

文章编号:1007-5321(2019)01-0053-08

DOI:10.13190/j.jbupt.2018-045

基于块的二维核四元数主成分分析

陈北京^{1,2,3}, 杨建浩³, 范春年^{1,3}, 苏庆堂⁴, 王定成^{1,3}

(1. 南京信息工程大学 江苏省网络监控工程中心, 南京 210044;

2. 南京信息工程大学 江苏省大气环境与装备技术协同创新中心, 南京 210044;

3. 南京信息工程大学 计算机与软件学院, 南京 210044; 4. 鲁东大学 信息与电气工程学院, 烟台 264025)

摘要: 核四元数主成分分析(KQPCA)被成功应用于处理非线性四元数信号,然而,核矩阵维数太高使其对角化非常耗时,目前二维形式的KQPCA(2DKQPCA)并没有成功实现. 对此,采用基于块处理和并行计算的思想,提出基于块的2DKQPCA(B2DKQPCA),实现真正意义上的2DKQPCA. 基于时间复杂度、应用性能和分块矩阵应为四元数Hermitian矩阵的综合考虑,B2DKQPCA重点处理主对角线、反对角线和主对角线旁3个方向的小块. 然后,结合B2DKQPCA与RGB-D图像四元数表示方法,将B2DKQPCA应用于RGB-D目标识别领域. 在2个公开库上的实验结果表明,提出的基于列向B2DKQPCA的RGB-D识别算法优于现有基于主成分分析算法和基于卷积神经网络的一些算法.

关键词: 核主成分分析; 四元数; 彩色图像; RGB-D 目标识别

中图分类号: TN911.22

文献标志码: A

Block-Wise Two Dimensional Kernel Quaternion Principal Component Analysis

CHEN Bei-jing^{1,2,3}, YANG Jian-hao³, FAN Chun-nian^{1,3}, SU Qing-tang⁴, WANG Ding-cheng^{1,3}

(1. Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology,

Nanjing University of Information Science and Technology, Nanjing 210044, China;

3. School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China;

4. School of Information and Electrical Engineering, Ludong University, Yantai 264025, China)

Abstract: Currently, kernel quaternion principal component analysis (KQPCA) has been proposed and successfully applied to process linear quaternion signals. However, two dimensional version of KQPCA (2DKQPCA) has not been successfully implemented due to the quite time-consuming problem for diagonalizing the high dimensional kernel matrix. So, using the block-based idea and the parallel computing idea, the block-wise 2DKQPCA (B2DKQPCA) is proposed to implement 2DKQPCA really. After the overall consideration of computational complexity, application performance and quaternion Hermitian block, B2DKQPCA mainly processes the blocks of three directions: main-diagonal direction, anti-diagonal direction and side-diagonal direction. Then, B2DKQPCA is applied into RGB-D object recognition by combining B2DKQPCA and quaternion representation of RGB-D images. Experimental results on two publicly available datasets demonstrate that the proposed RGB-D object recognition algorithm based on the column direction B2DKQPCA outperforms some existing algorithms using principal component analysis

收稿日期: 2018-03-20

基金项目: 国家自然科学基金项目(61572258, 61772281, 61602253, 61672294); 江苏高校优势学科建设工程项目(PAPD); 江苏高校“青蓝工程”项目; 江苏省自然科学基金项目(17KJB520021)

作者简介: 陈北京(1981—), 男, 副教授, 博士生导师, E-mail: nbuimage@126.com.

and some existing algorithms using convolutional neural network.

Key words: kernel principal component analysis; quaternion; color image; RGB-D object recognition

在过去的 20 年中,彩色图像四元数表示方法 (QR, quaternion representation) 被广泛应用于处理彩色图像^[1-11]: 彩色图像的每个像素采用一个纯四元数来表示, 3 个通道作为它的 3 个虚部. 其优势就是将彩色图像视为一个矢量整体并考虑了 3 个通道之间的内在联系^[1-2, 10]. 但是, 在采用四维四元数表示三维彩色图像时, 其第 4 维会产生存储冗余和额外计算成本^[12]. 为此, Chen 等^[7]将色彩信息与深度信息相结合, 对该表示方法进行了改进, 提出了一种基于四元数的 RGB-D 图像表征方法 (RGBD_QR), 并将 RGBD_QR 应用于 RGB-D 目标识别领域. RGBD_QR 不仅额外考虑了在目标识别中具有重要作用的深度信息, 而且将深度信息作为实部引入四元数中以避免存储冗余. 因此, 笔者也将继续采用 RGBD_QR.

