

基于旋转主方向梯度直方图特征的 判别稀疏图映射算法*

童莹^{1)2)†} 沈越泓¹⁾ 魏以民¹⁾

1) (中国人民解放军陆军工程大学, 通信工程学院, 南京 210007)

2) (南京工程学院, 信息与通信工程学院, 南京 211167)

(2019年2月20日收到; 2019年7月29日收到修改稿)

非约束环境下采集的人脸图像复杂多变, 将其直接作为字典原子用于稀疏表示分类 (sparse representation based classification, SRC), 识别效果不理想. 针对该问题, 本文提出一种基于旋转主方向梯度直方图特征的判别稀疏图映射 (discriminative sparse graph embedding based on histogram of rotated principal orientation gradients, DSGE-HRPOG) 算法, 用于构建类内紧凑、类间分离的低维判别特征字典, 提高稀疏表示分类准确性. 首先, 采用旋转主方向梯度直方图 (histogram of rotated principal orientation gradients, HRPOG) 特征算子提取非约束人脸图像的多尺度多方向梯度特征, 有效去除外界干扰和像素间冗余信息, 构建稳定、鉴别的 HRPOG 特征字典; 其次, 引入判别稀疏图映射 (discriminative sparse graph embedding, DSGE) 算法, 以类内重构散度最小、类间重构散度最大为目标计算特征字典的最佳低维投影矩阵, 进一步增强低维特征字典的判别性、紧致性; 最后, 提出投影矩阵和稀疏重构关系交替迭代优化算法, 将维数约简过程伴随在稀疏图构建过程中, 使分类效果更理想. 在 AR, Extended Yale B, LFW 和 PubFig 这 4 个数据库上进行大量实验, 验证了本文算法在实验环境数据库和真实环境数据库上的有效性.

关键词: 非约束人脸识别, 流形学习, 稀疏保持投影, 方向梯度直方图

PACS: 42.30.Sy, 42.30.Tz, 42.30.-d

DOI: 10.7498/aps.68.20190224

1 引言

随着互联网技术和人工智能技术的飞速发展, 实时采集人脸图像进行身份认证、安全监控和信息提取已成为机器视觉领域最热门的研究课题之一^[1,2]. 然而, 非控环境中采集的人脸图像受光照、姿态、表情、遮挡、年龄等多种因素影响, 导致人脸图像复杂多变, 严重抑制了非约束人脸识别的准确性.

2009年, Wright等^[3]受生物视觉稀疏性启发, 首次提出稀疏表示分类模型 (sparse representation

based classification, SRC), 并将其成功应用于非约束人脸识别中, 很好地解决了人脸图像中光照、遮挡、噪声等因素的影响. 分析其原因, 关键在于稀疏表示分类中并不需要精确的特征表达, 只需遵循“简约性”原则, 从字典中选择“有限”原子对待测样本进行线性表示, 并期望达到所选“原子”仅与待测样本同类. 由此可见, 构建学习紧致、判别的字典在稀疏表示分类中具有极其重要的作用. Yang等^[4]提出 Fisher 判别字典学习 (fisher discrimination dictionary learning, FDDL) 方法, 将 Fisher 判别准则引入结构字典学习中, 使得同

* 国家自然科学基金 (批准号: 61703201, KYTYJJG206), 江苏省自然科学基金 (批准号: BK20170765) 和南京工程学院青年创新基金 (批准号: CKJB201602) 资助的课题.

† 通信作者. E-mail: tongying@njit.edu.cn

类字典原子具有相似性而异类字典原子具有差异性. Vu 等^[5]在此基础上引入低秩共享字典 (low-rank shared dictionary, LRSD) 用于表示人脸本质特征, 与表示人脸差异性的类特定字典 (class-specific dictionary, CSD) 相结合, 进一步提高了字典空间的判别性. 为使学习得到的字典具有更丰富的语义信息, Babae 等^[6]摒弃了二值标签信息而采用相对属性信息, 提出相对属性引导字典学习 (relative attribute guided dictionary learning, RAGDL) 算法. 同时, 文献^[7–10]在字典学习方面也进行了深入研究.

分析发现, 字典学习 (dictionary learning, DL) 方法虽然在构建学习紧致、判别字典空间上取得一定效果, 但仍受 2 个方面限制: 1) 稀疏表示分类^[3]要求字典原子的维度远远小于字典原子个数, 从而保证解空间存在最优稀疏解, 然而, 以 $M \times N$ 像素大小的样本图像作为字典原子构建字典空间, 其维度远大于样本个数, 容易导致出现“小样本”问题; 2) 字典学习方法通过增加各种约束条件优化字典满足类内紧凑和类间分离, 然而, 由于原始样本图像的高维性、复杂性和冗余性, 使得优化效果大打折扣, 严重抑制了待测样本稀疏表示分类的准确性.

鉴于此, 大量学者提出采用降维 (dimensionality reduction, DR) 方法克服上述不足, 一方面通过降低字典空间维度可以避免“小样本”问题, 减少稀疏编码计算复杂度, 加速搜寻最优解; 另一方面增加投影约束条件, 使高维样本字典在低维子空间更具紧致性、判别性, 稀疏表示更准确, 这已成为稀疏表示分类中必不可少的步骤.

以主成分分析 (principle components analysis, PCA)^[11]、线性鉴别分析 (linear discriminant analysis, LDA)^[12] 为代表的经典降维方法被首先用于稀疏表示分类模型中^[3,4,13], 但受数据全局线性分布条件的限制, 当处理复杂扭曲的非线性分布的真实数据时, 这类方法效果不佳. 鉴于此, 广大学者提出采用基于核的方法^[7,14–16]和流形学习方法^[17–21]进行数据非线性降维. 核方法的本质是利用核函数映射样本到高维空间进行线性划分, 但起关键作用的核函数的选择机制和意义并不明确. 因此, 以等距映射 (isometric mapping, ISOMAP)^[17]、局部线性嵌入 (locally linear embedding, LLE)^[18]、拉普拉斯特征映射 (Laplacian eigenmaps, LE)^[19]、

平行向量场嵌入 (parallel vector field embedding, PFE)^[20]、测地线距离学习 (geodesic distance function learning, GDL)^[21] 等为代表的非线性流形学习方法更受青睐. 这类方法符合人眼视觉感知机制, 可以有效挖掘出隐藏在错综复杂的高维数据中的低维流形本质结构. He 等^[22,23]针对 LE 和 LLE 的隐式映射问题进行改进, 在不改变原有目标函数的基础上增加线性约束, 提出具有显式映射函数的局部保持投影 (locality preserving projections, LPP) 和近邻保持嵌入 (neighborhood preserving embedding, NPE) 算法, 有效克服了样本外扩展问题 (out-of-sample extension problem)^[24], 使得流形学习方法在图像、视频、文本等高维、海量、真实数据中的应用成为可能. 随后, 大量改进算法被陆续提出, 例如, 指数判别局部保持投影 (exponential discriminant locality preserving projection, EDLPP)^[25]、二维判别局部保持投影 (two-dimensional discriminant LPP, 2DDLPP)^[26]、双向二维近邻保持判别嵌入 (bilateral two-dimensional neighborhood preserving discriminant embedding, B2DNPDE)^[27]、快速正交局部保持投影 (fast and orthogonal LPP, FOLPP)^[28]、监督近邻保持嵌入 (supervised NPE, SNPE)^[29] 等.

Yan 等^[30]对 LPP, NPE 及改进算法的实现原理进行分析, 将它们归纳到图映射框架 (graph embedding framework, GEF) 中. 分析发现, GEF 的核心思想是通过构建高维样本近邻分布图, 寻找使低维子空间数据仍保持高维样本分布特性的最优投影矩阵. 因此, 在 GEF 中如何构建一个准确的高维样本近邻分布图是这类方法的关键. 近邻图的构建包含 2 个关键步骤: 图顶点选择和图边权分配. 传统方法多采用 k-nearest^[31] 或 ϵ -ball^[23] 方法搜索样本近邻点, 并用热核函数^[31] 或逆欧式距离^[32] 计算近邻点之间的权值. 然而, 实际应用中真实数据的分布是未知的、复杂的, 选择合适的近邻点非常困难, 并且样本之间距离测度的可区分性也会随着样本维数的增加而减弱, 这在一定程度上抑制了基于 GEF 的降维方法的广泛应用.

2010 年, Qiao 等^[33]受信号稀疏表示 (sparse representation, SR)^[3–5] 的启发, 首次将稀疏技术引入图映射框架中, 提出稀疏保持投影 (sparsity preserving projections, SPP) 算法. 这是一种不以距离测度为区分, 而以保持样本间稀疏重构关系为

目标的全新降维方法. 它通过优化求解每一个样本在全局超完备字典下的线性稀疏表示, 自适应得到与该样本近邻的少量重构样本以及样本之间的重构权值, 有效克服了传统图映射方法中预定义近邻图的弊端. 这种以 SPP 为代表的自适应稀疏图映射 (sparse graph embedding, SGE) 方法极大提高了真实复杂数据低维投影的准确性, 受到广大学者的高度关注.

Lai 等^[34]将稀疏表示技术引入 LLE 框架中, 提出稀疏线性映射 (sparse linear embedding, SLE) 算法, 并扩展到核空间, 在 3 个人脸库和 2 个目标库上验证了其降维有效性, 尤其是在小样本情况下表现更突出. Yin 等^[35]将稀疏表示与 LPP 相结合, 提出局部稀疏保持投影 (local sparsity preserving projection, LSPP) 算法, 并成功应用于生物大数据的降维与识别中. Zhang 等^[36]结合近邻图和稀疏图优点, 提出稀疏邻域保持投影 (sparsity and neighborhood preserving projections, SNPP) 算法. 分析发现, 基于全局样本的稀疏关系约束会破坏样本局部结构, 导致样本间的近邻关系描述并不准确. Lu 等^[37]在 SPP 基础上引入类别标签, 提出鉴别稀疏邻域保持嵌入 (discriminant sparsity neighborhood preserving embedding, DSNPE) 算法, 分别考虑待测样本与同类样本的重构关系, 以及异类样本对待测样本的重构影响, 有效提高了待测样本稀疏近邻图的准确性, 相比 SPP 算法识别性能大大提升. 同样地, Wei 等^[38]提出加权判别稀疏保持映射 (weighted discriminative sparsity preserving embedding, WDSPE) 算法, Lou 等^[39]则将稀疏表示技术引入 LPP 框架中, 提出正则图稀疏判别分析 (graph regularized sparsity discriminant analysis, GRSDA) 算法. 鉴于传统方法中多将降维和分类作为 2 个独立步骤分别建模, 存在降维效果与分类器不匹配的问题, Yang 等^[40]提出稀疏表示分类引导判别投影 (sparse representation classifier steered discriminative projection, SRC-DP) 算法, 以同类重构残差最小、异类重构残差最大为目标, 联合优化投影矩阵以及稀疏表示分类系数, 使 SRC 分类器在投影子空间获得最佳性能. 类似地, Zheng 等^[41]提出特征加权组稀疏判别投影 (feature weighted group sparse classification steered discriminative projection, FWGSDP) 算法, 并采用迭代重约束稀疏编码优化

方法对该模型进行高效求解. Gao 等^[42]基于 Fisher 判别准则引导数据降维和分类. Zhang 等^[43]将 SRC-DP 进一步扩展到核空间, 从多个不同角度验证算法有效性. 可见, 由于稀疏表示技术的引入, 以 SPP 为代表的稀疏图映射方法具有非常好的图自适应性和噪声鲁棒性, 将其与 SRC 分类器相结合, 用于构建低维、紧致、判别字典可以显著提高分类器的精度和速度, 是近年来值得研究和探索的热点方向之一.

