

# 联合多模态表示的人脸识别方法研究

张杜娟, 陈飞, 丁莉

(西安医学院 卫生管理学院, 陕西 西安 710021)

**摘要:** 针对人脸图像识别问题, 提出联合多模态表示的方法。采用二维经验模态分解(BEMD)对人脸图像进行分解, 获得多层次二维内蕴模函数(BIMF), 作为人脸的多模态表示。在分类阶段, 采用多任务压缩感知模型(MTCS)对多层次BIMF进行表征, 求解不同BIMF的稀疏系数矢量, 进而计算不同类别对于当前测试样本的重构误差, 根据重构误差判断样本所属类别。通过在ORL和Yale-B人脸图像数据集上开展实验, 所提方法平均识别率分别达到75.42%和94.07%, 且对噪声干扰具有良好的稳健性。

**关键词:** 人脸识别; 二维经验模态分解; 多模态表示; 多任务压缩感知

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 1674-6236(2021)24-0175-04

**DOI:** 10.14022/j.issn1674-6236.2021.24.036

## Research on face recognition method based on multimodal representations

ZHANG Dujuan, CHEN Fei, DING Li

(School of Health Services Management, Xi'an Medical University, Xi'an 710021, China)

**Abstract:** For face image recognition problem, a method using multimodal representations is proposed. The Bidimensional Empirical Mode Decomposition (BEMD) is applied to decompose the face image with multi-layer Bidimensional Intrinsic Mode Functions(BIMF), which are used as the multimodal representations for the face image. In the classification phase, the multi-task compressive sensing (MTCS) is adopted to represent the multi-layer BIMF to obtain their individual sparse coefficient vectors. Further, the reconstruction errors of different classes for the test image can be calculated. According to the reconstruction errors, the face label can be decided. Experiments are conducted on the ORL and Yale-B face image datasets, the proposed method could achieve average recognition rates of 75.42% and 94.07%, respectively, also with good noise robustness.

**Keywords:** face recognition; bidimensional empirical mode decomposition; multimodal representations; multi-task compressive sensing

人脸识别技术在日常生活中运用广泛, 成为现代社会管理的有力手段<sup>[1-3]</sup>。人脸识别技术自提出以来, 长期成为模式识别和人工智能领域的热点问题。基于人脸图像的识别算法主要从特征提取和分类决策两个角度入手。前者是从高维度图像中抽取与人脸特性直接相关的描述方式; 后者根据提取特

征进行未知样本所属类别的判断。当前, 应用于人脸识别的特征种类多样, 具有代表性的是像素分布类和图像分析类。像素分布类主要考察人脸图像的灰度分布规律, 通过注入主成分分析(Principal Component Analysis, PCA)等手段获得有效的低维度特征矢量<sup>[4-5]</sup>。图像分析类则通过图像细节的处理获得人脸的点、线、面以及局部问题特征<sup>[6-8]</sup>。对比而

**收稿日期:** 2020-11-07 **稿件编号:** 202011035

**基金项目:** 陕西省教育厅专项科研计划项目(20JK0886); 陕西省大学生创新创业训练计划项目(S202111840040); 西安医学院大学生创新创业训练计划项目(121521040)

**作者简介:** 张杜娟(1981—), 女, 陕西西安人, 硕士, 讲师。研究方向: 深度学习技术、图像技术等。

言,前一类特征具有提取效率高、特征形式规则的优势;后一类特征则具有物理意义清晰、解译性更强的优势。根据获取的特征,分类决策阶段针对性选取或设计判断机制。人脸识别算法的决策机制见证了模式识别领域分类器的发展进步。早期的分类器主要有近邻分类器、支持向量机(Support Vector Machine, SVM)<sup>[9]</sup>、神经网络等。后续,基于压缩感知的稀疏表示分类器(Sparse Representation-based Classification, SRC)在人脸识别中得到了验证<sup>[10-11]</sup>。近年来,以卷积神经网络(Convolutional Neural Network, CNN)为代表的深度模型成为人脸识别最为常用的工具<sup>[12-13]</sup>,并取得了较好的研究成果。

