.41$	Satin/1.00	Lead/0.49	Cardboard/1.00	
Polyethylene Green $/1.00$	Cherry/1.00	Cotton Towel/ 1.00	Iron/0.49	Textbook Paper/ 1.00	
Polyethylene Yellow/1.00 $$	Hard Maple/ 1.00	Wool/1.00	Titanium/1.00	Constr. Paper Orange/1.00	
Polyethylene $\operatorname{Red}/1.00$	Red Oak $/0.85$	Denim/1.00	Copper/1.00	Constr. Paper Red/1.00 $$	
ThermoTeflon/1.00	Walnut/1.00	Cotton Canvas/ 1.00	Brass/0.49	Newspaper $/0.97$	
Polyethylene Blue/1.00	Red Cedar $/0.88$	Gauze/1.00	Steel/1.00	Receipt Paper/ 1.00	
PET/1.00	White $Oak/0.88$	Flannel/1.00	Aluminum Foil $/0.74$	Constr. Paper Green/1.00	
PP/0.99	Ash/1.00	Linen/1.00	$\operatorname{Zinc}/0.49$	Printer Paper/1.00	
ThermoPolypropylene/1.00	Hickory/0.63	Cotton Sweater/1.00	Aluminum/0.96	Magazine Paper $/0.95$	
HDPE/1.00	Curly Maple/ 1.00	Felt/1.00	Magnesium/1.00	Notebook Paper/1.00	

表 2 OSL-TCSR 对不同材料下每个物体的识别表现

Table 2 OSL-TCSR's recognition performance of each object with different materials

表 3 光谱通道划分

	Table 3 Spectral channel division	
Spectroscopy	Wavelength range (nm)	Spectral channel
	317–380	Near-ultraviolet
Lumini	381-780	Visible-light
	781-856	Near-infrared
SCIO	740–780	Visible-light
5010	781 - 1070	Near-infrared

际的识别结果.

6.3 光谱通道分析

Lumini 和 SCiO 的光谱特性彼此不同, Lumini 的波长测量范围为 $\lambda = 317-856$ nm, 该光谱涵盖 了整个可见光区以及部分近紫外区和近红外区.而 SCiO 的波长测量范围为 $\lambda = 740-1070$ nm, 只能 在可见光区和近红外区中工作.本文根据光谱波段特性将两种光谱划分为多个通道, 如表 3 所示. 然 后, 利用通道特性,本文设定了 6 种通道组合模式, 以更加充分验证 OSL-TCSR 的有效性.在 6.2 节, 由于 KGSRSN 相对于其他算法有着更好的识别性能, 故本节将其作为对比算法来验证 OSL-TCSR 的 物体识别表现.两种算法在不同通道组合下的物体识别准确率如表 4 所示, 根据表中结果做出如下分 析: (1) 相对于其他通道组合, 组合 5 中 OSL-TCSR 和 KGSRSN 都有着较好的物体识别准确率, 分别 为 91.4% 和 90.3%. 说明 Lumini 的有效光谱特征主要体现在可见光区, SCiO 体现在近红外区. (2) 在 光谱通道较差的组合中 (组合 1, 3 和 6), OSL-TCSR 相对于 KGSRSN 有着更显著的识别效果, 说明 OSL-TCSR 在较差光谱特征中鲁棒性较强, 这是因为 OSL-TCSR 在样本特征较差的环境中利用到了 数据本身特点, 充分探索了两种光谱数据的耦合特征和时序特征.

