

基于近端凸差分方法的多层卷积变换学习算法

郭泳澄^{1,2}, 唐健浩^{1,2}, 李珍妮^{1,2†}, 吕俊¹

(1. 广东工业大学 自动化学院, 广东 广州 510006; 2. 广东省物联网信息技术重点实验室, 广东 广州 510006)

摘要: 卷积变换学习(CTL)结合了无监督学习与卷积神经网络的优点, 通过无监督的方式训练卷积核, 是一种新兴的稀疏表示方法. 现有的单层CTL模型仅通过一层稀疏编码, 不仅难以有效提取输入信号的深层语义信息, 并且, 基于 ℓ_0 范数的CTL模型得到的稀疏解虽然稀疏度强, 但它的求解是一个NP-hard难题, 而基于 ℓ_1 范数的CTL模型则存在稀疏度不足和参数过度惩罚的问题. 针对以上问题, 本文提出了一种基于log正则化函数的多层CTL模型(CTL-log): 为了提取输入信号更具鉴别性与丰富语义的稀疏特征, 对单层的CTL模型进行多层拓展, 同时使用稀疏度强, 偏差性小的非凸log正则化函数作为CTL模型的稀疏约束方法. 通过使用近端凸差分方法对模型的非凸优化问题进行优化求解, 开发出基于近端凸差分方法的多层卷积变换学习算法. 实验表明, 本文提出的基于近端凸差分方法的多层卷积变换学习算法所使用的log正则化稀疏约束效果优于现有的CTL模型, 且多层CTL-log的特征提取效果相较于单层取得了提升, 在支持向量机(SVM)分类器的分类精度提升了2个百分点左右.

关键词: 稀疏表示; 卷积变换学习; 近端凸差分方法; log正则化; 特征提取; 机器学习

引用格式: 郭泳澄, 唐健浩, 李珍妮, 等. 基于近端凸差分方法的多层卷积变换学习算法. 控制理论与应用, 2023, 40(11): 2019 – 2027

DOI: 10.7641/CTA.2022.11020

Multi-layer convolutional transform learning algorithm based on proximal difference of convex method

GUO Yong-cheng^{1,2}, TANG Jian-hao^{1,2}, LI Zhen-ni^{1,2†}, LÜ Jun¹

(1. School of Automation, Guangdong University of Technology, Guangzhou Guangdong 510006, China;

2. Guangdong Provincial Key Laboratory of Internet of Things Information Technology, Guangzhou Guangdong 510006, China)

Abstract: Convolutional Transform Learning(CTL) combines the advantages of unsupervised learning and convolutional neural network, learning filters in an unsupervised way, which is a new sparse representation method. However, the existing single-layer CTL model is difficult to effectively extract the deep semantic information of input signals through only one layer of sparse coding. Further more, the ℓ_0 -norm can enforce strong sparsity, but the ℓ_0 -norm-constrained CTL is an NP-hard optimization problem. And the ℓ_1 -norm-constrained CTL presents some drawbacks too, such as its inadequate sparsity and the overpenalization for large elements in the sparse vector. In order to solve these problems of the existing CTL model, This paper presents a multi-layer CTL model based on log regularizer(CTL-log): In order to extract the sparse features of input signals that are more discriminative and rich in semantics, the single-layer CTL model is extended by multiple layers. simultaneously, a log regularizer is used as sparse constraint of CTL model which can not only obtain accurate representations but also yield strong sparsity. Finally, we propose to employ the proximal difference of convex algorithm to efficiently address the nonconvex composite optimization, leading to a proximal difference of convex method based multi-layer convolutional transform learning algorithm. The experimental results demonstrate that the performance of the proposed CTL-log is better than the existing CTL model. And compared with the single-layer CTL-log, the multi-layer CTL-log has a comprehensive improvement in feature extraction, and the classification accuracy of SVM classifier is improved by about 2 percentage points.

Key words: sparse representation; convolutional transform learning; proximal difference of convex algorithm; log regularizer; feature extraction; machine learning

Citation: GUO Yongcheng, TANG Jianhao, LI Zhenni, et al. Multi-layer convolutional transform learning algorithm based on proximal difference of convex method. *Control Theory & Applications*, 2023, 40(11): 2019 – 2027

收稿日期: 2021-10-25; 录用日期: 2022-08-01.

†通信作者. E-mail: lizhenni2012@gmail.com.

本文责任编辑: 柯良军.

国家自然科学基金项目(62073086), 广东省自然科学基金项目(2022A1515011445)资助.

Supported by the National Natural Science Foundation of China (62073086) and the Natural Science Foundation of Guangdong Province (2022A1515011445).

1 引言

近年来,稀疏表示受到越来越多学者的关注与研究,并在信号处理、机器学习、人工智能等领域得到了广泛的应用^[1-5].稀疏表示模型主要有两大类,一种是字典学习(dictionary learning, DL)^[6-8],通过从信号学习一个过完备字典,并用字典中尽可能少的原子的线性组合表示原始信号,达到对原始信号稀疏表示的目的.然而,字典学习模型的稀疏编码问题是一个NP-hard难题^[9],通常需要采用贪婪算法等方法近似求解^[10-11],需要付出较大的计算代价;另一种是变换学习(transform learning, TL)^[12-14],通过从信号学习得到的变换基直接作用于信号提取特征,并对特征进行稀疏约束完成对信号的稀疏表示.字典学习与变换学习都属于传统的基于块的稀疏表示模型,即受限于字典或变换基的大小在处理信号时需先将一完整信号提取为多个小尺寸的重叠信号块,然后对这些信号块分别进行稀疏表示,最后再将得到的稀疏特征重新合成得到完整信号的稀疏特征.这种分块处理的方法导致在稀疏编码过程中忽视了块与块之间的相关性与连续性,并且在处理大规模数据时,需要大量的内存空间用于存储大量的信号块,导致算法的使用受限.

近期,随着卷积神经网络(convolutional neural network, CNN)的快速发展以及在各种类型数据分析取得的成功,同时为了解决传统变换学习模型的缺陷,研究者将卷积神经网络与变换学习相结合,提出了卷积变换学习(convolutional TL, CTL)^[15-18].CTL将卷积神经网络的卷积操作融入了变换学习模型,通过学习一组平移不变的卷积核代替原来的变换基,直接对原始信号进行卷积提取特征.与传统基于块的变换学习模型相比,CTL体现出以下几大优势:1)解决了传统变换学习对信号块之间相关性与连续性的忽视问题,在信号处理过程中充分保留了原始信号的完整内部结构;2)通过对信号整体特征的稀疏约束,避免了稀疏编码的高度冗余;3)直接处理原始信号而不需要截取和存储大量的信号块,适合用于处理大规模数据,拓展了传统变换学习的应用场景.

目前,与传统稀疏表示模型相比兼具高性能与高效率特点的CTL开始逐渐吸引学者们的研究.已有的CTL稀疏表示方法有基于 ℓ_0 范数与基于 ℓ_1 范数稀疏约束的CTL两种,2018年, Maggu 等^[15]首次将卷积操作融合进了变换学习,提出了基于 ℓ_1 范数的CTL,应用于人脸图像的稀疏特征提取与分类任务,并与传统稀疏表示方法进行实验对比,验证了CTL算法的优越性. Chun等^[16-17]使用 ℓ_0 范数作为稀疏约束,提出了卷积解析算子学习(convolutional analysis operator learning, CAOL),在CT图像的重构任务上取得了优秀的效果.随后, Gupta等^[18]用无稀疏约束的CTL训练一维卷积核,搭建了处理股票回归预测任务的多层框架.