文中着眼人脸识别中的特征提取和分类决策问题,提出联合多模态表示的新方法。采用二维经验模态分解(Bidimensional Empirical Mode Decomposition, BEMD)对人脸图像进行分解<sup>[14-15]</sup>,获得多层次二维内蕴模函数(Bidimensional Intrinsic Mode Function, BIMF)。相比单一的原始图像,多层次的BIMF可梯次化描述人脸的全局、局部、细节等特性,因此区分不同类别的能力更强。对于分解得到的多层次BIMF,文中选用多任务压缩感知模型(Multi-Task Compressive Sensing, MTCS)<sup>[15-16]</sup>对其进行表征。MTCS在贝叶斯理论框架下求解多任务稀疏表示问题,在考察不同任务之间关联的条件下获得概率意义上的最优解。因此,通过结合BEMD和MTCS在特征提取和分类决策方面的优势,所提方法可以取得更为稳健的人脸识别性能。基于ORL和Yale-B数据集开展实验分析,结果验证了方法的可靠性。

## 1 多模态表示

BEMD是EMD在二维领域的扩展,可直接用于图像分解<sup>[14-15]</sup>。对于输入图像 $I(x, y)$ ,经过BEMD分解后得到如下形式:

$$I(x, y) = \sum_{j=1}^L c_j(x, y) + r_L(x, y) \quad (1)$$

其中,  $c_j(x, y), j=1, 2, \dots, L$  表示分解得到的多层次BIMF;  $r_L(x, y)$  为分解残差,多为图像中的噪声杂波分量等。

文献[15]详细论证了BEMD的实施流程和分解结果。现有研究成果表明了BEMD在图像分析解译中的有效性。经过分解得到的多模态表示,即多层

次BIMF,可更为有效地反映其中被描述对象的多角度性质。文中将BEMD用于人脸图像分解,将其获得的BIMF用于后续的分类决策过程,获得更具鉴别力的特征描述。

## 2 多任务压缩感知分类决策

多任务压缩感知是压缩感知理论的拓展运用,主要用于解决多个关联稀疏表示问题的联合求解。其基本思想是在贝叶斯框架下通过对不同成分进行有效统计建模获得具有最大后验的结果。对于 $Y = \{y_i\}_{i=1, \dots, L}$  共有 $L$ 个不同的稀疏表示任务,它们独立描述的情况如下:

$$y_i = A_i x_i + n_i \quad (2)$$

其中,  $A_i \in R^{N_i \times N}$  表示第 $i$ 稀疏表示问题的字典;  $x_i$  为待求解的稀疏系数矢量;  $n_i$  为噪声分量,采用均值为零、方差为 $\xi_0$ 的高斯分布进行建模。

在此基础上,  $y_i$  的概率密度函数可表达如下:

$$p(y_i | x_i, \xi_0) = (2\pi/\xi_0)^{-N_i/2} \exp\left(-\frac{\xi_0}{2} \|y_i - A_i x_i\|_2^2\right) \quad (3)$$

在利用这 $L$ 个稀疏表示问题存在内在关联的条件下,进一步转化为如下形式:

$$p(y_i | \xi, \xi_0) = \prod_{j=1}^N N(x_{ij} | 0, \xi_j^{-1} \xi_0^{-1}) \quad (4)$$

其中,  $x_{ij}$  是  $x_i$  中的第 $j$ 个元素;  $\xi = [\xi_1, \dots, \xi_N]^T$ 。

MTCS通过最大化边缘概率密度可获得 $\xi$ 的最佳估计,形式如下:

$$\begin{aligned} L(\xi) &= \sum_{i=1}^L \log(p(y_i | \xi)) \\ &= -\frac{1}{2} \sum_{i=1}^L [(N_i + 2a) \log(y_i^T B_i^{-1} y_i + 2b) + \log|B_i|] + c \end{aligned} \quad (5)$$