受基于统计局部特征的鲁棒核表示 (robust kernel representation with statistical local features, RKR-SLF)^[44]以及基于 Gabor 特征的鲁棒表示分类^[45](gabor feature based robust representation and classification, RRC-GF) 等算法的启发, 我们发现, 采用 LBP, Gabor 等鉴别力和鲁棒性较强的局部特征代替传统的冗余数据样本构建特征字典, 可以显著提高稀疏表示分类的准确性. 因此, 本文考虑将局部特征算子与降维方法相结合, 提出一种基于旋转主方向梯度直方图特征的判别稀疏图映射 (discriminative sparse graph embedding based on histogram of rotated principal orientation gradients, DSGE-HRPOG) 算法, 实现流程如图 1 所示. 首先, 从特征描述角度设计一种旋转主方向梯度直方图 (histogram of rotated principal orientation gradients, HRPOG) 特征算子, 相比传统方向梯度直方图 (histogram of oriented gradients, HOG) 特征算子, 它从多个尺度、多个方向捕捉人脸纹理的梯度变化信息, 可以更准确提取出人脸本质特征, 显著提高了特征字典的鉴别性和鲁棒性. 其次, 从降维角度提出判别稀疏图映射 (discriminative sparse graph embedding, DSGE) 算法, 在 DSNPE 算法^[37]的基础上, 通过引入类内、类间紧凑度约束, 有效增强了待测样本与同类非近邻样本的重构关系, 并削弱了异类伪近邻样本的重构影响; 同时, 在低维投影阶段又增加了全局约束因子, 利用样本全局分布中隐含的鉴别信息使 HRPOG 特征字典在低维子空间更判别、更紧致. 最后, 受 SRC-DP^[40]的启发, 提出一种投影矩阵和稀疏重构关系交替的迭代优化算法, 将维数约简过程伴随在稀疏图构建过程中, 使分类效果更理想. 经实验环境下采集的人脸数据库 (AR 和 Extended Yale B) 和真实环境下采集的人脸数据库 (LFW 和 PubFig) 验证, DSGE-HRPOG 算法可以有效提取这些高度扭曲

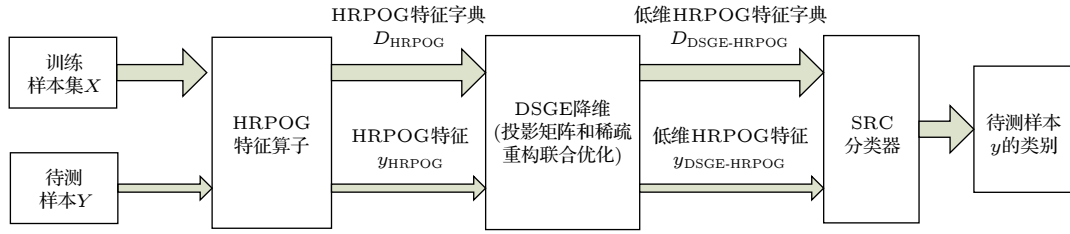


图 1 本文算法的实现流程

Fig. 1. Flow chart of the proposed algorithm.

的人脸数据的低维流形本质结构, 大大增强特征字典的紧致性、判别性和鲁棒性, 使 SRC 分类性能更突出.

2 HRPOG 特征算子

2.1 方向梯度直方图 (histogram of oriented gradients, HOG)

2005 年, Dalal 和 Triggs^[46] 提出方向梯度直方图 (histogram of oriented gradients, HOG) 用于行人检测, 其基本原理是统计图像中各个像素点的梯度幅值和梯度方向, 用于表征目标形态、轮廓变化等信息. 相比 LBP, Gabor 等局部纹理特征算子, HOG 算子提取的方向梯度变化信息更符合人眼视觉感知特性, 受光照、旋转、噪声等外界干扰影响更小, 提取目标特征效果更佳、更鲁棒^[47,48]. 实现步骤如下:

- 1) 将图像 $f(x, y)$ 划分成大小相同的无重叠单元格 (cell);
- 2) 将 $0-180^\circ$ 分为 9 个区间, 平均每 20° 为一个区间, 统计每个单元格中所有像素的梯度幅值和梯度方向, 将属于同一梯度方向区间的像素点的梯度幅值按照权重累加, 组成该单元格的 9 维方向梯度直方图;
- 3) 将 4 个相邻的无重叠的单元格组合成一个块 (block), 级联块中单元格的方向梯度直方图构成该块的方向梯度直方图特征;
- 4) 级联所有块的方向梯度直方图, 得到整幅图像的 HOG 特征.

2.2 HRPOG 特征算子

分析发现, 在 HOG 算子实现过程中, 图像 $f(x, y)$ 中任意像素的梯度 $\vec{\nabla}f(x, y)$ 是一个具有大小和方向的矢量, 由 $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 顺序表示:

$$\vec{\nabla}f(x, y) = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]^T. \quad (1)$$

可见, $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 计算是否准确直接决定了 HOG 特征的代表能力. 在传统 HOG 算子中, $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 计算公式如下:

$$\begin{cases} \frac{\partial f}{\partial x} = h_x * f(x, y) = f(x + 1, y) - f(x - 1, y), \\ \frac{\partial f}{\partial y} = h_y * f(x, y) = f(x, y + 1) - f(x, y - 1), \end{cases} \quad (2)$$

这里, h_x 和 h_y 为梯度卷积模板, 记为 $h_x = [-1 \ 0 \ 1]$, $h_y = [-1 \ 0 \ 1]^T$.

可以看出, 依据 (2) 式计算图像梯度受 2 个方面限制: 1) $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 仅考虑了中心像素点周围 2 个像素点的灰度变化情况, 提取的图像标量场的微观信息并不丰富, 也不准确; 2) $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 仅描述了图像在水平方向和垂直方向上的灰度变化程度, 其他方向上的灰度变化并未考虑. 鉴于此, 本文提出一种 HRPOG 特征算子, 一方面, 增加梯度卷积模板 h_x 和 h_y 的计算范围, 充分考虑中心像素点周围邻域像素点的灰度变化情况; 另一方面, 旋转模板 h_x 和 h_y , 使 HRPOG 算子可以提取图像在不同方向上的梯度特征, 具有旋转不变性. 具体设计细节如下:

首先, 将梯度卷积模板 h_x 和 h_y 设计为 3×3 大小 (图 2). 模板中, 主方向 (箭头指向) 上的卷积权值设为 3, 其余卷积权值设为 1 或 -1, 使模板中所有卷积权值的和为零. 可以看出, 相比传统 HOG 算子, HRPOG 算子的梯度卷积模板既充分考虑了中心像素点周围邻域像素点的灰度变化情况, 同时, 卷积权值设置也可以有效增强图像在主方向 (箭头指向) 上的梯度变化信息. 这里, 为了区分 HRPOG 算子中设置不同尺度的梯度卷积模板的情况, 定义 3×3 尺度下的 HRPOG 算子为

3-HRPOG. 依据图 2 得到 3-HRPOG 算子的梯度矢量 $(\vec{\nabla}f)_{3\text{-HRPOG}}$ 计算公式如下:

$$(\vec{\nabla}f)_{3\text{-HRPOG}} = \begin{bmatrix} \left(\frac{\partial f}{\partial x}\right)_{3\text{-HRPOG}} \\ \left(\frac{\partial f}{\partial y}\right)_{3\text{-HRPOG}} \end{bmatrix}, \quad (3)$$

其中,

$$\left\{ \begin{array}{l} \left(\frac{\partial f}{\partial x}\right)_{3\text{-HRPOG}} = h_x * f(x, y) = f(x-1, y+1) \\ \quad + 3f(x, y+1) + f(x+1, y+1) \\ \quad - f(x-1, y-1) - f(x-1, y) \\ \quad - f(x, y-1) - f(x+1, y-1) \\ \quad - f(x+1, y) \\ \left(\frac{\partial f}{\partial y}\right)_{3\text{-HRPOG}} = h_y * f(x, y) = f(x+1, y-1) \\ \quad + 3f(x+1, y) + f(x+1, y+1) \\ \quad - f(x, y-1) - f(x-1, y-1) \\ \quad - f(x-1, y) - f(x-1, y+1) \\ \quad - f(x, y+1) \end{array} \right. \quad (4)$$

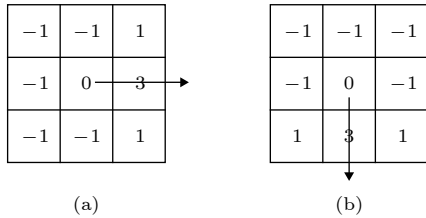


图 2 3-HRPOG 算子的梯度卷积模板示意图 (a) h_x 模板; (b) h_y 模板

Fig. 2. Gradient convolution masks of 3-HRPOG feature descriptor: (a) h_x mask; (b) h_y mask.