然而,在图像稀疏特征的提取方面,目前仅有文献[15]提出的基于 ℓ_1 范数的CTL稀疏特征提取模型. ℓ_0 范数与 ℓ_1 范数是稀疏表示领域常用的稀疏约束方法,其中 ℓ_0 范数是最直接的稀疏约束方法,通过约束特征中的非零元素数量来达到稀疏约束的效果,然而 ℓ_0 范数的求解是一个NP-hard优化难题,随着矩阵维度的增加求解计算量呈指数级增长,在多项式时间内只能找到次优解^[19].因此,常用的一种解决方法是通过采用贪婪算法(如正交匹配追踪^[20]、硬阈值法^[21]等)选择局部最优值对 ℓ_0 范数求近似解,但是此类求取近似解的方法通常难以使稀疏表示模型得到精确度足够理想的稀疏特征.另一种常用的方法则是使用 ℓ_1 范数代替 ℓ_0 范数的方法来求取 ℓ_0 范数的凸松弛解^[22],因为 ℓ_1 范数的求解是一个凸优化问题,求解计算量与矩阵维度为线性关系,通过采用软阈值法可以简单求得封闭解,虽然相比于直接对 ℓ_0 范数求解有效降低了计算复杂度^[20],但是 ℓ_1 范数存在稀疏度不足和大元素过度惩罚的缺陷^[23-24],同样容易导致模型获得的稀疏解精确度不足的问题.近年,有研究者提出使用非凸log正则化函数作为稀疏约束^[25],并且证明了相较于 ℓ_1 范数,log正则化在获取强稀疏解的同时可以有效缓解大参数过度惩罚的问题,有效减小稀疏解的误差.然而,目前还没有关于使用log正则化作为卷积变换学习特征提取稀疏约束方法的研究.根据以往的稀疏表示理论与文献研究表明,稀疏特征的稀疏性与精确性是影响稀疏特征质量的重要因素,在获取高稀疏解的同时减小稀疏特征与原始信号的偏差,是提取高质量高精度稀疏特征的关键,也是提升稀疏表示模型性能的关键.而对于信息丰富复杂的输入信号,单层的稀疏表示模型难以有效提取信号的深层语义信息,获取更具鉴别性的深层稀疏特征的能力受限.因此,研究稀疏性更强、精确度更高的多层CTL稀疏表示方法是其进一步提升和开发的一个关键的研究点.

针对上述问题,本文提出了一种基于log正则化函数的多层CTL稀疏表示模型.通过搭建多层的CTL稀疏特征提取框架,对单层的CTL稀疏特征进行进一步深层次的稀疏编码,从而从输入信号提取更具鉴别性与深层语义信息的稀疏特征.同时,针对现有的基于 ℓ_0 范数与基于 ℓ_1 范数的CTL算法所使用的稀疏约束方法难以有效获取具有高精确度的稀疏特征问题,本文使用稀疏度强,同时偏差性小的非凸log正则化方法^[25]作为CTL的稀疏约束,提升CTL模型的稀疏特征提取的精确度.针对log正则化函数的非凸优化问题,本文通过使用近端凸差分算法(proximal difference of convex algorithm, PDCA)^[26]对模型进行求解优化,开发出基于近端凸差分方法的多层CTL算法稀疏表示算法.

归纳而言,本文的主要贡献有以下几点:

1) 为了提取深层的稀疏特征,提升稀疏特征的提

取质量, 将单层的CTL模型进行多层次拓展, 通过对单层模型的稀疏特征进行进一步稀疏编码, 实现对输入信号提取更具丰富语义与鉴别性的深层稀疏特征.

2) 针对已有的基于 ℓ_0 范数与基于 ℓ_1 范数CTL模型提取稀疏特征存在精确度不足的缺陷, 使用稀疏强、偏差小的log正则化方法作为CTL模型的稀疏约束方法, 提升CTL模型稀疏特征的提取质量.

3) 针对log正则化函数的非凸优化问题, 使用近端凸差分算法对模型目标函数的非凸优化问题进行求解优化, 开发基于近端凸差分方法的多层CTL算法.

2 相关工作

2.1 变换学习

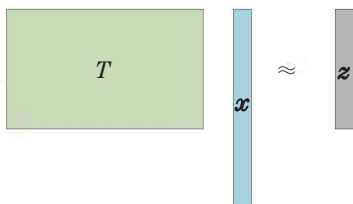
给定一输入信号 $\mathbf{x} \in \mathbb{R}^n$, 变换学习通过学习一组变换基 $T \in \mathbb{R}^{m \times n}$ 对信号进行解析, 得到信号的解析特征, 如图1(a)所示, 其数学过程表示为 $T\mathbf{x} \approx \mathbf{z}$, 其中 $\mathbf{z} \in \mathbb{R}^m$ 为信号 \mathbf{x} 对应的解析特征. 具体地, 变换学习模型算法可用如下函数表示:

$$\min_{T, \mathbf{z}} f(T, \mathbf{z}) + J_z(\mathbf{z}) + J_T(T), \quad (1)$$

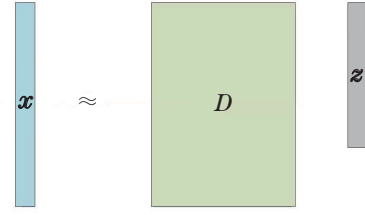
其中: $f(T, \mathbf{z}) = \|T\mathbf{x} - \mathbf{z}\|_2^2$; $J_z(\mathbf{z})$ 为稀疏正则化函数, 通过对解析特征进行稀疏正则化得到信号的稀疏特征; $J_T(T)$ 为变换基组的正则化函数, 用于约束促进模型学习得到的变换基的多样性, 以及防止算法优化过程中出现平凡解(如: $T = 0$).

与传统的字典学习算法将信号表示为字典 D 与其对应的权重向量或矩阵的线性组合 $\mathbf{x} \approx D\mathbf{z}$ 不同, 如图1(b)所示, 变换学习使用变换基组直接对信号进行解析得到信号的解析特征, 并且在解析过程中, 所有的变换基都以均等的权重和使用率参与信号的解析变换过程. 因此变换学习具有更灵活丰富的表示能力. 此外, 在稀疏编码方面, 变换学习的稀疏正则化只需通过对解析特征进行一次简单的阈值操作来求解得到信号的稀疏特征, 是一种更高效的稀疏表示算法.

虽然变换学习相对于字典学习是一种更为高效的稀疏表示模型, 但是它与字典学习都属于基于块的传统稀疏表示模型, 如文章前言所论述, 传统的变换学习模型与字典学习模型也都存在稀疏编码冗余、忽视信号的整体性、丢失特征间的关联性问题, 因此对变换学习模型的开发与应用也受到了这些模型缺陷的限制.