其中,  $B_i = I + A_i A_i^T$ ;  $I$  表示单位矩阵;  $c$  为常数。

MTCS采用迭代更新的策略不断获得概率意义上的近似解,求得不同稀疏表示问题的表示系数的均值和方差。最终结果在概率意义上具有最优性,对于后续的分类决策具有重要的意义。

在MTCS求解得到各个稀疏表示问题的系数矢量的前提下,可分别利用不同类别训练样本对测试样本进行重构,通过比较重构误差进行类别决策:

$$\text{identity}(Y) = \min_k \sum_{i=1}^L \|y_i - A_i^{(k)} x_i^{(k)}\| \quad (6)$$

其中,  $\mathbf{A}_i^{(k)}$  为第  $i$  个稀疏表示问题中与第  $k$  类关联的局部字典;  $\mathbf{x}_i^{(k)}$  为对应的系数矢量。

文中采用 MTCS 对人脸图像分解得到的多模态表示进行表征。从 MTCS 的原理可以看出,其在概率框架下获得具有最大后验的最优估计结果。因此,其可充分考察不同 BIMF 的独立性和关联性,获得最佳的重构结果。图 1 给出了所提方法的具体实施流程。训练集在 BEMD 的分解下构建多层次的 BIMF 字典,测试样本相应分解得到对应层次的 BIMF。在 MTCS 的表征下,各个训练类别均会对测试样本产生一个重构误差,最终通过比较不同类别的误差大小即可判定测试样本的所属人脸类别。具体地,文中根据重复实验结果,选取分解结果中的前四层(即图 1 中  $M=4$ ) BIMF 代入识别算法,MTCS 的最佳参数通过自寻优获得。

### 3 实验与分析

#### 3.1 实验数据集

ORL 和 Yale-B 人脸图像数据集是当前开展人脸识别算法验证的权威数据集,具有很强的代表性。其中,ORL 数据集收集了 40 张的人脸图像,每

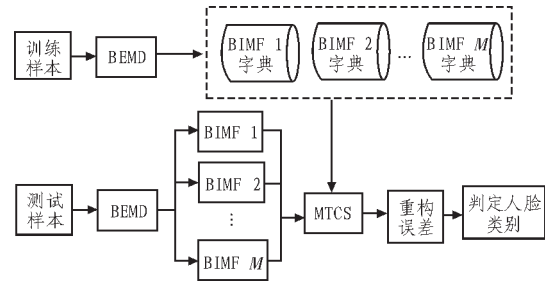


图1 联合多模态表示的人脸识别流程

人包含 10 个样本。Yale-B 数据集收集了 10 张的人脸图像,每人 45 张。图 2 给出了两个数据集的部分人脸样本示意。具体开展实验时,在 ORL 数据集中,随机选取每人 5 个样本作为训练集,其余 5 个样本用于测试。在 Yale-B 数据集中随机选取每人 20 个样本作为训练集,其余 25 个样本用于测试。

实验过程选择已有的一些人脸识别方法与文中方法作对比研究,主要包括文献[9]中提出的基于灰狼算法优化的 SVM(记为优化 SVM)、文献[3]设计的改进 SRC(记为改进 SRC)以及文献[13]采用的多模态深度特征(记为 DeepFace)。实验共分为 3 个部分,分别为原始样本、噪声样本和遮挡样本,可据此全面测试所提方法的综合识别性能。