其次, 在图 2 的基础上, 依据梯度卷积模板的主方向 (箭头指向), 将 h_x 和 h_y 分别沿逆时针以 45° 为间隔旋转一周, 依次得到 8 组旋转梯度卷积模板, 记为 3-Group(i) ($i = 1, 2, \dots, 8$), 如图 3 所示. 采用 (5) 式分别计算图像在每组旋转梯度卷积模板主方向上的梯度矢量, 统计得到对应的主方向梯度直方图. 将 8 组主方向梯度直方图依次累加, 得到最终的 3-HRPOG 特征向量.

$$\left\{ \begin{array}{l} \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=0^\circ} = [h_{x1} * f(x, y), h_{y1} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=45^\circ} = [h_{x2} * f(x, y), h_{y2} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=90^\circ} = [h_{x3} * f(x, y), h_{y3} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=135^\circ} = [h_{x4} * f(x, y), h_{y4} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=180^\circ} = [h_{x5} * f(x, y), h_{y5} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=225^\circ} = [h_{x6} * f(x, y), h_{y6} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=270^\circ} = [h_{x7} * f(x, y), h_{y7} * f(x, y)]^T \\ \vec{\nabla}f_{3\text{-HRPOG}} |_{\theta=315^\circ} = [h_{x8} * f(x, y), h_{y8} * f(x, y)]^T. \end{array} \right. \quad (5)$$

图 4 进一步说明了 3-HRPOG 算子的旋转不变性. 图 4(a) 是一幅二值图像 (紫色表示 0, 绿色表示 1), 将其沿逆时针方向旋转 45° , 得到图 4(b), 分别采用传统 HOG 算子以及 3-HRPOG 算子计算 2 种情况下的梯度矢量. 分析发现, 传统 HOG 算子受图像旋转影响, 梯度矢量由 $[1, 0]^T$ 变为 $[1, -1]^T$, 而 3-HRPOG 算子的 8 组梯度矢量值不变, 仅出现顺序发生变化, 可见, 3-HRPOG 算子可以提取图像在各个不同方向上的梯度变化信息, 不受图像旋转影响.

2.3 多尺度旋转主方向梯度直方图 (multi-scale histogram of rotated principal orientation gradients, Ms-HRPOG) 特征算子

研究表明, 人眼视觉系统具有多尺度分析能力, 通过调节视网膜细胞感受野范围可以提取不同尺度下的目标信息, 从而更加全面准确地“认识”目标 [49,50]. 受此启发, 本文在 2.2 节 3×3 尺度的 3-HRPOG 算子基础上 (图 3), 将卷积模板尺寸扩展为 5×5 大小, 且卷积权值设置方法不变, 进一步设计出 5×5 尺度的 5-HRPOG 算子 (图 5), 并将 2 种算子相互结合, 提出一种 Ms-HRPOG 特征算子, 特征提取流程如图 6 所示, 其中, 5-HRPOG 特征提取步骤与 3-HRPOG 特征提取一致, 将二者级联得到最终的 Ms-HRPOG 特征向量.

3 稀疏表示和稀疏保持投影

稀疏表示 (sparse representation, SR)^[3] 是继小波变换、多尺度几何分析后, 又一新的信号表示

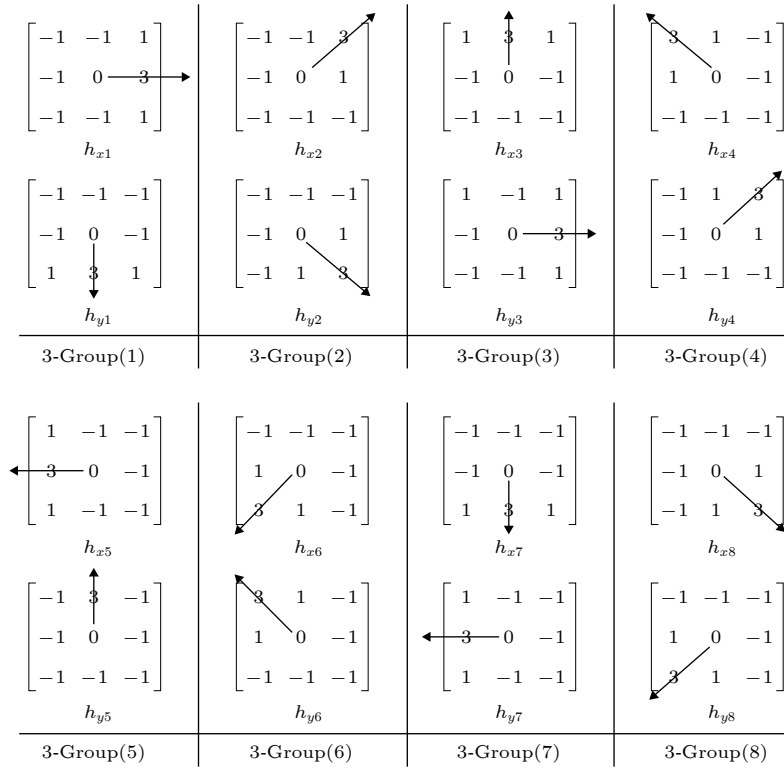


图 3 3-HRPOG 算子的旋转梯度卷积模板

Fig. 3. Rotated gradient convolution masks of 3-HRPOG feature descriptor.

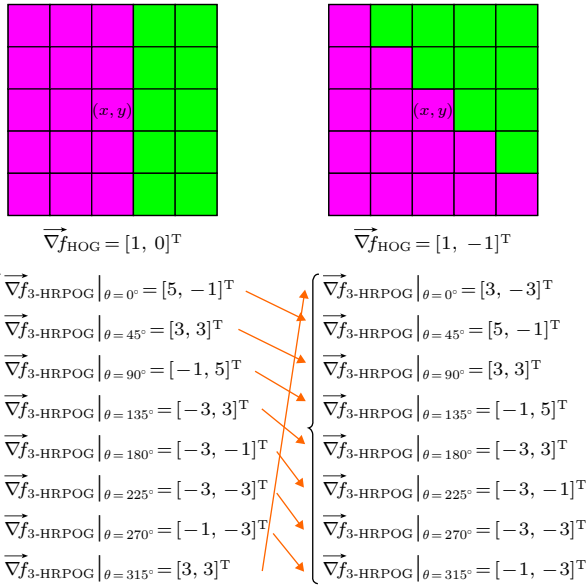


图 4 旋转不变性分析 (a) 原图及 HOG 和 3-HRPOG 的梯度矢量值; (b) 旋转 45° 图像及 HOG 和 3-HRPOG 的梯度矢量值

Fig. 4. Rotation invariance analysis: (a) Original binary image and gradient vectors of HOG and 3-HRPOG; (b) rotated 45° binary image and gradient vectors of HOG and 3-HRPOG.

方法, 基本思想是用超完备字典 D 中少量 (稀疏) 原子线性组合来近似表示待测样本 y :

$$\min_{\alpha} \|y - D\alpha\|_2^2 + \|\alpha\|_1, \quad (6)$$

这里采用 l_1 范数 $\|\bullet\|_1$ 约束向量 α 中非零编码个数, 使其最少 (最稀疏). 乔立山等^[33] 受 SR 启发, 提出了稀疏保持投影 (sparsity preserving projections, SPP) 算法, 通过在低维子空间保持高维样本之间的稀疏重构关系, 寻找最佳低维投影, 目标函数如下:

$$\begin{cases} \min_{s_i} \|s_i\|_1 \\ \text{s.t.} \quad \|x_i - X_i s_i\|_2 < \varepsilon, \\ \mathbf{1}^T s_i = 1 \end{cases} \quad (7)$$

$$\begin{cases} \min_P \sum_{i=1}^N \|P^T x_i - P^T X s_i\|_2^2, \\ \text{s.t.} \quad P^T X X^T P = I \end{cases} \quad (8)$$

其中, $X_i = [x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_N] \in R^{mn \times N}$ 是样本集 X 中去除样本 x_i 的超完备字典, 约束式 $\|x_i - X_i s_i\|_2 < \varepsilon$ 松弛了重构条件, $\mathbf{1}^T s_i = 1$ 使样本 x_i 的稀疏重构权值 s_i 归一化, 这里 $\mathbf{1} = [1, 1, \dots, 1]^T \in R^{N \times 1}$. 可见, SPP 中 s_i 描述了高维空间中样本 x_i 与剩余样本的稀疏重构关系, 通过在低维空间中约束保持该局部属性, 由 (8) 式优化得到最佳低维投影矩阵 $P \in R^{mn \times d}$.

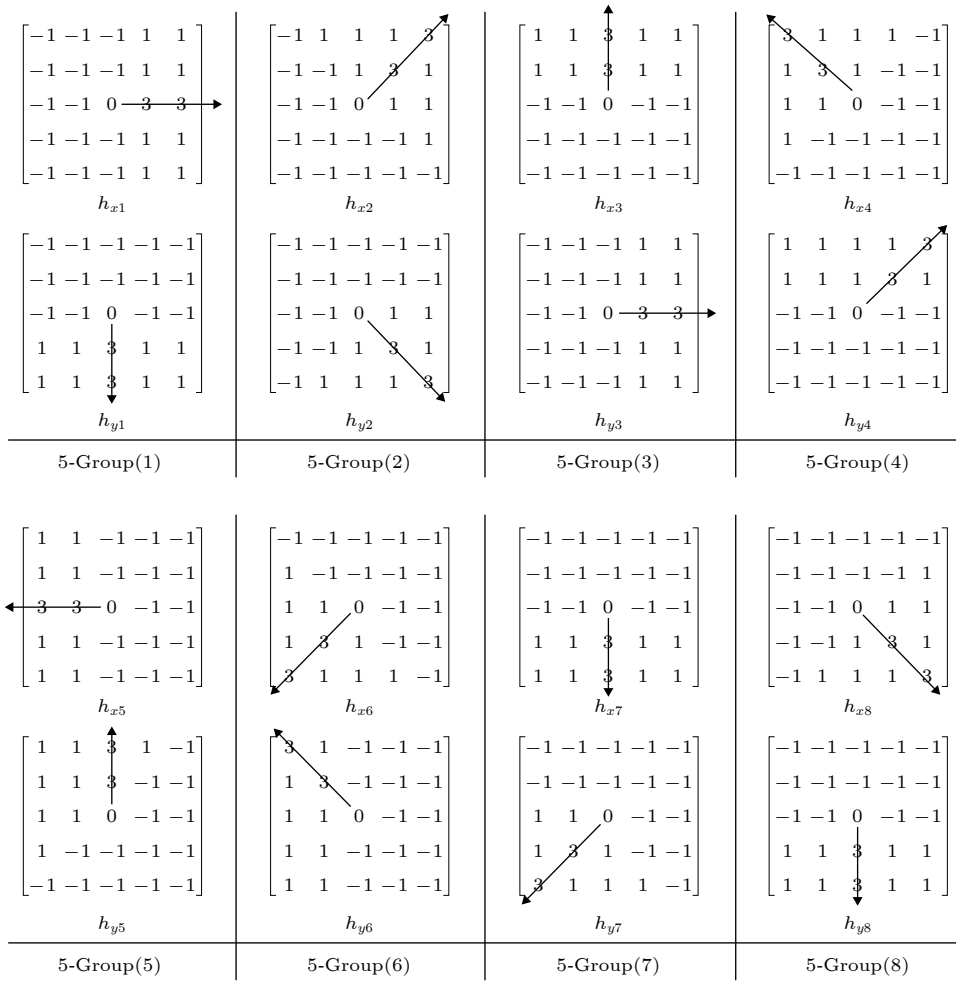


图 5 5-HRPOG 算子的旋转主方向梯度模板

Fig. 5. Rotated dominant direction gradient masks of 5-HRPOG feature descriptor.