(a) 变换学习原理图



(b) 字典学习原理图

图 1 变换学习与字典学习

Fig. 1 Transform learning and dictionary learning

2.2 卷积变换学习

卷积变换学习将卷积神经网络的卷积特征提取操作引入了变换学习, 如图2所示, 通过学习具有平移不变性的卷积核组(卷积变换基组)直接对原始信号进行卷积提取特征, 克服了第2.1小节所述传统基于块的变换学习算法的缺陷, 卷积变换学习模型算法可由如下函数表示:

$$\min_{\{\mathbf{d}_k\}, \{\mathbf{z}_{l,k}\}} f(\{\mathbf{d}_k\}, \{\mathbf{z}_{l,k}\}) + J_z(\mathbf{z}_{l,k}) + J_D(D), \quad (2)$$

$$f(\{\mathbf{d}_k\}, \{\mathbf{z}_{l,k}\}) = \sum_{l=1}^L \sum_{k=1}^K \frac{1}{2} \|\mathbf{d}_k \times \mathbf{x}_l - \mathbf{z}_{l,k}\|_2^2, \quad (3)$$

其中: $\{\mathbf{x}_l \in \mathbb{C}^N, l = 1, \dots, L\}$ 为一组包含 L 个输入信号的数据集, $D := [\mathbf{d}_1 \dots \mathbf{d}_K]$ 为由 K 个卷积核 $\{\mathbf{d}_k \in \mathbb{C}^R, k = 1, \dots, K\}$ 组成的卷积核组, 每个卷积核都作用于所有输入数据进行卷积特征提取, 即每个输入信号都产生相应的 K 个特征, $\{\mathbf{z}_{l,k} \in \mathbb{C}^N, l = 1, \dots, L, k = 1, \dots, K\}$ 为稀疏特征集合; $J_z(\mathbf{z})$ 为稀疏正则化函数; $J_D(D)$ 为卷积核组的多样性正则化函数.

CTL通过学习卷积核组直接从原始信号提取特征的方式与CNN一致, 可以将其视为一个通过无监督学习的单层结构CNN特征提取模型, 因此, 与神经网络相比, CTL具有如下优势: 1) 在模型学习方面, CTL通过无监督的方式从数据的特征与内部结构学习卷积核组, 在模型的学习过程中不依赖于数据的标签, 而神经网络训练时需要大量的标签数据反向传播学习神经元, 导致了当处理数据标签稀少或者没有标签时神经网络的应用受限; 2) 在数据特征提取方面, 与CNN相比, CTL可以通过对卷积核组实施正则化约束的方式约束卷积核的多样性, 从而减小特征提取的重复冗余, 提升模型提取数据特征的丰富性.

3 基于近端凸差分方法的多层卷变换学习算法

本文通过搭建具有多层稀疏结构的卷积变换学习特征提取框架, 并使用稀疏度强, 同时偏差性小的非凸log正则化函数作为模型的稀疏约束, 提出了一种基于log正则化稀疏约束的多层卷积变换学习模型(多

层CTL-log),如图2所示.首先使用log正则化稀疏约束提升模型提取稀疏特征的精确性,其次,通过构建多层稀疏结构提取更具鉴别性与丰富语义的稀疏特

征.最后,采用近端凸差分方法优化求解log正则化函数的非凸非光滑优化问题,开发基于近端凸差分方法的多层卷变换学习算法.

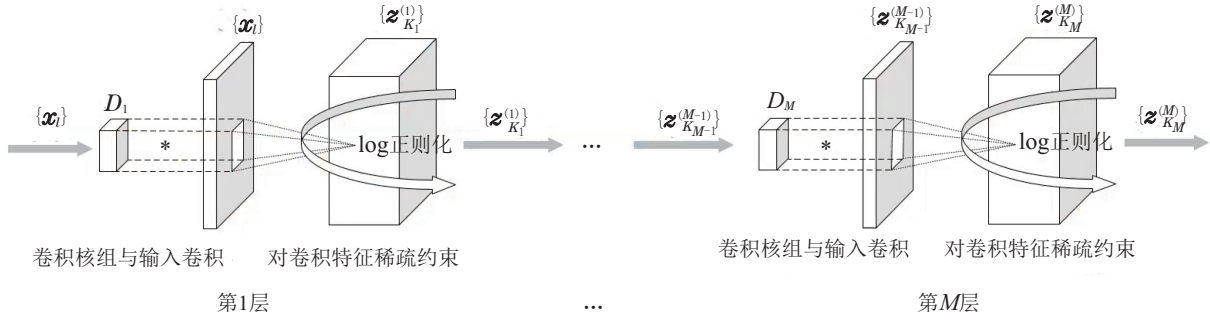


图2 多层卷积变换学习模型

Fig. 2 Multi-layer convolutional transformation learning model

3.1 基于log正则化的多层卷积变换学习模型

将多层卷积变换学习模型层数设为 M , 给定一组输入数据 $\{x_l \in \mathbb{C}^N, l = 1, \dots, L\}$, 其目标是学习 M 组卷积核组 $\{D_1, D_2, \dots, D_M\}$, 对输入数据进行 M 层稀疏特征提取, 第 m 层的卷积核组由 K_m 个卷积核组成, 即 $D_m := [d_1^{(m)} \dots d_{K_m}^{(m)}]$, 其中 $\{d_k^{(m)} \in \mathbb{C}^{R_m}\}$, 每一层的每个输入提取得到的稀疏特征为 $\{z_{K_m}^{(m)} \in \mathbb{C}^N\}$, 并将每一层得到的稀疏特征作为下一层的输入继续进行稀疏编码, 多层卷积变换学习模型的目标函数可表示为

$$\min_{D_m := [d_1^{(m)} \dots d_{K_m}^{(m)}], \{z_k^{(m)}\}} \sum_{l=1}^L \sum_{m=1}^M \sum_{k=1}^{K_m} [J_z(\{z_k^{(m)}\}) + \frac{1}{2} \|d_k^{(m)} * x^{(m)} - z_k^{(m)}\|_2^2], \quad (4)$$

s.t. $J_D(D_m)$,

其中:

$$\{x^{(m)}\} = \{d_{K_{(m-1)}}^{(m-1)}\}, \{x^{(1)}\} = \{x_l\}, \quad (5)$$

$$J_z(\{z_k^{(m)}\}) = \sum_{i=1}^N \lambda \log(1 + \frac{\|z_k^{(m)}\|_i}{\alpha}), \quad (6)$$

$$J_D(D) = \delta_\zeta(D) = \{D : DD^H = \frac{1}{R}I\}. \quad (7)$$

上式中, $J_z(z)$ 是log稀疏正则化函数, 其中参数 λ 是稀疏系数; α 为log函数常系数, 通过调整稀疏系数 λ 平衡稀疏正则化项与近似误差项之间的大小关系; $\delta_\zeta(D)$ 是一个关于到紧集 ζ 投影的指示函数, 通过到这个集合的投影来对卷积核组施加正交性约束^[17].

3.2 近端凸差分方法优化求解

因为log正则化函数具有非凸性和非光滑性, 所以对模型目标函数的求解也是一个非凸非光滑的优化问题. 本小节通过使用近端凸差分分解算法^[23]对模型进行分层求解. PDCA的思想首先通过凸差分分解方法(DCA)将优化函数中的非凸部分分解为凸函数的组

合, 将非凸非光滑优化问题转换为凸非光滑的优化问题, 继而使用近端梯度下降法继续求解, 达到对原非凸非光滑优化问题的优化求解.