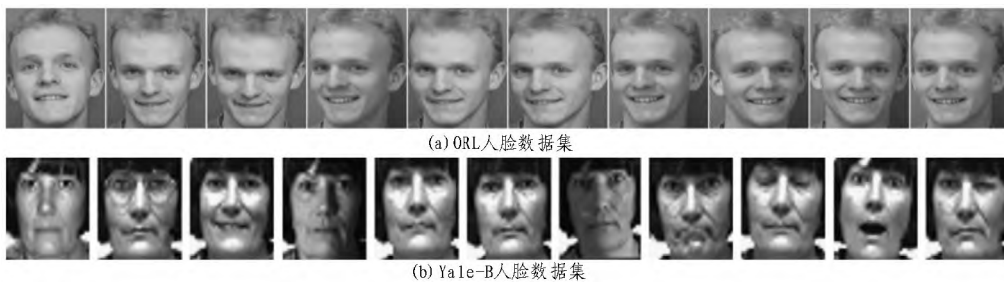


图2 人脸图像样本示意

#### 3.2 结果与分析

##### 3.2.1 原始样本

该实验基于 ORL 和 Yale-B 原始样本开展测试,按照 3.1 节中的实验对比设置,分别测试各类方法的性能,统计如表 1 所示。横向对比 ORL 和 Yale-B 两个数据集,明显发现 ORL 的整体识别难度更大,所有方法的平均识别率均相对较低,主要由于图像样本的获取条件相对苛刻。纵向对比各类方法,所提方法在两个数据集上均保持最佳性能,体现了其有效性。相比优化 SVM 和改进 SRC 方法,该文除了选用合适的分类机制,还在特征提取方面作了充分考虑。多层次 BIMF 相比单纯利用原始人

脸图像可提供更为充分的识别信息。和 DeepFace 相比,文中通过结合 BEMD 和 MTCS 两者的优势进一步提升了最终性能。

表1 原始样本识别结果对比

方法	平均识别率(%)	
	ORL	Yale-B
所提方法	75.42	94.07
优化 SVM	67.82	90.74
改进 SRC	68.24	91.02
DeepFace	66.52	90.16

##### 3.2.2 噪声样本

上述实验中的原始样本多在合作条件下获取,



总体噪声水平较低,不能很好地检验各方法在噪声干扰条件下的性能。而实际情况下,人脸图像获取可能会有各种干扰存在,为此,实验首先通过模拟噪声条件的方式获得不同信噪比(SNR)下的测试集,进而考核各类方法的噪声稳健性。图3给出了全部四类方法在不同信噪比条件下的平均识别率,从中可以看出噪声干扰对于人脸识别性能的显著影响。所提方法通过结合 BEMD 分解的多模态表示和 MTCS 的优势,在各个噪声水平均保持了最佳性能。BEMD 分解过程实际也在进行噪声残差的优化剔除,MTCS 在贝叶斯框架下进行稀疏表示问题求解,充分考虑了噪声分布的影响。因此,两者的有效结合可以提升识别算法对于噪声干扰的稳健性。

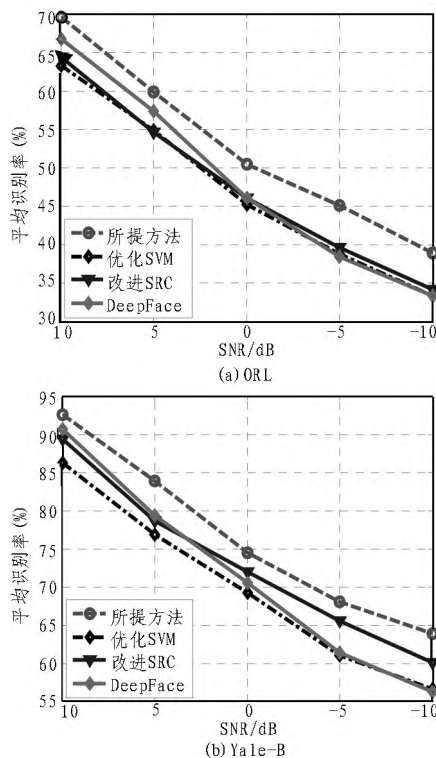


图3 噪声样本识别结果对比

### 3.2.3 遮挡样本

如文献[17]所述,现实中采集得到的人脸图像往往伴随遮挡的情形。为此,下面开展遮挡样本的测试。首先,按照文献[17]的思路采用局部遮挡的方式构造不同遮挡比例条件下的测试样本。然后,分别在不同的遮挡比例下对各类方法进行测试,统计识别结果如图4所示。遮挡情形的出现直接影响了所有4类方法的性能,其中DeepFace最为剧烈。所提方法在各个遮挡比例下均保持最佳性能,体现其遮