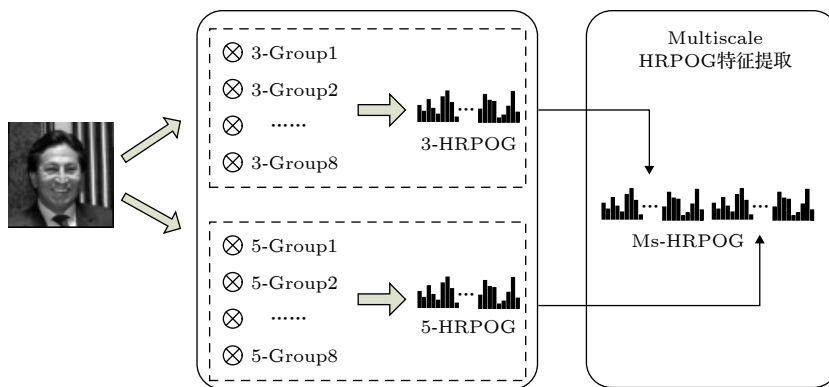


图 6 Ms-HRPOG 特征提取示意图

Fig. 6. The sketch of Ms-HRPOG feature descriptor.

以 LFW 数据库为例, 图 7 给出了根据 (7) 式计算得到的某一样本图像的稀疏重构权值. 图中, 前 10 个为同类样本的重构权值, 后 90 个为异类样本的重构权值. 由图 7 可以看出, 图中仅同类的第 4 个样本和第 8 个样本参与了重构 (有权值), 而剩余的同类样本均未参与重构 (权值为 0); 相反, 某

些异类样本在重构中具有较大权值, 如第 29, 39, 48, 65 个样本等. 造成这种结果的原因是, 真实采集的人脸图像复杂多变, 同类样本受光照、遮挡、表情、年龄、姿态等影响, 彼此之间存在差异, 导致部分同类非近邻样本重构权值为 0, 而不同类样本间又具有人脸结构相似特征, 造成部分异类伪近邻

样本参与重构. 可见, SPP 算法在非约束人脸识别中仍有局限性, 不能准确描述样本的近邻重构关系.

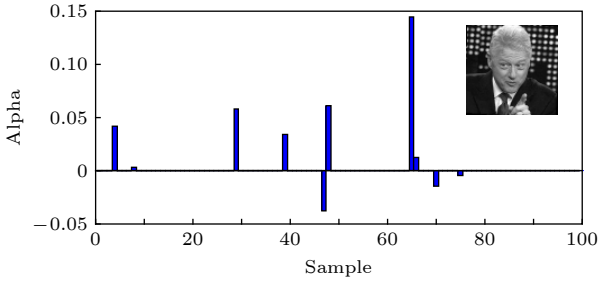


图 7 LFW 数据库中某一图像的 SPP 稀疏重构权值

Fig. 7. Sparsity reconstruction weights of one sample with SPP algorithm on the LFW database.

4 判别稀疏图映射算法

鉴于 SPP 的局限性, 本文提出 DSGE 算法, 从 2 个方面进行改进. 一方面, 在稀疏图构建阶段, 利用类别标签分别构建类内、类间重构关系矩阵, 并且引入类内、类间紧凑度约束因子, 用以增强待测样本与同类非近邻样本的重构关系, 以及削弱异类伪近邻样本的重构影响; 另一方面, 在低维投影阶段, 增加全局约束因子, 利用样本全局分布中隐含的鉴别信息, 以类间重构散度最大、类内重构散度最小为目标计算最佳投影矩阵, 使低维子空间更紧致、更判别.

4.1 类内、类间重构关系矩阵

首先, 分析待测样本与同类样本之间的重构关系. 设训练样本集 X 由 C 类样本子集 X_1, X_2, \dots, X_C 组成, 表示为

$$X = [X_1, X_2, \dots, X_C] \\ = \left[\underbrace{x_{11}, x_{12}, \dots, x_{1n_1}}_{n_1}, \dots, \underbrace{x_{c1}, x_{c2}, \dots, x_{cn_c}}_{n_c} \right] \\ \in R^{mn \times N}, \quad (9)$$

其中, n_k 表示为第 k 类样本子集个数, 满足 $\sum_{k=1}^C n_k = N$. 引入类别标签, 优化求解待测样本 x_{ij} 的类内稀疏重构权值 θ_{ij}^w , 目标函数如下:

$$\begin{cases} \min_{\theta_{ij}^w} \|x_{ij} - X_{l(x_{ij})} \theta_{ij}^w\|_2 + \lambda_1 \|\theta_{ij}^w\|_1 \\ \quad + \lambda_2 \|\theta_{ij}^w - \Gamma_{l(x_{ij})}\|_1 \\ \text{s.t. } \mathbf{1}^T \theta_{ij}^w = 1, \end{cases} \quad (10)$$

这里, $l(x_{ij})$ 表示待测样本 x_{ij} 的类别, 取 $l(x_{ij}) = i$, 定义类内超完备字典

$X_{l(x_{ij})} = [x_{i1}, \dots, x_{ij-1}, 0, x_{ij+1}, \dots, x_{in_i}] \in R^{mn \times n_i}$, 表示为除 x_{ij} 外的剩余同类样本子集. 定义 $\Gamma_{l(x_{ij})}$ 为类内紧凑度约束因子, 记为与 x_{ij} 同类的所有样本的类内稀疏重构权值的平均向量, 最小化 $\|\theta_{ij}^w - \Gamma_{l(x_{ij})}\|_1$ 可以使同类样本的稀疏重构权值接近中心值, 增强待测样本与同类非近邻样本的重构关系.

依据 (10) 式分别计算样本子集 X_i ($i = 1, 2, \dots, C$) 中同类样本的类内稀疏重构权值 $\theta_{ij}^w \in R^{n_i \times 1}$, 从而得到样本子集 X_i 的类内重构权值矩阵 $\Theta_i^w = [\theta_{i1}^w, \theta_{i2}^w, \dots, \theta_{in_i}^w] \in R^{n_i \times n_i}$, 由此构建样本集 X 的类内重构关系矩阵 $\mathbf{A}^w \in R^{N \times N}$:

$$\mathbf{A}^w = \text{diag}(\Theta_1^w, \Theta_2^w, \dots, \Theta_C^w). \quad (11)$$

分析发现, (11) 式保证了待测样本与同类样本的重构关系, 然而大量研究表明, 异类样本在样本集中的空间分布对低维投影的准确性也有很大影响, 因此, 有必要进一步考虑待测样本与异类样本的重构关系. 定义目标函数如下:

$$\begin{cases} \min_{\theta_{ij}^b} \|x_{ij} - X_{\overline{l(x_{ij})}} \theta_{ij}^b\|_2 + \lambda_3 \|\theta_{ij}^b\|_1 \\ \quad + \lambda_4 \|\theta_{ij}^b - \Gamma_{\overline{l(x_{ij})}}\|_1 \\ \text{s.t. } \mathbf{1}^T \theta_{ij}^b = 1, \end{cases} \quad (12)$$

其中, $X_{\overline{l(x_{ij})}}$ 为类间超完备字典, 由与 x_{ij} 不同类别的训练样本组成. 设 x_{ij} 属于第 k 类样本子集, 则 $X_{\overline{l(x_{ij})}} = [X_1, X_2, \dots, X_{k-1}, X_{k+1}, \dots, X_C] \in R^{mn \times (N - n_k)}$. 根据 (12) 式优化得到待测样本 x_{ij} 的类间稀疏重构权值 θ_{ij}^b . 定义 $\Gamma_{\overline{l(x_{ij})}}$ 为类间紧凑度约束因子, 记为与 x_{ij} 同类样本的类间稀疏重构权值的平均向量, 最小化 $\|\theta_{ij}^b - \Gamma_{\overline{l(x_{ij})}}\|_1$ 可以进一步削弱异类伪近邻样本对待测样本的重构影响.

为了便于描述样本集 X 的类间重构关系, 将 θ_{ij}^b 扩展为 N 维列向量, 即

$$\theta_{ij}^{b\text{-extended}} = \begin{bmatrix} \beta_{i1}, \beta_{i2}, \dots, \beta_{i(n_1 + \dots + n_{i-1})}, \underbrace{0, \dots, 0}_{n_i \uparrow 0} \\ \beta_{i(n_1 + \dots + n_i + 1)}, \dots, \beta_{iN} \end{bmatrix}^T \in R^{N \times 1},$$

则样本子集 X_i 的类间重构权值矩阵表示为

$$\theta_i^{b\text{-extended}} = [\theta_{i1}^{b\text{-extended}}, \theta_{i2}^{b\text{-extended}}, \dots, \theta_{in_i}^{b\text{-extended}}] \in R^{N \times n_i},$$

由此得到样本集 X 的类间重构关系矩阵 $B^b \in R^{N \times N}$:

$$B^b = [\Theta_1^{b\text{-extended}}, \Theta_2^{b\text{-extended}}, \dots, \Theta_c^{b\text{-extended}}]. \quad (13)$$

4.2 DSGE 低维投影

首先, 为了使低维子空间中同类样本尽可能聚合, 在最小化类内重构误差基础上, 增加全局类内约束条件, 使得低维子空间数据既保留了样本间局部类内稀疏重构关系, 又考虑了同类样本子集的全局紧凑度分布特性. 最小化类内重构散度目标函数如下:

$$\min_P \sum_{i=1}^N \left(\left\| P^T x_i - \sum_{j=1}^N P^T \alpha_{ji}^w x_j \right\|_F^2 \right) + \|P^T X_w\|_F^2. \quad (14)$$