如前文所述, 在多层卷积变换学习模型中, 每一层得到的稀疏特征即为下一层的输入, 整个模型目标函数的优化问题可以分解为多个单层的逐层优化问题, 即

$$\min_{D_m := [d_1^{(m)} \dots d_{K_m}^{(m)}], \{z_k^{(m)}\}} F_m(\{d_k^{(m)}\}, \{z_k^{(m)}\}) = \sum_{k=1}^{K_m} [\frac{1}{2} \|d_k^{(m)} * x^{(m)} - z_k^{(m)}\|_2^2 + J_z(\{z_k^{(m)}\})], \quad (8)$$

s.t. $J_D(D_m)$.

每一层的优化问题都是一个非凸非光滑优化问题, 因此使用PDCA对模型每一层的目标函数 F_m 进行逐层优化. 如式(8)所示, 目标函数的优化包含了 $d_k^{(m)}$ 与 $z_k^{(m)}$ 两个变量的求解, 采用交替更新的策略, 即交替固定一个变量, 更新另一个变量的方式, 对目标函数进行迭代优化, 从而将式(8)的优化转换为式(9)与式(10)的交替迭代优化.

$$\min_{\{z_k^{(m)}\}} F_z(\{z_k^{(m)}\}) = \sum_{k=1}^{K_m} [\frac{1}{2} \|d_k^{(m)} * x^{(m)} - z_k^{(m)}\|_2^2 + \sum_{i=1}^N \lambda \log(1 + \frac{\|z_k^{(m)}\|_i}{\alpha})], \quad (9)$$

$$\min_{D_m := [d_1^{(m)} \dots d_{K_m}^{(m)}]} F_d(\{d_k^{(m)}\}) = \sum_{k=1}^{K_m} \frac{1}{2} \|d_k^{(m)} * x^{(m)} - z_k^{(m)}\|_2^2, \quad (10)$$

s.t. $J_D(D_m)$.

3.2.1 变量 $z_k^{(m)}$ 的更新

更新 $z_k^{(m)}$ 时, 固定 $d_k^{(m)}$, 求解优化式(9), 因为稀疏正则化项 J_z 是一个非凸函数, 所以式(9)的优化是一个非凸优化问题. 首先, 采用DCA将 J_z 分解为两个凸函

数组合, 即

$$J_z(\{\mathbf{z}_k^{(m)}\}) = J_z^1(\{\mathbf{z}_k^{(m)}\}) - J_z^2(\{\mathbf{z}_k^{(m)}\}), \quad (11)$$

其中:

$$J_z^1(\{\mathbf{z}_k^{(m)}\}) = \lambda \|\mathbf{z}_k^{(m)}\|_1, \quad (12)$$

$$J_z^2(\{\mathbf{z}_k^{(m)}\}) = \lambda \|\mathbf{z}_k^{(m)}\|_1 - \sum_{i=1}^N \lambda \log(1 + \frac{[\mathbf{z}_k^{(m)}]_i}{\alpha}), \quad (13)$$

通过对式(13)求偏导 $\partial J_z^2(\{\mathbf{z}_k^{(m)}\})$, 得到

$$\{\mathbf{y}_k^{(m)}\} \in \partial J_z^2(\{\mathbf{z}_k^{(m)}\}) = \sum_{i=1}^N \lambda \text{sgn}([\mathbf{z}_k^{(m)}]_i) (1 - \frac{1}{[\mathbf{z}_k^{(m)}]_i + \alpha}), \quad (14)$$

根据DCA, 式(11)可转换为

$$J_z(\{\mathbf{z}_k^{(m)}\}) = J_z^1(\{\mathbf{z}_k^{(m)}\}) - \langle \{\mathbf{y}_k^{(m)}\}, \{\mathbf{z}_k^{(m)}\} \rangle, \quad (15)$$

从而, 式(9)的非凸问题的优化求解可转换为以下的凸优化问题:

$$\min_{\{\mathbf{z}_k^{(m)}\}} F_z(\{\mathbf{z}_k^{(m)}\}) = \sum_{k=1}^{K_m} \frac{1}{2} \|\mathbf{d}_k^{(m)} \times \mathbf{x}^{(m)} - \mathbf{z}_k^{(m)}\|_2^2 + J_z^1(\{\mathbf{z}_k^{(m)}\}) - \langle \{\mathbf{y}_k^{(m)}\}, \{\mathbf{z}_k^{(m)}\} \rangle. \quad (16)$$

由于式(16)中存在部分连续光滑与部分非光滑, 其中非光滑部分为 $J_z^1(\{\mathbf{z}_k^{(m)}\})$, 并将连续光滑部分表示为

$$G_z(\{\mathbf{z}_k^{(m)}\}) = \frac{1}{2} \|\mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{z}_k^{(m)}\|_2^2 - \langle \{\mathbf{y}_k^{(m)}\}, \{\mathbf{z}_k^{(m)}\} \rangle, \quad (17)$$

继而通过使用近端梯度下降法求解以下问题更新变量 $\{\mathbf{z}_k^{(m)}\}$:

$$\{\mathbf{z}_k^{(m)}\} \in \arg \min_{\{\mathbf{z}_k^{(m)}\}} F_z(\{\mathbf{z}_k^{(m)}\}) = \arg \min_{\{\mathbf{z}_k^{(m)}\}} \sum_{k=1}^{K_m} [G_z(\{\mathbf{z}_k^{(m)}\}) + J_z^1(\{\mathbf{z}_k^{(m)}\})]. \quad (18)$$

设在第 t 次迭代时 $\{\mathbf{z}_k^{(m),t}\}$ 的梯度下降更新步长为 η_z , 根据近端梯度下降法, 令 \hat{G}_z 为

$$\hat{G}_z(\{\mathbf{z}_k^{(m)}\}) = G_z(\{\mathbf{z}_k^{(m),t}\}) + \langle \{\mathbf{z}_k^{(m)}\} - \{\mathbf{z}_k^{(m),t}\}, \nabla_{\mathbf{z}_k^{(m)}} G_z(\{\mathbf{z}_k^{(m),t}\}) \rangle, \quad (19)$$

其中:

$$\nabla_{\mathbf{z}_k^{(m)}} G_z = \mathbf{z}_k^{(m)} - \mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{y}_k^{(m)}, \quad (20)$$

则 $\{\mathbf{z}_k^{(m)}\}$ 可由以下公式进行迭代更新:

$$\{\mathbf{z}_k^{(m),t+1}\} \in \arg \min_{\{\mathbf{z}_k^{(m)}\}} \sum_{k=1}^{K_m} [\hat{G}_z(\{\mathbf{z}_k^{(m)}\}) + J_z^1(\{\mathbf{z}_k^{(m)}\})]$$

$$\frac{1}{2\eta_z} \|\{\mathbf{z}_k^{(m)}\} - \{\mathbf{z}_k^{(m),t}\}\|_2^2, \quad (21)$$

从而 $\{\mathbf{z}_k^{(m)}\}$ 的近端梯度下降更新可表示为

$$\{\mathbf{z}_k^{(m),t+1}\} \in \text{Prox}_{\eta_z, J_z^1}(\{\mathbf{z}_k^{(m),t}\} - \eta_z \nabla_{\mathbf{z}_k^{(m)}} G_z(\{\mathbf{z}_k^{(m)}\})), \quad (22)$$