挡稳健性。基于 BEMD 分解的多模态表示通过互补关联描述可在遮挡情形下保持高鉴别力。MTCS 则可以利用多模态表示的优势获得最终的稳健决策结果。

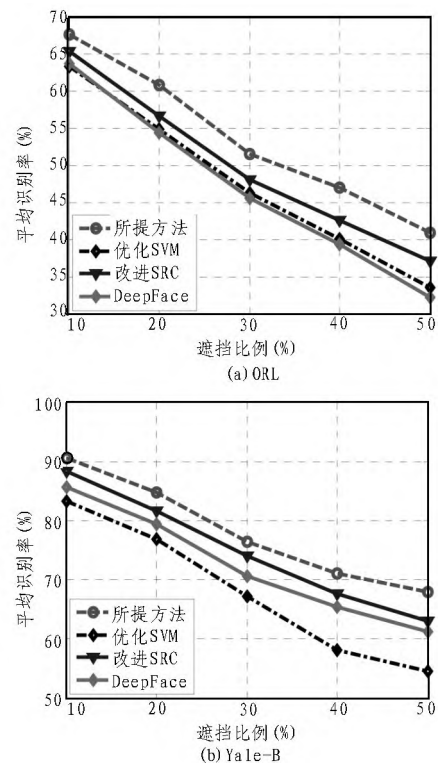


图4 遮挡样本识别结果对比

## 4 结论

文中将 BEMD 和 MTCS 联合运用于人脸识别问题,通过 MTCS 对多模态表示的有效处理获得更为稳健的人脸识别结果。特征提取和分类决策有效性的提升使得文中方法具有性能优势。实验依托 ORL 和 Yale-B 数据集在原始样本、噪声样本和遮挡样本上开展,通过与现有几类方法的对比分析,可以证实所提方法的性能优势。

### 参考文献:

- [1] 廖延娜,马超.基于稀疏表示的人脸识别系统设计与实现[J].电子设计工程,2016,24(17):153-155.
- [2] 张一夫,陈天飞,张建松.基于树莓派的在线人脸识别考勤系统[J].电子设计工程,2019,27(22):152-156.
- [3] 刘嘎琼.改进稀疏表示的人脸识别在高校管理中的应用[J].计算机与数字工程,2018,46(11):2333-2336.

(下转第183页)