这里为了描述方便, 不考虑样本类别, 定义 x_i 为样本集 X 中的第 i 个样本, x_j 为第 j 个样本, α_{ji}^w 对应类内重构关系矩阵 Λ^w 中相应元素, 描述了样本 x_i 与 x_j 的类内重构关系. 定义全局类内约束因子 $X_w = [X_1 - \bar{X}_1, \dots, X_i - \bar{X}_i, \dots, X_c - \bar{X}_c]$, X_i 是第 i 类样本子集, \bar{X}_i 则是第 i 类样本子集的均值矩阵, 约束 $\|P^T X_w\|_F^2$ 最小, 可以使得各类样本子集类内更紧凑. 推导目标函数 (14) 得到

$$\begin{aligned} & \sum_{i=1}^n \left(\left\| P^T x_i - \sum_{j=1}^n P^T \alpha_{ji}^w x_j \right\|_F^2 \right) + \|P^T X_w\|_F^2 \\ &= \sum_{i=1}^n \left(\|P^T x_i - P^T X \Lambda_i^w\|_F^2 \right) + \|P^T X_w\|_F^2 \\ &= \sum_{i=1}^n \left[\text{tr} \left(P^T (x_i - X \Lambda_i^w) (x_i - X \Lambda_i^w)^T P \right) \right] \\ & \quad + \text{tr} (P^T X_w X_w^T P) \\ &= \text{tr} \left[P^T \left(\sum_{i=1}^n (x_i - X \Lambda_i^w) (x_i - X \Lambda_i^w)^T \right) P \right] \\ & \quad + \text{tr} (P^T X_w X_w^T P) \\ &= \text{tr} \left[P^T X \left(I - \Lambda^w - (\Lambda^w)^T + (\Lambda^w)^T \Lambda^w \right) X^T P \right] \\ & \quad + \text{tr} (P^T X_w X_w^T P). \end{aligned} \quad (15)$$

取 $L^w = I - \Lambda^w - (\Lambda^w)^T + (\Lambda^w)^T \Lambda^w$, 则目标函数 (14) 转换为

$$\min_P \text{tr} [P^T (X L^w X^T + X_w X_w^T) P]. \quad (16)$$

同理, 为了使低维子空间中异类样本尽可能分离, 在最大化类间重构误差基础上, 又增加全局类间约束条件, 得到最大化类间重构散度目标函数如下:

$$\max_P \sum_{i=1}^N \left(\left\| P^T x_i - \sum_{j=1}^N P^T \beta_{ji}^b x_j \right\|_F^2 \right) + \|P^T X_b\|_F^2, \quad (17)$$

其中, β_{ji}^b 为类间重构关系矩阵 B^b 中相应元素, 描述了样本 x_i 与 x_j 的类间重构关系. 定义全局类间约束因子 $X_b = [\bar{X}_1 - \bar{X}, \dots, \bar{X}_i - \bar{X}, \dots, \bar{X}_c - \bar{X}]$, 取 \bar{X} 为所有样本的均值矩阵, 约束 $\|P^T X_b\|_F^2$ 最大, 可以使得各类样本子集尽可能分离, 低维子空间数据更具有判别性. 依据目标函数 (14) 推导过程, 目标函数 (17) 也可转换 (18) 式, 其中, $L^b = I - B^b - (B^b)^T + (B^b)^T B^b$.

$$\max_P \text{tr} [P^T (X L^b X^T + X_b X_b^T) P]. \quad (18)$$

结合 (16) 式和 (18) 式, 基于最大边界准则 (maximum margin criterion, MMC) 得到 DSGE 低维投影目标函数:

$$\begin{aligned} \max_P J(P) &= \text{tr} [P^T (X L^b X^T + X_b X_b^T) P] \\ & \quad - \text{tr} [P^T (X L^w X^T + X_w X_w^T) P] \end{aligned} \quad (19)$$

由拉格朗日乘法, 在 $P^T P = I$ 约束条件下, (19) 式转换为如下特征方程:

$$[X L^b X^T + X_b X_b^T - X L^w X^T - X_w X_w^T] p = \lambda p. \quad (20)$$

求取前 d 个最大特征值对应的特征向量, 构成最佳低维映射矩阵 $P^* = [p_1, p_2, \dots, p_d] \in R^{m \times d}$.

5 基于 DSGE-HPROG 的算法

综上所述, 稀疏表示可以理解为函数逼近问题, 选择一个好的基函数对 SRC 分类器非常重要, 因此, 本文将 HRPOG 和 DSGE 相结合, 提出一种 DSGE-HPROG 算法, 用以增强低维特征字典的鉴别性、紧致性和稳定性, 使 SRC 分类效果更理想. 同时, 考虑 DSGE 算法中稀疏重构关系与

SRC 分类器的稀疏表示系数具有相同物理意义, 本文又借鉴 SRC-DP^[40] 思想, 将低维投影和稀疏图构建联合优化, 使 SRC 分类器在低维投影子空间获得最佳识别效果, 具体实现步骤如算法 1 所示.

算法 1 DSGE-HRPOG 算法

输入: 训练样本 $X = [X_1, X_2, \dots, X_c] \in \mathbb{R}^{m \times N}$, 参数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, k_{\max}$ 和 PCA Ratio.
输出: 最佳投影矩阵 P^* .

步骤 1. 采用 HRPOG 算子 (单尺度或多尺度) 提取训练样本 X 的特征, 并采用 PCA 对 HRPOG 特征向量进行预处理降维, 构建特征字典 D_{HRPOG}

步骤 2. 初始化投影矩阵 P_0 , 取迭代次数 $k = 1$;

步骤 3. 将特征字典 D_{HRPOG} 投影到变换空间 $Z = P_{k-1}^T D_{\text{HRPOG}}$;

步骤 4. 依据 (10–13) 式计算变换空间 $Z = [Z_1, Z_2, \dots, Z_c]$ 中任一样本 z_{ij} 与同类样本的重构关系矩阵 A^w , 以及与异类样本的重构关系矩阵 B^b ;

步骤 5. 计算 $L^w = I - A^w - (A^w)^T + (A^w)^T A^w$ 和 $L^b = I - B^b - (B^b)^T + (B^b)^T B^b$, 代入特征方程 (20) 中, 取前 d 个最大特征值对应的特征向量构成投影矩阵 P_k ;

步骤 6. 取 $k = k + 1$, 计算收敛条件 $|J(P_k) - J(P_{k-1})| / |J(P_k)|$;

步骤 7. 若 $\frac{|J(P_k) - J(P_{k-1})|}{|J(P_k)|} < \varepsilon$ 或迭代次数 k 大于设定值 k_{\max} , 则最佳投影矩阵 $P^* = P_k$, 反之, 重复步骤 3—步骤 7.

6 实验结果

为了验证 DSGE-HRPOG 算法的有效性, 分别在 AR, Extended Yale B, LFW 和 PubFig 这

4 个人脸数据库上进行实验仿真. 实验主要参数设置为 $\lambda_1 = \lambda_3 = 0.5, \lambda_2 = \lambda_4 = 1, k_{\max} = 20$. 所有实验均采用 PCA 进行数据预处理 (取 PCA Ratio = 0.98), 在 Win10 64 位操作系统、8 GB 内存和 MATLAB R2017a 仿真环境下完成.

6.1 AR 数据库

AR 数据库是在严格控制实验环境条件下采集的具有不同面部表情、照明条件和遮挡 (包括太阳镜和围巾) 的人脸正视图, 图像尺寸为 50×40 像素大小. 本文选取 120 类人 (65 男 55 女) 在 2 个阶段拍摄的图像, 其中每人每个阶段包含 7 张无遮挡图像 (包含表情、光照影响) 和 6 张遮挡图像 (3 张墨镜遮挡, 3 张围巾遮挡), 样本如图 8 所示. HRPOG 算子的参数设置为 1 个 cell = 2×2 pixels, 1 个 block = 2×2 cells, bin = 9.

6.1.1 表情、光照、时间的影响

本实验主要分析表情、光照、时间等非遮挡因素对人脸识别的影响. 取 AR 数据库中每类人在 Session1 中 7 张无遮挡图像作为训练样本, Session2 中 7 张无遮挡图像作为测试样本进行实验仿真 (表 1). 由表 1 可以看出, 在本文提出的 DSGE-HRPOG 算法框架下, 无论采用单尺度特征算子 (3-HRPOG 和 5-HRPOG), 还是多尺度特征算子 (Ms-HRPOG) 构建特征字典, DSGE-HRPOG 的识别率均高于其他算法, 与 DSNPE^[37], DP-NFL^[51] 和 SRC-DP^[40] 相比, 识别率最大提升了 12.74%, 17.01% 和 13.61%. 这说明, 本文提出的 DSGE-HRPOG 算法在降低人脸识别中表情、光照、时间等非遮挡影响时, 具有绝对的优势. 同时, 本实验又基于原始图像构建样本字典, 并采用 DSGE 算法进行字典降维, 得到 DSGE-pixels 的识别率为 76.79%, 相比 SPP^[33], DSNPE^[37], DP-

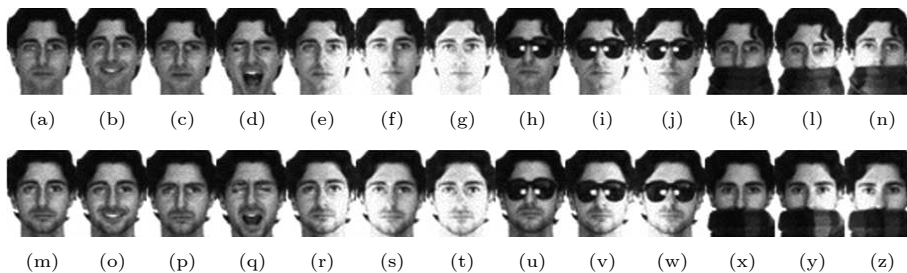


图 8 AR 数据库部分样本图像
Fig. 8. Samples of one person in the AR database.

表 1 AR 数据库在表情、光照和时间干扰因素下的实验结果

Table 1. Experimental results on the AR database with the interference factors of expression, illumination and time.

| Method | LPP ^[22] | NPE ^[23] | SPP ^[33] | DSNPE ^[37] | DP-NFL ^[51] | SRC-DP ^[40] | DSGE-pixels | DSGE-HRPOG | | |
|--------------------|---------------------|---------------------|---------------------|-----------------------|------------------------|------------------------|-------------|------------|---------|----------|
| | | | | | | | | 3-HRPOG | 5-HRPOG | Ms-HRPOG |
| Recognition Rate/% | 67.14 | 68.69 | 68.21 | 76.07 | 71.8 | 75.2 | 76.79 | 88.45 | 88.21 | 88.81 |
| Dimension | 115 | 311 | 220 | 140 | 63 | 63 | 322 | 777 | 754 | 774 |

NFL^[51] 和 SRC-DP^[40], 识别率分别提升了 8.58%, 0.72%, 4.99% 和 1.59%. 这也从另一个角度说明, 摒弃特征字典的作用, 本文提出的 DSGE 算法在去除数据冗余、构建低维判别字典方面也具有一定优势, 但将 DSGE 与 HRPOG 特征字典相结合, 识别效果更佳, 如表 1 中所示, DSGE-HRPOG 算法的识别率比 DSGE-pixels 最大提升了 12.02%.