其中: $J_z^1 = \lambda \|\cdot\|_1$; $\text{Prox}_{\eta_z, J_z^1}$ 为近端算子, 即使用近端算子计算软阈值 $\text{Prox}_{\eta_z, \lambda \|\cdot\|_1}$ 为

$$[\text{Prox}_{\eta_z, \lambda \|\cdot\|_1}(\mathbf{z}_k^{(m)})]_i = \begin{cases} [\mathbf{z}_k^{(m)}]_i - \lambda \text{sgn}[\mathbf{z}_k^{(m)}]_i, & [\mathbf{z}_k^{(m)}]_i > \lambda, \\ 0, & \text{其他}. \end{cases} \quad (23)$$

3.2.2 变量 $\mathbf{d}_k^{(m)}$ 的更新

更新 $\mathbf{d}_k^{(m)}$ 时, 固定 $\mathbf{z}_k^{(m)}$, 求解优化式(10), 因为式(10)是一个凸且光滑的优化问题, 所以可直接使用梯度下降法优化, 再进行正交投影 $J_D(D_m)$. 定义

$$G_d(\{\mathbf{d}_k^{(m)}\}) = \frac{1}{2} \|\mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{z}_k^{(m)}\|_2^2, \quad (24)$$

设在第 t 次迭代时 $\{\mathbf{d}_k^{(m),t}\}$ 的梯度下降更新步长为 η_d , 根据梯度下降法, $\{\mathbf{d}_k^{(m),t}\}$ 的梯度下降更新公式为

$$\{\mathbf{d}_k^{(m),t+\frac{1}{2}}\} = \{\mathbf{d}_k^{(m),t}\} - \eta_d \nabla_{\mathbf{d}_k^{(m)}} G_d(\{\mathbf{d}_k^{(m)}\}), \quad (25)$$

其中:

$$\nabla_{\mathbf{d}_k^{(m)}} G_d = (\mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{z}_k^{(m)})^H * \mathbf{x}^{(m)}. \quad (26)$$

然后对卷积核组 $D_m^{t+\frac{1}{2}} := [\mathbf{d}_1^{(m),t+\frac{1}{2}} \dots \mathbf{d}_{K_m}^{(m),t+\frac{1}{2}}]$ 进行正交投影约束. 令 P_{η_d, J_D} 表示在标准正交约束 $DD^H = \frac{1}{R}I$ 上的投影映射, 该投影通过计算线性算子的奇异值分解来实现, 即

$$P_{\eta_d, J_D}(D_m^{t+\frac{1}{2}}) = UI_{R_m \times K_m} V^H, \quad (27)$$

其中 $I_{R_m \times K_m}$ 是主对角线上为单位矩阵的对角矩阵, 通过正交约束投影可以促使列满秩, 避免得到平凡解, 使卷积核组具有紧框架性质. 从而 $\{\mathbf{d}_k^{(m)}\}$ 的迭代更新可表示为

$$\{\mathbf{d}_k^{(m),t+1}\} \in P_{\eta_d, J_D}(\{\mathbf{d}_k^{(m),t}\} - \eta_d \nabla_{\mathbf{d}_k^{(m)}} G_d(\{\mathbf{d}_k^{(m)}\})). \quad (28)$$

3.2.3 模型整体优化训练步骤

综上所述, 基于log稀疏约束的多层卷积变换学习模型训练优化的算法过程具体步骤如表1所示.

算法的计算复杂度主要取决于输入信号维度 N 与个数 L 、卷积核维度 R 与个数 K . 在算法的每次迭代中, 更新稀疏特征的复杂度主要由 $\mathbf{d} * \mathbf{x}$ 决定, 即 $O(RKNL)$; 更新卷积核组的复杂度主要由 $(\mathbf{d} * \mathbf{x})^H * \mathbf{x}$ 决定, 即 $O(RKN^2L)$. 所以, 算法每次迭代的主要复杂度为 $O(RKN^2L)$.

表1 基于log正则化的多层卷积变换学习算法

Table 1 Mutil-layer convolutional transform learning based on log regularizer algorithm

输入: 数据集 $\{\mathbf{x}_l \in \mathbb{C}^N : l = 1, \dots, L\}$, 稀疏参数 λ , 梯度步长 η_z, η_d .

- 1 初始化 $\{D_1, D_2, \dots, D_M\}$, $D_m := [\mathbf{d}_1^{(m)} \dots \mathbf{d}_{K_m}^{(m)}]$, $\{\mathbf{d}_k^{(m)} \in \mathbb{C}^{R_m}\}$.
- 2 **For** $m = 1 : M$
- 3 $t = 0$.
- 4 **重复**
- 5 计算 $\{\mathbf{y}_k^{(m)} \in \partial J_z^2(\{\mathbf{z}_k^{(m)}\})$;
- 6 计算 $\{\mathbf{z}_k^{(m), t+\frac{1}{2}}\} = \{\mathbf{z}_k^{(m), t}\} - \eta_z(\mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{y}_k^{(m)})$;
- 7 计算 $\{\mathbf{z}_k^{(m), t+1}\} = \text{Prox}_{\eta_z, \lambda \|\cdot\|_1}(\{\mathbf{z}_k^{(m), t+\frac{1}{2}}\})$.
- 8 计算 $\{\mathbf{d}_k^{(m), t+\frac{1}{2}}\} = \{\mathbf{d}_k^{(m), t}\} - \eta_d((\mathbf{d}_k^{(m)} * \mathbf{x}^{(m)} - \mathbf{z}_k^{(m)})^H * \mathbf{x}^{(m)})$;
- 9 计算 $\{\mathbf{d}_k^{(m), t+1}\} = P_{\eta_d, J_D}(\{\mathbf{d}_k^{(m), t+\frac{1}{2}}\})$;
- 10 $t = t + 1$;
- 11 **至目标函数收敛**
- 12 $m = m + 1$;
- 13 $\mathbf{x}^{(m)} = \mathbf{z}_{K_{m-1}}^{(m-1)}$.
- 14 **End**

输出: 稀疏特征 $\{\mathbf{z}_{K_M}^{(M)}\}$, $\{D_1, D_2, \dots, D_M\}$.

4 实验分析

本节对本文提出的算法进行实验分析. 首先, 在第4.1节介绍实验所使用的公开数据集的详细信息, 然后, 在第4.2节实验设置介绍实验的流程、对比方法, 以及评估指标, 最后, 在第4.3节呈现实验数据并对结果进行分析讨论.

4.1 实验数据集

为了验证本文所提算法的有效性, 笔者在 YALE^[27], Extended Yale B^[28] 和 AR-Face^[29] 3 个公开人脸数据集上进行了实验, 数据集的具体信息及其处理与划分方法如下:

1) YALE数据集: YALE人脸数据集包含15个人, 每个人在不同表情、姿态和光照下的11张人脸图像, 共165张图片, 每张图片原始大小为 64×64 像素. 实验中, 将图片裁剪缩小为 32×32 像素, 首先, 打乱所有图片, 随机划分70%作为训练集, 30%作为测试集, 作为一组数据划分. 此外, 在每个人的11张图像中随机选择 p 张划分到训练集, 剩下的 $11 - p$ 张划分为测试集, 即训练集的图像数为 $15p$ 张, 测试集的图像数为 $15(11 - p)$ 张, 作为一组YALE- p 数据划分, 通过此规则, 生成新的YALE-2, \dots , YALE-8共7组数据.