Table 4 Comparison of recognition accuracy (%) of different spectral channel combinations						
Lumini	Near-ultraviolet	Visible-light	Near-infrared	Near-ultraviolet	Visible-light	Near-infrared
SCiO	Visible-light	Visible-light	Visible-light	Near-infrared	Near-infrared	Near-infrared
Groups	group1	group2	group3	group4	group5	group6
OSL-TCSR	72.7	86.8	73.4	87.7	91.4	86.2
KGSRSN	60.7	86.7	69.4	84.7	90.3	78.7

表 4 不同光谱通道组合识别准确率 (%) 比较

表 5 参数 λ_1, λ_2 和 λ_3 对物体识别性能影响

Table 5 Influence of parameters λ_1 , λ_2 and λ_3 on object recognition performance

λ_1	λ_2	λ_3	Accuracy (%)
10^{-2}	10^{-4}	10^{-4}	90.1
10^{-3}	10^{-4}	10^{-4}	90.4
10^{-4}	10^{-4}	10^{-4}	92.5
10 ⁻⁴	10^{-2}	10^{-4}	89.8
10^{-4}	10^{-3}	10^{-4}	90.8
10^{-4}	10^{-4}	10^{-4}	92.5
10-4	10^{-4}	10^{-2}	87.3
10^{-4}	10^{-4}	10^{-3}	88.5
10^{-4}	10^{-4}	10^{-4}	92.5

6.4 参数敏感度分析

在本文提出的 OSL-TCSR 算法中, 参数 λ_1 , λ_2 和 λ_3 对物体识别结果有着重要影响, 其中 λ_1 有利 于促进相关性特征重构数据和原始重构数据的联合编码向量矩阵行稀疏化, λ_2 有利于促进联合编码矩 阵整体稀疏化, λ_3 有利于探索多光谱数据的时序特征, 3 种参数不同数值下对物体识别性能影响如表 5 所示. 在该表中, 采用固定变量法分别观察 λ_1 , λ_2 和 λ_3 惩罚参数对设计的 OSL-TCSR 模型的物体识 别性能影响, 以选择最佳物体识别准确率下的参数值. 为了分析 λ_1 的作用, 这里固定 $\lambda_2 = \lambda_3 = 10^{-4}$, 将 λ_1 的值从 10^{-2} 减小到 10^{-4} , 发现 OSL-TCSR 的物体识别结果增大, 这是由于较大的 λ_1 可能导致 相关性特征重构数据和原始重构数据的联合编码向量的各行过于稀疏, 不利于保证两种数据的空间结 构耦合信息; 然后, 为了分析 λ_2 的作用, 这里固定 $\lambda_1 = \lambda_3 = 10^{-4}$, 将 λ_2 的值从 10^{-2} 减小到 10^{-4} , 发现 OSL-TCSR 的物体识别结果也在增大, 较大的 λ_2 可能导致相关性特征重构数据和原始重构数据 的联合编码向量过于稀疏, 使得光谱数据不利于被字典和编码向量进行有效线性组合, 从而较大的 λ_2 值将取得较差的物体识别表现; 最后, 为了分析 λ_3 的作用, 这里固定 $\lambda_1 = \lambda_2 = 10^{-4}$, 当 $\lambda_3 = 10^{-2}$ 时, 发现 OSL-TCSR 的物体识别准确率只有 87.3%, 原因是过大的 λ_3 会导致相邻子编码向量相等. 这 不但没有利用光谱数据内任意相邻两个波段子数据间的相关性特征, 而且认为光谱序列的任意相邻两 个波段子数据相同, 忽略了光谱数据本身的特点. 同时, 过大的 λ_3 导致的相邻子编码向量相等, 也不 利于编码向量稀疏化.

7 结论

本文提出一种新颖的 OSL-TCSR 方法以解决单样本学习下物体识别问题, 它首先通过聚类典型 关联分析将两种光谱的原始数据并行投影到共享子空间中得到相关性特征数据; 其次将原始数据的编 码向量向相关性特征数据的公共编码向量进行二次投影, 并通过不断迭代至收敛后, 重构获得具有充 分耦合联系的多种光谱数据; 然后将重构后的光谱数据导入设计的时序约束稀疏表示模型中, 以考虑 多光谱数据的时序连续性. 最后使用多种分析实验验证, 结果表明了提出的 OSL-TCSR 的有效性. 此 外, OSL-TCSR 还可灵活迁移至多种应用场景, 比如材料识别或纹理识别等, 未来我们还将进一步地 探索考虑数据时序约束特征的主动式多模态融合用于单样本物体识别研究, 使算法能够主动选择不同 模态去融合, 避免不必要的数据冗余.