6.1.2 遮挡的影响

本实验主要分析遮挡因素对人脸识别的影响, 包括眼镜遮挡、围巾遮挡以及混合遮挡 3 个方面. 下面有针对性地设计 3 个实验, 依次进行实验仿真 (表 2).

表 2 AR 数据库在遮挡干扰因素下的实验结果

Table 2. Experimental results of AR database with the occlusion interference.

| | Experiment | | |
|-------------------------|------------|-------|-------|
| | 1/% | 2/% | 3/% |
| LPP ^[22] | 71.39 | 68.68 | 69.46 |
| NPE ^[23] | 72.64 | 71.81 | 71.08 |
| SPP ^[33] | 75.90 | 72.92 | 74.07 |
| DSNPE ^[37] | 80.28 | 78.26 | 78.14 |
| SRC-DP ^[40] | 78.35 | 76.50 | 77.80 |
| SRC-FDC ^[42] | 80.90 | 79.90 | 80.30 |
| DSGE-pixels | 79.03 | 78.75 | 82.65 |
| DSGE-HRPOG (3-HRPOG) | 88.54 | 89.51 | 90.53 |
| DSGE-HRPOG (5-HRPOG) | 89.31 | 89.58 | 90.98 |
| DSGE-HRPOG (Ms-HRPOG) | 89.31 | 90.00 | 91.06 |

实验 1: 取 AR 数据库中每类人在 Session1 中 7 张无遮挡图像和任意 1 张眼镜遮挡图像作为训练样本, 而 Session2 中 7 张无遮挡图像和 3 张眼镜遮挡图像, 以及 Session1 中剩余的 2 眼镜遮挡图像作为测试样本.

实验 2: 取 AR 数据库中每类人在 Session1 中 7 张无遮挡图像和任意 1 张围巾遮挡图像作为训练样本, 而 Session2 中 7 张无遮挡图像和 3 张

围巾遮挡图像, 以及 Session1 中剩余的 2 围巾遮挡图像作为测试样本.

实验 3: 取 AR 数据库中每类人在 Session1 中 7 张无遮挡图像和任意 1 张眼镜遮挡图像以及任意 1 张围巾遮挡图像作为训练样本, 而 Session2 中 7 张无遮挡图像、3 张眼镜遮挡图像、3 张围巾遮挡, 以及 Session1 中剩余的眼镜和围巾遮挡图像作为测试样本.

分析表 2 中 3 个实验结果, 可以看出, 本文提出的 DSGE-HRPOG 算法在眼镜遮挡、围巾遮挡以及混合遮挡 3 种真实遮挡情况下, 均具有最佳识别效果, 其中, 基于 Ms-HRPOG 特征字典的识别率最高, 分别为 89.31%, 90.00% 和 91.06%, 比次优算法 SRC-FDC^[42, 52] 高 8.41%, 10.1% 和 10.76%. 这说明, DSGE-HRPOG 算法通过结合 HRPOG 特征字典以及 DSGE 低维投影的优势, 可以有效消除人脸识别中各种遮挡因素的影响, 具有较强的遮挡鲁棒性. 同时, 从表 2 中也可以看出, 基于多尺度特征字典 (Ms-HRPOG) 的 DSGE-HRPOG 的识别率均高于基于单尺度特征字典 (3-HRPOG 和 5-HRPOG) 的情况, 这也说明, 采用多尺度特征融合策略有助于进一步提升特征字典的鉴别能力, 增强系统的稳定性.

6.1.3 混合影响

前 2 个实验分别分析了遮挡因素以及非遮挡因素对人脸识别的影响, 本实验将遮挡因素和非遮挡因素综合考虑, 分析混合因素 (遮挡、表情、光照、时间) 对人脸识别的影响. 随机取每类人 26 张 (2 个 Session) 图像中的 13 张图像作为训练样本, 剩余图像作为测试样本进行实验仿真, 交叉验证 10 次 (表 3). 分析发现, 当样本中随机包含各种不同干扰因素时, 本文提出的 DSGE-HRPOG 算法在消除人脸识别中遮挡、表情、光照、时间等混合影响时, 仍具有一定优势. 相比文献 [53] 和 [54] 中提出的 2 个先进算法, 基于 Ms-HRPOG 特征字典的 DSGE-HRPOG 算法的平均识别率提升最多,

分别提升了 0.96% 和 1.7%。同时, 文献 [53] 和 [54] 算法的标准差分别为 0.53 和 0.93, 而 DSGE-HRPOG 的标准差最大为 0.17, 远远小于文献值, 这也进一步说明了本文提出的 DSGE-HRPOG 算法不受样本选择和干扰因素变化的影响, 具有分类稳定性。

表 3 AR 数据库在混合干扰因素下的实验结果
Table 3. Experimental results on the AR database with the mix interference factors.

| | Mean/% | Std/% | Dimension |
|--------------------------|--------|-------|-----------|
| LPP ^[22] | 95.90 | 0.38 | 141 |
| NPE ^[23] | 95.87 | 0.55 | 311 |
| SPP ^[33] | 90.39 | 0.90 | 151 |
| DSNPE ^[37] | 98.02 | 0.33 | 200 |
| Wang ^[54] | 97.85 | 0.93 | — |
| Gao ^[53] | 98.59 | 0.53 | — |
| DSGE-pixels | 98.45 | 0.27 | 202 |
| DSGE-HRPOG (3-HRPOG) | 99.45 | 0.17 | 1350 |
| DSGE-HRPOG (5-HRPOG) | 99.37 | 0.12 | 1297 |
| DSGE-HRPOG (Ms-HRPOG) | 99.55 | 0.13 | 1385 |

6.2 Extended Yale B 数据库

Extended Yale B 数据库包含 38 类人在不同光照条件下拍摄的人脸正视图, 每类人约 64 张图像, 共 2414 张, 图像尺寸为 32×32 像素大小, 部分样本图像如图 9 所示。这里, HRPOG 算子的参数设置与 AR 数据库一致。



图 9 Extended Yale B 数据库部分样本图像

Fig. 9. Samples of one person in the Extended Yale B database.

在 Extended Yale B 数据库上, 首先分析不同强度光照对人脸识别的影响。随机选取每类人的 10 张图像作为训练样本, 剩余图像作为测试样本进行实验仿真 (表 4)。由表 4 可以看出, 在多尺度特征字典情况下 (Ms-HRPOG), DSGE-HRPOG 算法可以达到 92.48% 的识别率, 相比近年来提出的先进算法 GRSDA^[39] 和 RCDA^[52], 分别提升了 9.78% 和 0.48%, 识别效果最佳。而在单尺度特征

字典 (3-HRPOG 和 5-HRPOG) 情况下, DSGE-HRPOG 的识别率分别为 91.35% 和 89.77%, 略低于 RCDA^[52]。可见, 在少量训练样本条件下, 多尺度特征更有助于消除人脸识别中不同强度的光照影响。

其次, 进一步验证 DSGE-HRPOG 算法在 Extended Yale B 数据库上的遮挡鲁棒性。本文又随机选取每类人 14 张图像, 在上面分别添加大小随机、位置随机的黑白点噪声块, 得到部分遮挡样本图像 (图 10)。这里, 噪声块大小与图像大小比例在 0.05—0.15 任意取值。

下面设计 2 个实验, 用以讨论遮挡样本数量不同时, DSGE-HRPOG 算法的性能。实验 1: 随机选取每类人 32 张图像作为训练样本, 其中包含 14 张遮挡图像, 剩余图像作为测试样本, 交叉验证 10 次; 实验 2: 随机选取每类人 32 张图像作为训练样本, 其中包含 7 张遮挡图像, 剩余图像作为测试样本, 交叉验证 10 次, 实验仿真结果如表 5 所示。从表 5 中可看出, 在光照和遮挡混合干扰情况下, DSGE-HRPOG 算法仍具有最佳性能。当遮挡样本数为 7 张时, 基于 Ms-HRPOG 特征字典的平均识别率为 98.10%, 当遮挡样本数增大到 14 张时, 识别率仍有 97.98%, 二者仅相差 0.12%。同样地, 基于 3-HRPOG 和 5-HRPOG 特征字典的平均识别率也高于其他算法, 并且在不同遮挡样本数量条件下, 实验结果仅相差 0.43% 和 0.08%, 而其余算法则有近 1% 左右的差值。这充分说明, 无论是基于单尺度特征字典还是多尺度特征字典, DSGE-HRPOG 算法不受遮挡样本数量和质量的影

6.3 LFW 和 PubFig 数据库

前面讨论是基于实验环境下采集的人脸数据库 (AR 和 Extended Yale B) 进行实验仿真和分析的, 实验结果具有一定的局限性, 本节将进一步对真实环境中采集的 LFW 和 PubFig 数据库上进行实验仿真, 使实验结果更具有说服力。

LFW (labeled faces in the Wild database) 数据库^[55–57] 是从 Internet 上采集的真实人脸数据库, 共有 13233 张 5749 类人脸图像, 包含了光照、表情、姿态、遮挡、年龄、种族等多种混合干扰, 对于人脸识别来说非常具有挑战性。部分样本如图 11(a) 所示, 图像尺寸为 128×128 pixels。本文从中选取

表 4 Extended Yale B 数据库在光照干扰因素下的实验结果
Table 4. Experimental results of Extended Yale B database with the illumination interference.

| Method | LPP ^[22] | NPE ^[23] | SPP ^[33] | DSNPE ^[37] | GRSDA ^[39] | RCDA ^[52] | DSGE-pixels | DSGE-HRPOG | | |
|--------------------|---------------------|---------------------|---------------------|-----------------------|-----------------------|----------------------|-------------|------------|---------|----------|
| | | | | | | | | 3-HRPOG | 5-HRPOG | Ms-HRPOG |
| Recognition Rate/% | 87.86 | 89.31 | 85.79 | 85.74 | 82.7 | 92 | 86.03 | 91.35 | 89.77 | 92.48 |
| Dimension | 65 | 160 | 95 | 85 | 266 | — | 83 | 355 | 345 | 351 |



图 10 Extended Yale B 数据库部分遮挡样本图像
Fig. 10. Occlusion samples of one person in the Extended Yale B database.