2) E-YALE-B数据集: Extended Yale B数据集包含38个人, 每人包含64张在不同光照情况下拍摄的人脸图像, 共有2432张图像, 去除其在官方采样过程被

损坏的18张图像, 每张图像被裁剪为 192×168 像素. 实验中, 将图片裁剪缩小为 48×42 像素, 并打乱所有图片, 随机划分70%作为训练集, 30%作为测试集.

3) AR-Face数据集: AR-Face数据库包含4000多张人脸图像, 来自126个不同的对象(70名男性和56名女性)在不同的面部表情、照明和遮挡(太阳眼镜和围巾)条件下拍摄的正面图像. 实验中, 随机选择其中100名对象(50名男性和50名女性), 每人26张图像, 共2600张图像进行实验, 每张图像裁剪为540像素(27×20), 并随机选择2000张图像作为训练集, 其余600张图像作为测试集.

4.2 实验设置

为了验证基于log正则化卷积变换学习相较于已有基于 l_0 范数卷积变换学习^[16-17]与基于 l_1 范数卷积变换学习^[15]对于提取稀疏特征质量的提升, 以及多层CTL-log特征提取相较于单层CTL-log的有效性, 本文对于在以上数据集的实验进行了以下设置:

1) 首先进行单层的CTL实验对比, 在相同的模型结构下比较CTL-log, CTL- l_0 , CTL- l_1 的特征提取效果, 验证log正则化对于提升CTL模型稀疏特征提取的效果; 然后将单层CTL-log拓展为双层, 验证双层CTL-log相对于单层的效果.

2) 实验数据集划分为训练集和测试集, 训练集用于训练CTL模型的卷积核组完成CTL模型的训练, 并提取训练集稀疏特征; 再通过训练完成的CTL模型提取测试集的稀疏特征.

3) 实验的评估指标为分类精度, 将训练集的稀疏特征输入分类器进行训练学习, 再输入测试集稀疏特征进行分类测试, 从分类精度高低体现模型提取的稀疏特征质量优劣, 即模型稀疏特征提取效果的好坏. 使用的分类器为支持向量机(support vector machine, SVM).

4.3 实验结果

如上一小节实验设置所述, 本小节首先进行单层实验不同稀疏约束方法的对比, 然后进行单层模型与双层模型的对比, 验证多层模型的有效性.

4.3.1 单层对比

如上一节实验设置所述, 此小节首先在单层模型结构下, 对比不同的稀疏约束下CTL模型(CTL-log, CTL- l_0 , CTL- l_1)的特征提取效果. 为了保证实验对比合理公平, 3种CTL模型除了稀疏约束项以外, 模型的其它结构均保持一致. 卷积核数 $K = 5$, 卷积核尺寸 $R = 5 \times 5$, 即卷积核规模为 $5 \times 5 \times 5$.

第3.2.3节中算法表的输入参数 λ, η_d, η_z 设置如下: 在数据集YALE, YALE-2~YALE-8的实验中, 梯度步长 $\eta_d = 4e - 13$, $\eta_z = 3e - 2$, 稀疏参数 λ 的最终寻优区间为 $(1e - 5, 3e - 4)$; 在数据集AR-Face的实验

中, 梯度步长 $\eta_d = 5e - 14$, $\eta_z = 8e - 2$, 稀疏参数 λ 的最终寻优区间为(0.05, 0.2); 在数据集 E-YALE-B 的实验中, 梯度步长 $\eta_d = 4e - 12$, $\eta_z = 6e - 2$, 稀疏参数 λ 的最终寻优区间为(0.01, 0.02).

实验中, 通过改变模型的稀疏约束项系数调整模型进行多次实验, 分别得到模型在各个实验数据集下的最优分类结果进行比较, 实验结果如表2所示.

表 2 单层提取特征分类结果

Table 2 Classification results of single layer extraction features

| 数据集 | 算法 | | |
|----------|---------------|---------------|---------------|
| | CTL- l_0 | CTL- l_1 | CTL-log |
| YALE | 0.9000 | 0.9200 | 0.9400 |
| YALE-2 | 0.7704 | 0.7704 | 0.7704 |
| YALE-3 | 0.8417 | 0.8167 | 0.8250 |
| YALE-4 | 0.8571 | 0.8667 | 0.8762 |
| YALE-5 | 0.8778 | 0.8889 | 0.9000 |
| YALE-6 | 0.8800 | 0.8933 | 0.9200 |
| YALE-7 | 0.9000 | 0.9167 | 0.9333 |
| YALE-8 | 0.9333 | 0.9556 | 0.9556 |
| E-YALE-B | 0.9945 | 0.9945 | 0.9917 |
| AR-Face | 0.9550 | 0.9467 | 0.9633 |

根据表2实验结果显示, 在所有10项数据集的分类对比实验中, CTL-log在其中8项上取得了最高或并列最高的分类准确率, 表明了对于输入信号稀疏特征的提取, CTL-log的效果显然优于CTL- l_0 与CTL- l_1 , 验证了与已有的CTL- l_0 与CTL- l_1 相比, 引入了具有误差小且稀疏性强的log稀疏正则化的CTL-log模型具有从输入信号数据提取更高质量稀疏特征的能力. 突

出使用log正则化函数作为CTL模型的稀疏约束的优势.

在运行时间方面, 以YALE数据集的实验为例进行分析, 在实验中CTL- l_0 , CTL- l_1 与CTL-log在相同的迭代次数条件下的运行时间分别为110 s, 82 s, 87 s, CTL- l_1 与CTL-log的运行时间明显快于CTL- l_0 , 而CTL-log相比于CTL- l_1 要略微慢一些, 这是因为相比于CTL- l_1 , CTL-log在每次迭代中都要进行一次凸差分(difference of convex, DC)分解.

4.3.2 双层拓展

本小节笔者将上一小节在稀疏特征提取效果的优势得到充分验证的CTL-log拓展为双层: 将单层稀疏编码得到稀疏特征进一步进行深一层稀疏编码, 在保持上一小节的单层CTL-log模型结构不变的条件下, 将其作为双层CTL-log的首层, 进行相同的实验策略, 对比在第2层不同的卷积核组规模下($3 \times 3 \times 3$, $5 \times 5 \times 5$, $7 \times 7 \times 7$ 与 $9 \times 9 \times 9$)稀疏特征的提取效果, 寻找确定模型第2层的最佳结构, 并与单层结构CTL-log对比, 验证双层CTL-log稀疏编码得到的稀疏特征质量相比于单层的提升效果. 其中, 第3.2.3节中算法表的最终输入参数在第1层的设置与第4.3.1节保持一致, 第2层的参数设置具体如下: 在数据集YALE, YALE-2 ~ YALE-8的实验中, 梯度步长 $\eta_d = 6e - 18$, $\eta_z = 1.7e - 2$, 稀疏参数 λ 的最终寻优区间为($5e - 6$, $3e - 5$); 在数据集AR-Face的实验中, 梯度步长 $\eta_d = 4e - 12$, $\eta_z = 5e - 2$, 稀疏参数 λ 的最终寻优区间为($5e - 3$, $1e - 2$); 在数据集E-YALE-B的实验中, 梯度步长 $\eta_d = 3e - 11$, $\eta_z = 3e - 2$, 稀疏参数 λ 的最终寻优区间为(0.01, 0.02). 实验结果如表3所示.