参考文献 -

- 1 Song A, Han Y, Hu H, et al. A novel texture sensor for fabric texture measurement and classification. IEEE Trans Instrum Meas, 2014, 63: 1739–1747
- 2 Xiong P, He K, Song A, et al. A novel multi-modal one-shot learning method for texture recognition. IEEE Access, 2019, 7: 182538–182547
- 3 Erickson Z, Luskey N, Chernova S, et al. Classification of household materials via spectroscopy. IEEE Robot Autom Lett, 2019, 4: 700–707
- 4 Liu H, Yu Y, Sun F, et al. Visual-tactile fusion for object recognition. IEEE Trans Automat Sci Eng, 2017, 14: 996–1008
- 5 Abderrahmane Z, Ganesh G, Crosnier A, et al. A deep learning framework for tactile recognition of known as well as novel objects. IEEE Trans Ind Inf, 2020, 16: 423–432
- 6 Xi X G, Tang M Y, Zhang Z H, et al. Lower limb motion recognition based on the fusion of sEMG and acceleration signal. Acta Electr Sin, 2017, 45: 2735–2741 [席旭刚, 汤敏彦, 张自豪, 等. 融合表面肌电和加速度信号的下肢运动 模式识别研究. 电子学报, 2017, 45: 2735–2741]
- 7 Xiong P W, Tong X B, Song A G, et al. Robotic cross-modal generative adversarial network based on variational Bayesian Gaussian mixture noise model. Sci Sin Inform, 2021, 51: 104–121 [熊鹏文, 童小宝, 宋爱国, 等. 基于变分 贝叶斯高斯混合噪声模型的机器人跨模态生成对抗网络. 中国科学: 信息科学, 2021, 51: 104–121]
- 8 Strese M, Schuwerk C, Iepure A, et al. Multimodal feature-based surface material classification. IEEE Trans Haptics, 2017, 10: 226–239
- 9 Strese M, Brudermueller L, Kirsch J, et al. Haptic material analysis and classification inspired by human exploratory procedures. IEEE Trans Haptics, 2020, 13: 404–424
- 10 Takahashi K, Tan J. Deep visuo-tactile learning: estimation of tactile properties from images. In: Proceedings of International Conference on Robotics and Automation (ICRA), 2019. 8951–8957
- 11 Erickson Z, Xing E, Srirangam B, et al. Multimodal material classification for robots using spectroscopy and high resolution texture imaging. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020. 10452–10459
- 12 Liu H, Guo D, Sun F. Object recognition using tactile measurements: kernel sparse coding methods. IEEE Trans Instrum Meas, 2016, 65: 656–665
- 13 Roberge J P, Rispal S, Wong T, et al. Unsupervised feature learning for classifying dynamic tactile events using sparse coding. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2016. 2675–2681
- 14 Chen Z, Zuo W, Hu Q, et al. Kernel sparse representation for time series classification. Inf Sci, 2015, 292: 15–26
- 15 Liu H, Sun F. Material identification using tactile perception: a semantics-regularized dictionary learning method. IEEE ASME Trans Mechatron, 2018, 23: 1050–1058
- 16 Xiong P, Tong X, Song A, et al. Robotic multifinger grasping state recognition based on adaptive multikernel dictionary learning. IEEE Trans Instrum Meas, 2022, 71: 1–14
- 17 Liu H, Sun F, Fang B, et al. Cross-modal zero-shot-learning for tactile object recognition. IEEE Trans Syst Man