表 5 Extended Yale B 数据库在遮挡干扰因素下的实验结果
Table 5. Experimental results of Extended Yale B database with the occlusion interference.

| | Experiment 1/% | Experiment 2/% |
|------------------------|----------------|----------------|
| LPP ^[22] | 95.51 ± 0.40 | 96.78 ± 0.72 |
| NPE ^[23] | 96.43 ± 0.23 | 97.85 ± 0.31 |
| SPP ^[33] | 92.57 ± 0.84 | 93.05 ± 0.77 |
| DSNPE ^[37] | 94.18 ± 0.48 | 95.29 ± 0.54 |
| Gao ^[53] | 86.91 ± 1.07 | 88.23 ± 0.91 |
| DSGE-pixels | 95.83 ± 0.66 | 96.21 ± 0.21 |
| DSGE-HRPOG(3-HRPOG) | 97.30 ± 0.20 | 97.73 ± 0.35 |
| DSGE-HRPOG(5-HRPOG) | 96.85 ± 0.38 | 96.93 ± 0.60 |
| DSGE-HRPOG G(Ms-HRPOG) | 97.98 ± 0.50 | 98.10 ± 0.31 |

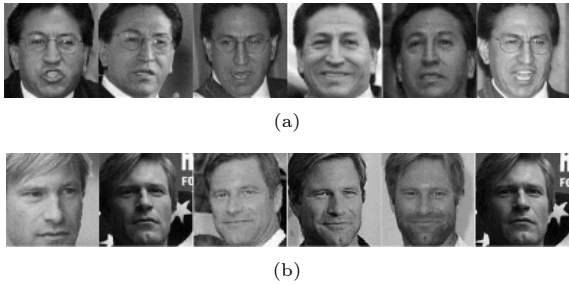


图 11 部分样本图像 (a) LFW 数据库部分样本; (b) PubFig 数据库部分样本
Fig. 11. Samples of one person: (a) LFW database; (b) PubFig database.

包含 10 张以上图像的人进行辨识, 得到 158 类人, 共 4324 张图像, 随机选取每类人的 5 张图像作为训练样本, 5 张图像作为测试样本进行实验 (表 6).

HRPOG 特征算子的参数设置为一个 cell = 4×4 pixels, 一个 block = 2×2 cells, bin = 9.

表 6 LFW 和 PubFig 数据库的实验结果
Table 6. Experimental results on the LFW database and PubFig database.

| | LFW/% | PubFig/% |
|-----------------------|-------|----------|
| LPP ^[22] | 35.32 | 24.00 |
| NPE ^[23] | 35.19 | 25.00 |
| SPP ^[33] | 31.52 | 29.00 |
| DSNPE ^[37] | 44.05 | 30.90 |
| WGSC ^[58] | 47.60 | 37.50 |
| RSRC ^[3] | 42.80 | 47.00 |
| RRC ^[57] | 53.20 | 42.20 |
| IRGSC ^[41] | 56.30 | 48.50 |
| DSGE-pixels | 51.52 | 38.60 |
| DSGE-HOG | 69.62 | 49.00 |
| DSGE-HRPOG(3-HRPOG) | 76.71 | 54.20 |
| DSGE-HRPOG(5-HRPOG) | 76.58 | 53.30 |
| DSGE-HRPOG(Ms-HRPOG) | 73.80 | 53.70 |

PubFig(public figures face database) 数据库^[56]与 LFW 数据库类似, 包括从互联网上采集到的 200 类知名人物的 58797 张图像. 数据库中的人脸都是真实环境下拍摄的, 包含部分遮挡(眼镜、帽子等饰物)、极端光照、较大的姿势变换 (> 45°)、不同种族、年龄等干扰因素. 部分样本如图 11(b)所示, 图像尺寸为 100×100 pixels. 本文从 PubFig 数据库中随机选取 100 类人, 每类人 20 张图像进行实验仿真, 其中每类人的 10 张图像作为训练样

本, 剩余图像作为测试样本 (表 6). HRPOG 特征算子的参数设置为一个 cell = 10×10 pixels, 一个 block = 2×2 cells, bin = 9.

从表 6 中可以看出, DSGE-HRPOG 算法在 LFW 和 PubFig 这 2 个数据库上的识别率均高于其他算法, 其中, 基于 3-HRPOG 特征字典的识别率最高, 达到 76.71% 和 54.20%, 比文献 [41] 提出的 IRGSC 算法提升了 20.41% 和 5.7%, 比文献 [57] 提出的 RRC 算法提升了 23.51% 和 12%, 比其他算法则提升更多. 然而, 表中 DSGE-pixels 的识别效果并不理想, 在 LFW 数据库上识别率仅为 51.52%, 低于 IRGSC[41] 和 RRC[57], 在 PubFig 数据库上识别率为 38.6%, 低于 IRGSC[41], RRC[57] 和 RSRC[3]. 可见, 在 LFW 和 PubFig 这 2 个具有挑战性的非约束人脸数据库上, DSGE 算法容易受样本影响, 近邻图构建不准确, 而 DSGE-HRPOG 算法通过结合 HRPOG 特征字典以及 DSGE 低维投影两方面的优势, 能够更准确挖掘出嵌入在真实复杂数据中的低维流形本质结构, 显著提高了系统的判别能力.

值得注意的是, 本文也采用传统 HOG 特征算子构建特征字典, 再与 DSGE 算法相结合, 得到 DSGE-HOG 的识别率为 69.62% 和 49.00%, 分类效果仅次于 DSGE-HRPOG 算法, 并优于其他非特征字典算法. 这进一步说明了, 在稀疏表示分类中, 首先采用特征算子提取图像特征用以构建特征字典, 对准确区分受不同因素干扰的真实环境采集的非约束人脸图像十分有效. 同时, 这也从另一角度验证了, 相比传统 HOG 特征算子, 本文提出的 HRPOG 特征算子从多个尺度、多个方向捕捉人脸纹理的梯度变化信息, 可以更准确提取出人脸本质特征, 从而显著提高特征字典的鉴别性和鲁棒性.

6.4 参数性能分析

本节主要讨论 DSGE-HRPOG 算法中最优参数选择, 包括初始投影矩阵 P_0 的选择、最大迭代次数 k_{\max} 的选择, 以及有联合优化和无联合优化条件下的分类结果比较. 以 PubFig 数据库为例, 图 12 为 P_0 取随机矩阵、单位矩阵以及 LDA 投影矩阵作为初始值时, 在不同特征字典条件下的识别结果. 从图 12 中可以看出, 无论采用单尺度特征字典 (3-HRPOG 或 5-HRPOG), 还是多尺度特征字典 (Ms-HRPOG), 取初始投影矩阵 P_0 为随机矩阵时,

联合优化得到的 P^* 最佳, 分类效果最好.

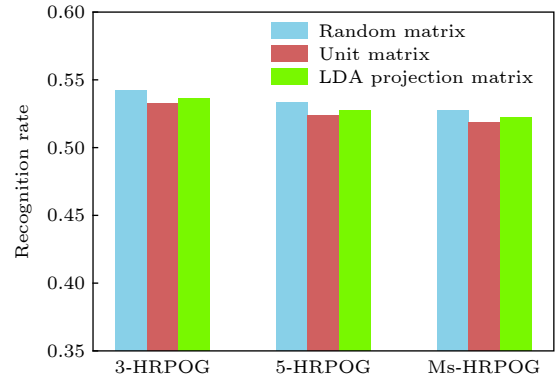


图 12 不同初始投影矩阵 P_0 的识别率

Fig. 12. Recognition rates based on different initial matrix P_0 .

图 13 为迭代次数与目标函数差值的曲线图, 从图 13 中可以看出, 随着迭代次数 k 的增加, 目标函数 $J(P_k)$ 与 $J(P_{k-1})$ 的差值越来越小, 当 $k \geq 20$ 后, 逐渐趋于稳定, 目标函数收敛. 因此, 本文选取最大迭代次数 $k_{\max} = 20$ 作为算法 1 的截止条件之一.

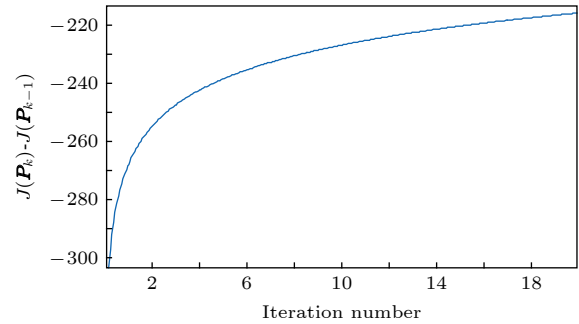


图 13 目标函数收敛曲线

Fig. 13. Convergence curve of the objective function.

最后, 表 7 给出有联合优化和无联合优化条件下, DSGE-HRPOG 算法在 PubFig 数据库上的识别结果及对应最佳投影维度. 这里取初始投影矩阵

表 7 PubFig 数据库上有联合优化和无联合优化的实验结果

Table 7. Experimental results with joint optimization and without joint optimization on the PubFig database.

| | DSGE-HRPOG | | |
|----------------------------|-------------|-------------|-------------|
| | 3-HRPOG | 5-HRPOG | Ms-HRPOG |
| with joint optimization | 54.20 (630) | 53.30 (473) | 53.70 (514) |
| without joint optimization | 53.50 (514) | 50.90 (423) | 53.20 (514) |

P_0 为随机矩阵, 最大迭代次数 $k_{\max} = 20$. 从表 7 中可以看出, 将维数约简过程伴随在稀疏图构建过程中, 可以使低维子空间数据逐渐逼近最佳判别分布, DSGE-HRPOG 分类效果更理想.

7 结 论

鉴于 SRC 分类器的局限性, 本文从字典构建、低维投影以及联合优化 3 个方面进行改进, 提出一种基于 DSGE-HRPOG 的算法. 首先, 利用鉴别力和鲁棒性更强的 HRPOG 特征算子提取样本本质特征, 去除非约束人脸图像中包含的各种干扰因素, 增强同类字典原子间的相关性和异类原子间的差异性; 然后, 再与 DSGE 算法相结合, 进一步提升了特征字典的紧致性、判别性和稳定性; 最后, 将稀疏图构建与低维投影联合优化, 在 AR, Extended Yale B, LFW 以及 PubFig 这 4 个人脸数据库上进行实验仿真, 实验结果均表明, DSGE-HRPOG 算法相比传统流形学习算法以及现有先进算法, 均表现出卓越性能. 需要强调的是, DSGE-HRPOG 算法提供了一种稀疏图映射与局部特征提取相结合的研究思路, 为非约束人脸识别的实用化提供了有效解决方案.