表 3 双层拓展实验结果

Table 3 Experimental results of two-layer expansion of CTL-log

| 数据集 | 模型结构 | | | | |
|----------|--------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|
| | 单层 $5 \times 5 \times 5$ | $5 \times 5 \times 5 + 3 \times 3 \times 3$ | $5 \times 5 \times 5 + 5 \times 5 \times 5$ | $5 \times 5 \times 5 + 7 \times 7 \times 7$ | $5 \times 5 \times 5 + 9 \times 9 \times 9$ |
| YALE | 0.9400 | 0.9400 | 0.9600 | 0.9400 | 0.9400 |
| YALE-2 | 0.7704 | 0.7778 | 0.7778 | 0.7704 | 0.7852 |
| YALE-3 | 0.8250 | 0.8500 | 0.8500 | 0.8250 | 0.8083 |
| YALE-4 | 0.8762 | 0.8857 | 0.9048 | 0.8762 | 0.8571 |
| YALE-5 | 0.9000 | 0.9000 | 0.9111 | 0.9111 | 0.9000 |
| YALE-6 | 0.9200 | 0.9333 | 0.9467 | 0.9333 | 0.9333 |
| YALE-7 | 0.9333 | 0.9333 | 0.9500 | 0.9667 | 0.9667 |
| YALE-8 | 0.9556 | 0.9778 | 0.9556 | 0.9778 | 0.9556 |
| E-YALE-B | 0.9917 | 0.9876 | 0.9945 | 0.9917 | 0.9903 |
| AR-Face | 0.9633 | 0.9933 | 0.9933 | 0.9967 | 0.9867 |

表3实验结果表明, 在单层结构与双层结构的方面, 在10项数据集的实验结果对比中, 所有最优的结果(包含并列最优)全部在双层CTL-log中产生, 表明

通过对单层CTL-log稀疏编码得到的稀疏特征的进一步稀疏编码提升了稀疏特征的提取质量; 在不同双层结构的内部对比方面, 在第2层的卷积核组规模分别

为 $3 \times 3 \times 3 \times 3$, $5 \times 5 \times 5$, $7 \times 7 \times 7$, $9 \times 9 \times 9$ 这4种不同结构下的双层CTL-log在10项数据集实验中得到最优结果的次数分别为2, 6, 4, 2, 从取得最优结果次数的角度看, 第2层卷积核组规模为 $5 \times 5 \times 5$ 时双层CTL-log的效果最好.

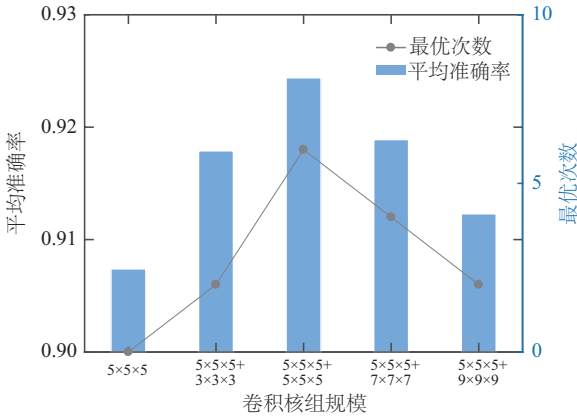


图3 不同CTL结构的分类平均准确率与最优次数对比

Fig. 3 Comparison of average accuracy and optimal times of different CTL structures

此外, 通过计算10项实验结果准确率的整体平均值, 从平均准确率角度对比各模型结构的效果, 如图3所示, 同样可以看出双层CTL-log的平均准确率也都高于单层CTL-log, 并且双层CTL-log卷积核组规模为 $5 \times 5 \times 5 + 5 \times 5 \times 5$ 时平均准确率与最优次数均为最高, 说明通过对CTL-log的双层拓展, 提取得了输

入信号更具鉴别性与丰富语义的稀疏特征, 提升了模型的稀疏特征提取质量, 且当第2层卷积核规模为 $5 \times 5 \times 5$ 时提升效果最好.

综上分析, 可以得出在保持第1层结构不变的情况下, 双层CTL-log第2层卷积核组规模设置为 $5 \times 5 \times 5$ 时, 模型对于稀疏特征的提取整体效果是最好的, 因此将第2层的卷积核规模设定为 $5 \times 5 \times 5$. CTL- l_0 , CTL- l_1 , CTL-log与双层CTL-log的最终对比如表4所示.

如前文所述, 与已有的CTL- l_0 , CTL- l_1 相比, 引入log正则化函数稀疏约束的CTL-log能够提取得到具有更高精确度, 质量更高的稀疏特征, 并且通过将CTL-log拓展为双层CTL-log对输入信号进一步的稀疏编码提取了输入信号更具鉴别性与丰富语义的深层次稀疏特征, 进而提升了模型提取稀疏特征的质量.

此外, 为了验证本文提出算法当下的有效性, 除了与CTL模型的实验对比外, 还与当下最新的人脸识别算法之一, 基于判别投影和表征的分类(discriminative projection and representation-based classification, DP-RC)^[30]算法进行对比, 实验结果如表4所示. 从DP-RC与双层CTL-log的实验结果对比可以看出, 在10项分类数据集分类结果的对比中, 双层CTL-log除了在YALE, YALE-8与DP-RC持平, AR-Face略低于DP-RC外, 在其余的7项结果都要明显优于DP-RC, 表明了与DP-RC算法相比, 本文的双层CTL-log算法具有更好人脸识别分类效果, 验证了本文提出算法在当下的有效性.

表4 不同CTL模型及DP-RC分类结果对比

Table 4 Comparison of classification results of different CTL models and DP-RC algorithm

| 数据集 | 算法 | | | | |
|----------|---------------|---------------|---------------|---------------|---------------|
| | CTL- l_0 | CTL- l_1 | CTL-log | 双层CTL-log | DP-RC |
| YALE | 0.9000 | 0.9200 | 0.9400 | 0.9600 | 0.9600 |
| YALE-2 | 0.7704 | 0.7704 | 0.7704 | 0.7778 | 0.6741 |
| YALE-3 | 0.8417 | 0.8167 | 0.8250 | 0.8500 | 0.7667 |
| YALE-4 | 0.8571 | 0.8667 | 0.8762 | 0.9048 | 0.8476 |
| YALE-5 | 0.8778 | 0.8889 | 0.9000 | 0.9111 | 0.8778 |
| YALE-6 | 0.8800 | 0.8933 | 0.9200 | 0.9467 | 0.9200 |
| YALE-7 | 0.9000 | 0.9167 | 0.9333 | 0.9500 | 0.9333 |
| YALE-8 | 0.9333 | 0.9556 | 0.9556 | 0.9556 | 0.9556 |
| E-YALE-B | 0.9945 | 0.9945 | 0.9917 | 0.9945 | 0.9931 |
| AR-Face | 0.9550 | 0.9467 | 0.9633 | 0.9933 | 0.9983 |

5 结论

针对已有的卷积变换学习模型存在的问题, 本文提出了基于近端凸差分方法的多层卷变换学习算法, 该算法通过对单层卷积变换学习模型的稀疏特征进行进一步的稀疏编码, 提取输入信号更具丰富语义与鉴别性的深层稀疏特征, 并且使用稀疏度强, 偏差性

小的非凸log正则化函数作为卷积变换学习模型的稀疏约束方法, 通过使用近端凸差分算法对模型的非凸优化问题进行优化求解, 构建基于log正则化函数稀疏约束的多层卷积变换学习特征提取框架.