Cybern Syst, 2020, 50: 2466-2474

- 18 Ren M Y, Liao R J, Fetaya E, et al. Incremental few-shot learning with attention attractor networks. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019. 5275–5285
- 19 Xing C, Rostamzadeh N, Oreshkin B N, et al. Adaptive cross-modal few-shot learning. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019. 4847–4857
- 20 Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017. 4080–4090
- 21 Abderrahmane Z, Ganesh G, Crosnier A, et al. Haptic zero-shot learning: recognition of objects never touched before. Robotics Autonomous Syst, 2018, 105: 11–25
- 22 Rasiwasia N, Mahajan D, Mahadevan V, et al. Cluster canonical correlation analysis. In: Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS), 2014. 823–831
- 23 Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imag Sci, 2009, 2: 183–202
- 24 Mairal J, Bash F, Ponce J, et al. Online dictionary learning for sparse coding. In: Proceedings of the 26th Annual International Conference on Machine Learning (ICML), 2009. 689–696
- 25 Yang M, Zhang L, Yang J, et al. Metaface learning for sparse representation based face recognition. In: Proceedings of IEEE International Conference on Image Processing (ICIP), 2010. 1601–1604
- 26 Efron B, Hastie T, Johnstone I, et al. Least angle regression. Ann Statist, 2004, 32: 407–499
- 27 Liu J, Ji S W, Ye J P. Multi-task feature learning via efficient L2,1-norm minimization. In: Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, 2009. 339–348
- 28 Liu J, Yuan L, Ye J P. An efficient algorithm for a class of fused lasso problems. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010. 323–332
- 29 Mandal D, Biswas S. Generalized coupled dictionary learning approach with applications to cross-modal matching. IEEE Trans Image Process, 2016, 25: 3826–3837
- 30 Zheng J, Qiu H, Sheng W, et al. Kernel group sparse representation classifier via structural and non-convex constraints. Neurocomputing, 2018, 296: 1–11

Object recognition based on one-shot learning with temporally constrained sparse representation

Xiaobao TONG^{1,2}, Pengwen XIONG^{2,3*}, Aiguo SONG⁴ & Peter X. LIU^{2,5}

1. School of Automation, Central South University, Changsha 410083, China;

2. School of Advanced Manufacturing, Nanchang University, Nanchang 330031, China;

3. Department of Automation, University of Science and Technology, Hefei 230026, China;

4. School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China;

5. Department of Systems and Computer Engineering, Carleton University, Ottawa K1S5B6, Canada

* Corresponding author. E-mail: steven.xpw@ncu.edu.cn

Abstract Compared with the contact haptic sensor, the non-contact sensor can avoid the noise generated during direct contact with the object. Therefore, more valuable haptic data can be obtained. However, it is difficult for existing algorithms to achieve high recognition accuracy under the one-shot learning regime for the data acquired by the non-contact sensor. To address this problem, we propose a novel one-shot learning with a temporally constrained sparse representation (OSL-TCSR) method for recognizing 50 objects from 5 materials. First, two original spectroscopy data items, denoted as Lumini and SCiO, are projected into the shared subspace, and the cluster correlation characteristics of the two spectroscopies are calculated using cluster canonical correlation analysis (C-CCA). Then, we use dictionary learning to calculate the coding vectors of the original data and the correlation feature data after the first projection of two spectroscopies, and the coding vector of the original data is projected into the coding vector of the correlation feature data. Further, we reconstruct the original data and the correlation feature data after two mappings to fully couple the two spectroscopies. Finally, the sparse representation method with temporal constraint is designed to calculate respectively the reconstructed original data and the correlation feature data to consider the temporal characteristics of each spectroscopy sequence. We compare it with the latest object recognition methods to prove that OSL-TCSR improves the accuracy of object recognition in the case of one-shot learning. Moreover, OSL-TCSR can also be applied flexibly to other scenarios, such as material identification or texture recognition.

Keywords object recognition, clustering correlation characteristics, one-shot learning, temporally constrained regularization