- [5] 司海青,林松,马胜,等.无人机遥感提取输电线路走廊信息[J].地理空间信息,2018,16(3):33-35,42.
- [6] 解宇阳,王彬,姚扬,等.基于无人机激光雷达遥感的亚热带常绿阔叶林群落垂直结构分析[J].生态学报,2020,40(3):196-207.
- [7] 郑雪娜,沈国杰,马鹏宇,等.基于ai智能图像的工控火电现场爆炸风险评估[J].中国安全生产科学技术,2018,14(12):135-138.
- [8] 冯敏,罗旺,余磊,等.适用于无人机巡检图像的输电线路螺栓检测方法[J].电力科学与技术学报,2018,33(4):135-140.
- [9] 熊小萍,许爽,蒙登越,等.基于Faster R-CNN的输电线路缺陷识别模型研究[J].自动化与仪器仪表,2020(3):7-12.
- [10] 黄俊波,蔡澍雨,贾永祥.基于虚拟现实技术在无人机巡检输电线路的研究[J].电子测量技术,2020,43(8):44-47.
- [11] 杨昱,孙昌雯,王大伟,等.基于无人机前端和SSD算法的输电线路部件检测模型对比研究[J].太原理工大学学报,2020,51(2):212-219.
- [12] 辛守庭,杜轻,徐超.用于旋翼无人机避障的数据预处理系统设计[J].飞行力学,2018,36(5):85-88,94.
- [13] 李俊鹏,张继伟,余艳稳.基于电网行业巡检无人机智能地面监控系统研究[J].信息技术,2020,44(6):134-138.
- [14] 顾德喜,李贵霖,时祥凯,等.基于百度AI开放平台的人脸识别门禁系统[J].单片机与嵌入式系统应用,2019,19(6):33-36.
- [15] 金立军,王恒,王文华,等.基于机器视觉和激光测距的输电线故障定位[J].同济大学学报(自然科学版),2018,46(12):1745-1753.
- [16] 罗隆福,李冬,钟杭.基于改进RRT的无人机电力杆塔巡检路径规划[J].湖南大学学报(自然科学版),2018,45(10):80-86.
- [17] 王舶仲,蒋毅舟,文超,等.基于生成对抗网络的隔离开关分合位置判别方法研究及应用[J].智慧电力,2019,47(10):77-84.
- [18] 马静怡,崔昊杨.基于改进RPN网络的电力设备图像识别方法研究[J].供用电,2020,37(1):57-61,72.
- [19] 彭福先,张玮,祝晓军,等.基于激光点云精确定位的输电线路无人机自主巡检系统研究[J].智慧电力,2019,47(7):117-122.
- (上接第178页)
- [4] 徐竞泽,吴作宏,徐岩,等.融合PCA、LDA和SVM算法的人脸识别[J].计算机工程与应用,2019,55(18):34-37.
- [5] 杨梅芳,石义龙.基于2DPCA+PCA与SVM的人脸识别[J].信息技术,2018(2):32-36.
- [6] 刘斌,米强,徐岩.LBP和MB-LBP加权融合的人脸识别[J].计算机工程与设计,2018,39(2):551-556.
- [7] 何云,吴怀宇,钟锐.基于多种LBP特征集成学习的人脸识别[J].计算机应用研究,2018,35(1):292-295.
- [8] 马振,刘凤连,汪日伟.基于子模式下LBP-HOG特征融合的单样本人脸识别方法[J].光电子·激光,2019,30(12):1309-1316.
- [9] 冯璋,裴东,王维.基于改进灰狼算法优化支持向量机的人脸识别[J].计算机工程与科学,2019,41(6):1057-1063.
- [10] 孔英会,尹紫薇,车麟麟.基于稀疏学习与显著性的人脸识别方法研究[J].计算机工程,2018,44(2):271-276.
- [11] 酒明远,陈恩庆,齐林,等.基于多核学习的多阶分数阶傅里叶变换域人脸识别[J].光电工程,2018,45(6):134-142.
- [12] 胡亚洲,周亚丽,张奇志.基于深度学习的人脸识别算法研究[J].计算机应用研究,2020,37(5):1432-1436.
- [13] Ding C, Tao D. Robust face recognition via multimodal deep face representation[J]. IEEE Transactions on Multimedia, 2015, 17(11):2049-2058.
- [14] Qin Y, Qiao L, Wang Q F, et al. Bidimensional empirical model decomposition method for image processing in sensing system[J]. Computers and Electrical Engineering, 2018(68):215-224.
- [15] 张宏武,康凯.结合二维内蕴模函数和贝叶斯多任务学习的SAR目标识别[J].电讯技术,2020,60(4):372-377.
- [16] Ji S, Dunson DB, Carin L. Multitask compressive sensing[J]. IEEE Transactions on Signal Processing, 2009, 57(1):92-106.
- [17] 董艳花,张树美,赵俊莉.有遮挡人脸识别方法综述[J].计算机工程与应用,2020,56(9):1-12.