在公开的人脸图像数据集的特征提取实验结果表明, 本文提出的多层log稀疏约束卷积变换学习模型的

人脸图像稀疏特征提取质量优于已有的基于 ℓ_0 范数与基于 ℓ_1 范数的卷积变化学习算法, 并且基于 \log 正则化函数的多层卷积变换学习通过层次的拓展有效提升了稀疏特征的提取质量, 取得了更高的分类精度。

在未来的工作中, 笔者将对提出的算法进行进一步深度层次的开发, 并拓展应用到心电信号数据的分类。

参考文献:

- [1] CHEN Yongbin, LI Yuanqing. Pattern localization of functional connectivity in parkinson's disease based on sparse representation. *Control Theory & Applications*, 2017, 34(6): 843 – 848.
(陈勇斌, 李远清. 基于稀疏表示的帕金森功能连接模式定位. 控制理论与应用, 2017, 34(6): 843 – 848.)
- [2] XIE K, ZHOU G, YANG J, et al. Eliminating the permutation ambiguity of convolutive blind source separation by using coupled frequency bins. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(2): 589 – 599.
- [3] GUO Junfeng, LI Yuliang. Sparse representation of robot image based on dictionary learning algorithm. *Acta Automatica Sinica*, 2020, 46(4): 820 – 830.
(郭俊峰, 李育亮. 基于学习字典的机器人图像稀疏表示方法. 自动化学报, 2020, 46(4): 820 – 830.)
- [4] WEN S, ZHENG Y, LU F. A sparse representation based joint demosaicing method for single-chip polarized color sensor. *IEEE Transactions on Image Processing*, 2021, 30: 4171 – 4182.
- [5] HUANG Y, YANG G, WANG K, et al. Learning joint and specific patterns: A unified sparse representation for off-the-person ECG biometric recognition. *IEEE Transactions on Information Forensics and Security*, 2021, 16: 147 – 160.
- [6] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311 – 4322.
- [7] HUANG K, WU Y, WANG C, et al. A projective and discriminative dictionary learning for high-dimensional process monitoring with industrial applications. *IEEE Transactions on Industrial Informatics*, 2021, 17(1): 558 – 568.
- [8] RONG Y, XIONG S, GAO Y. Double graph regularized double dictionary learning for image classification. *IEEE Transactions on Image Processing*, 2020, 29: 7707 – 7721.
- [9] HOCHBA, DORIT S. Approximation algorithms for NP-hard problems. *ACM SIGACT News*, 1997, 28(2): 40 – 52.
- [10] REZAIIFAR Y C P R, KRISHNAPRASAD P S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. *Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, CA, USA: IEEE, 1993: 40 – 44.
- [11] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 1998, 20(1): 33 – 61.
- [12] NAGPAL S, SINGH M, SINGH R, et al. Discriminative shared transform learning for sketch to image matching. *Pattern Recognition*, 2021, 114: 107815.
- [13] RAVISHANKAR S, BRESLER Y. Learning sparsifying transforms. *IEEE Transactions on Signal Processing*, 2013, 61(5): 1072 – 1086.
- [14] SHARMA S, MAJUMDAR A. Unsupervised detection of non-technical losses via recursive transform learning. *IEEE Transactions on Power Delivery*, 2021, 36(2): 1241 – 1244.
- [15] MAGGU J, CHOUZENOUS E, CHIERCHIA G, et al. Convolutional transform learning. *International Conference on Neural Information Processing, Siem Reap*, 2018, 25: 162 – 174.
- [16] CHUN I Y, HONG D, ADCOCK B, et al. Convolutional analysis operator learning: Dependence on training data. *IEEE Signal Processing Letters*, 2019, 26(8): 1137 – 1141.
- [17] CHUN I Y, FESSLER J A. Convolutional analysis operator learning: Acceleration and convergence. *IEEE Transactions on Image Processing*, 2019, 29(1): 2108 – 2122.
- [18] GUPTA P, MAGGU J, MAJUMDAR A, et al. DeConFuse: A deep convolutional transform based unsupervised fusion framework. *EURASIP Journal on Advances in Signal Processing*, 2020, 26(1): 1 – 32.
- [19] BAO C, HUI J, YU Q, et al. L_0 norm based dictionary learning by proximal methods with global convergence. *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014: 3858 – 3865.
- [20] ZENG J, LIN S, XU Z. Sparse regularization: Convergence of iterative jumping thresholding algorithm. *IEEE Transactions on Signal Processing*, 2016, 64(19): 5106 – 5118.
- [21] FAN J, LI R. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 2001, 96(456): 1348 – 1360.
- [22] NIU L, ZHOU R, TIAN Y, et al. Nonsmooth penalized clustering via ℓ_p regularized sparse regression. *IEEE Transactions on Cybernetics*, 2016, 47(6): 1423 – 1433.
- [23] GONG P, ZHANG C, LU Z, et al. A general iterative shrinkage and thresholding algorithm for non-convex regularized optimization problems. *International Conference on Machine Learning*, 2013, 28: 37 – 45.
- [24] WEN J, ZHOU Z, CHEN H. An optimal condition for the block orthogonal matching pursuit algorithm. *IEEE Access*, 2018, 6: 38179 – 38185.
- [25] MALIOUTOV D, ARAVKIN A. Iterative log thresholding. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy: IEEE, 2014: 7198 – 7202.
- [26] RAKOTOMAMONJY A, FLAMARY R, GASSO G. Dc proximal newton for nonconvex optimization problems. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 27(3): 636 – 647.
- [27] BELHUMEUR P, HESPANHA J, KRIEGMAN D. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711 – 720.
- [28] LEE K. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(5): 684 – 698.
- [29] MARTINEZ A, BENAVENTE R. The AR Face Database: CVC Technical Report, 1998.
- [30] DENG K, PENG Z, ZHU W. A discriminative projection and representation-based classification framework for face recognition. *SIAM Journal on Imaging Sciences*, 2020, 13(3): 1446 – 1466.

作者简介:

郭泳澄 硕士研究生, 目前研究方向为稀疏表示、信号处理, E-mail: 2111904303@mail2.gdut.edu.cn;

唐健浩 学士本科生, 目前研究方向为稀疏表示、图像处理, E-mail: TangJianhao2020@outlook.com;

李珍妮 副教授, 硕士生导师, 目前研究方向为信号处理、稀疏表示、深度学习, E-mail: lizhenni2012@gmail.com;

吕俊 副研究员, 硕士生导师, 目前研究方向为多模态多标签机器学习、微弱病理特征波提取, E-mail: ryjl@163.com.