近年来, 有许多学者^[59,60]将图像、视频等数据组织为张量形式进行稀疏图映射降维, 避免了将多维数据强制表示为向量形式而引发的一系列问题. 下一步, 我们将研究基于张量特征字典的稀疏重构技巧及优化算法, 对流形学习与字典学习进行改进和推广, 力争进一步提高非约束人脸识别准确率.

参考文献

- [1] Qian J J, Luo L, Yang J, Zhang F L, Lin Z C 2015 *Pattern Recognit.* **48** 3145
- [2] Chen Y, Yang J, Luo L, Zhang H M, Qian J J, Tai Y, Zhang J 2016 *Pattern Recognit.* **59** 26
- [3] Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y 2009 *IEEE Trans. Pattern Anal. Mach. Intell.* **31** 210
- [4] Yang M, Zhang L, Feng X C, Zhang D 2014 *Int. J. Comput. Vis.* **109** 209
- [5] Vu T H, Monga V 2016 *The 23rd IEEE International Conference on Image Processing* Phoenix, Arizona, USA, September 25-28, 2016 p4428
- [6] Babae M, Wolf T, Rigoll G 2016 *The 23rd IEEE International Conference on Image Processing* Phoenix, Arizona, USA, September 25-28, 2016 p704
- [7] Huang K K, Dai D Q, Ren C X, Lai Z R 2017 *IEEE Trans. Neural Netw.* **28** 1082
- [8] Zheng H, Tao D P 2015 *Neurocomputing* **162** 9
- [9] Cai S J, Zuo W M, Zhang L, Feng X C, Wang P 2014 *The 13th European Conference on Computer Vision* Zurich, Switzerland, September 6-12, 2014 p624
- [10] Yang J M, Yang M H 2017 *IEEE Trans. Pattern Anal. Mach. Intell.* **39** 576
- [11] Li J, Tao D 2012 *IEEE Trans. Image Process.* **21** 4830
- [12] Yan Y, Ricci E, Subramanian R 2014 *IEEE Trans. Image Process.* **23** 5599
- [13] Yang M, Zhang L, Shiu S C K, Zhang D 2013 *IEEE Trans. Neural Netw. Learn. Syst.* **24** 900
- [14] Yang M, Zhang L, Shiu S C K, Zhang D 2013 *Pattern Recognit.* **46** 1865
- [15] Georgakis C, Panagakis Y, Pantic M 2016 *IEEE Trans. Image Process.* **25** 2021
- [16] Zafeiriou S, Tzimiropoulos G, Petrou M, Stathaki T 2012 *IEEE Trans. Neural Netw.* **23** 526
- [17] Tenenbaum J B, De Silva V, Langford J C 2000 *Science* **290** 2319
- [18] Roweis S T, Saul L K 2000 *Science* **290** 2323
- [19] Belkin M, Niyogi P 2003 *Neural Comput.* **15** 1373
- [20] Lin B B, He X F, Zhang C Y, Ji M 2013 *J. Mach. Learn. Res.* **14** 2945
- [21] Lin B B, Yang J, He X F, Ye J P 2014 *Int. Conf. Mach. Learn.* 145
- [22] He X, Niyogi P 2004 *Advances in Neural Information Processing Systems* 153
- [23] He X, Cai D, Yan S 2005 *Proc. IEEE Int. Conf. Comput. Vis.* **2** 1208
- [24] Dornaika F, Raduncanu B 2013 *The 26th IEEE Conference on Computer Vision and Pattern Recognition* Portland, Oregon, USA, Jun 23-28, 2013 p862
- [25] Huang S C, Zhuang L 2016 *Neurocomputing* **218** 373
- [26] Wan M H, Yang G W, Gai S, Yang Z J 2017 *Multimed. Tools Appl.* **76** 355
- [27] Liang J Z, Chen C, Yi Y F, Xu X X, Ding M 2017 *IEEE Access* 17201
- [28] Wang R, Nie F P, Hong R C, Chang X J, Yang X J, Yu W Z 2017 *IEEE Trans. Image Process.* **26** 5019
- [29] Yuan X F, Ge Z Q, Ye L J, Song Z H 2016 *J. Chemometr.* **30** 430
- [30] Yan S C, Xu D, Zhang B Y, Zhang H J, Yang Q, Lin S 2007 *IEEE Trans. Pattern Anal. Mach. Intell.* **29** 40
- [31] Belkin M, Niyogi P 2013 *Neural Comput.* **15** 1373
- [32] Cortes C, Mohri M 2007 *Advances in Neural Information Processing Systems* Vancouver, Canada, December 3-8 2007 p305
- [33] Qiao L S, Chen S C, Tan X Y 2010 *Pattern Recognit.* **43** 331
- [34] Lai Z H, Wong W K, Xu Y, Yang J, Zhang D 2016 *IEEE Trans. Neural Netw. Learn. Syst.* **27** 723
- [35] Yin J, Lai Z H, Zeng W M, Wei L 2018 *Multimed. Tools Appl.* **77** 1069
- [36] Zhang Y P, Xiang M, Yang B 2016 *Neurocomputing* **173** 518
- [37] Lu G F, Jin Z, Zou J 2012 *Knowl-Based Syst.* **31** 119
- [38] Wei L, Xu F F, Wu A H 2014 *Knowl-Based Syst.* 136
- [39] Lou S J, Zhao X M, Chuang Y L, Yu H T, Zhang S Q 2016 *Neurocomputing* **173** 290
- [40] Yang J, Chu D L, Zhang L, Xu Y, Yang J Y 2013 *IEEE Trans. Neural Netw.* **24** 1023
- [41] Zheng J W, Yang P, Chen S Y, Shen G J, Wang W L 2017 *IEEE Trans. Image Process.* **26** 2408
- [42] Gao Q X, Wang Q Q, Huang Y F, Gao X B 2015 *IEEE Trans. Image Process.* **24** 5684
- [43] Zhang G Q, Sun H J, Xia G Y, Sun Q S 2016 *IEEE Trans.*

- Image Process.* **25** 4271
- [44] Ren C X, Dai D Q, Li X X, Lai Z R 2014 *IEEE Trans. on Image Processing* **23** 725
- [45] Yang M, Zhang L, Shiu S C K, Zhang D 2012 *IEEE Trans. on Information Forensics and Security* **7** 1738
- [46] Dalal N, Triggs B 2005 2005 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* San Diego, California, June 20-26 2005 p886
- [47] Tian S X, Bhattacharya U, Lu S J, Su B L, Wang Q Q, Wei X H, Lu Y, Tan C L 2016 *Pattern Recognit.* **51** 125
- [48] Tzimiropoulos G, Zafeiriou S, Pantic M 2012 *IEEE Trans. Pattern Anal. Mach. Intell.* **34** 2454
- [49] Ding C X, Choi J, Tao D C, Davis L S 2016 *IEEE Trans. Pattern Anal. Mach. Intell.* **38** 518
- [50] Weng D W, Wang Y L, Gong M M, Tao D C 2015 *IEEE Trans. Image Process.* **24** 2287
- [51] Yin J, Zeng W M, Wei L 2016 *Knowl-Based Syst.* **99** 112
- [52] Huang K K, Dai D Q, Ren C X 2017 *Pattern Recognit.* **62** 87
- [53] Liu Y, Gao Q X, Miao S, Gao X B, Nie F, Li Y S 2017 *IEEE Trans. Image Process.* **26** 684
- [54] Wang H, Nie F P, Huang H 2014 *The 31st International Conference on Machine Learning* Beijing, China, June 21-26, 2014 p1836
- [55] Learned-Miller E, Huang G B, Roy C A, Li H X, Hua G 2016 *Advances in Face Detection and Facial Image Analysis.* 189
- [56] Kumar N, Berg A C, Bellhumeur P N, Nayar S K 2009 *Proc. IEEE Int. Conf. Comput. Vis.* 365
- [57] Yang M, Zhang L, Yang J, Zhang D 2013 *IEEE Trans. Image Process.* 1753
- [58] Tang X, Feng G C, Cai J X 2014 *Neurocomputing* 402
- [59] Li F, Jiang M Y 2018 *Neural Process. Lett.* **47** 661
- [60] Tao D P, Guo Y N, Li Y T, Gao X B 2018 *IEEE Trans. Image Process.* **27** 325

Discriminative sparsity graph embedding based on histogram of rotated princial orientation gradients*

Tong Ying^{1)2)†} Shen Yue-Hong¹⁾ Wei Yi-Min¹⁾

1) (College of Communication Engineering, The Army Engineering University of PLA, Nanjing 210007, China)

2) (School of Information and Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China)

(Received 20 February 2019; revised manuscript received 29 July 2019)

Abstract

The unconstrained face images collected in the real environments include many complicated and changeable interference factors, and sparsity preserving projections (SPP) cannot well obtain the low-dimensional intrinsic structure embedded in the high-dimensional samples, which is important for subsequent sparse representation classifier (SRC). To deal with this problem, in this paper we propose a new method named discriminative sparsity graph embedding based on histogram of rotated principal orientation gradients (DSGE-HRPOG). Firstly, it extracts multi-scale and multi-directional gradient features of unconstrained face images by HRPOG feature descriptor and incorporates them into a discriminative feature dictionary of sparse representation classifier. Secondly, it seeks an optimal subspace of HRPOG feature dictionary in which the atoms in intra-classes are as compact as possible, while the atoms in inter-classes are as separable as possible by adopting the proposed DSGE dimensionality reduction method. Finally, an optimal algorithm is presented in which the low-dimensional projection and the sparse graph construction are iteratively updated, and the accuracy of unconstrained face recognition is further improved. Extensive experimental results on AR, Extended Yale B, LFW and PubFig databases demonstrate the effectiveness of our proposed method.

Keywords: unconstrained face recognition, manifold learning, sparsity preserving projection, histogram of oriented gradients

PACS: 42.30.Sy, 42.30.Tz, 42.30.-d

DOI: 10.7498/aps.68.20190224

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61703201, KYTYJJG206), the Natural Science Foundation of Jiangsu Province, China (Grant No.BK20170765), and the NIT fund for Young Scholar, China (Grant No.CKJB201602).

† Corresponding author. E-mail: tongying@njit.edu.cn