在模式识别、计算机视觉和多媒体信号处理中, 主成分分析 (PCA, principal component analysis) 是一个非常重要的主要针对高维数据的特征提取技术. 因此, 一方面, 基于传统的 QR 和四元数代数, Bihan^[4]、郎方年^[5]和 Sun 等^[6]先后提出了四元数 PCA (QPCA, quaternion PCA), 将传统的 PCA 推广应用于彩色图像, 随后 Sun 等^[6]还提出了双向二维 QPCA (BD2DQPCA, bidirectional two-dimensional QPCA); 另一方面, 为了有效地处理非线性情况, Schölkopf 等^[13]引入了核 PCA (KPCA, kernel PCA). KPCA 是 PCA 的非线性推广, 它采用核技巧将输入数据映射到高维特征空间, 从而将非线性问题转变为线性问题, 然后在映射空间进行传统的 PCA^[13].

最近, Chen 等^[7]提出了主要用于处理非线性四元数信号的核 QPCA (KQPCA, kernel QPCA), 并且采用了 RGBD_QR. 但是, 尽管 Chen 等^[7]给出了二维 KQPCA (2DKQPCA, two-dimensional KQPCA) 的定义, 但由于其核矩阵维数太高导致对角化对于普通计算机难以实现等原因, 并没有实现真正的 2DKQPCA, 而是采用近似的思路: 用平均行 (或平均列) 代表一幅图像. 这实质上实现的是变相的一维 KQPCA^[7], 破坏了图像的二维结构. 因此, 笔者将采用分块和并行计算的思路, 提出基于块的 2DKQPCA (B2DKQPCA, block-wise 2DKQPCA), 尽可能地考虑图像的二维结构, 实现真正的 2DKQPCA, 并结合

B2DKQPCA 与 RGBD_QR 应用于 RGB-D 的目标识别.

1 现有的 2DKQPCA

1.1 四元数及 RGBD_QR

作为复数的推广, 四元数在 1843 年由著名数学家 Hamilton 引入. 一个四元数由一个实部和 3 个虚部构成:

$$q = q_r + q_i i + q_j j + q_k k \quad (1)$$

其中: q_r, q_i, q_j 和 $q_k \in \mathbf{R}$, i, j 和 k 为 3 个满足如下运算规则的基本四元数单位.

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k,$$

$$jk = -kj = i, \quad ki = -ik = j$$

如果式 (1) 中实部 $q_r = 0$, 则称 q 为纯四元数. 四元数 q 的模定义如下:

$$|q| = \sqrt{(q_r)^2 + (q_i)^2 + (q_j)^2 + (q_k)^2} \quad (2)$$

为了解决现有彩色图像四元数表示方法存在的冗余问题, Chen 等^[7]提出新的 RGBD_QR: 通过引入对光照和色彩变化具有很强鲁棒性的深度信息作为四元数的实部分量, 采用四元数表示 RGB-D 图像. 令 $g(u, v)$ 为 1 个 RGB-D 图像函数, 其每个像素表示 1 个四元数:

$$g(u, v) = g_D(u, v) + g_R(u, v)i + g_G(u, v)j + g_B(u, v)k \quad (3)$$

其中 $g_D(u, v)$ 、 $g_R(u, v)$ 、 $g_G(u, v)$ 和 $g_B(u, v)$ 分别为像素点 (u, v) 的深度、红色、绿色和蓝色分量. 这里需要注意的是, 在将深度信息和色彩信息整合前, 深度信息的数值被归一化为 $[0, 255]$. RGBD_QR 至少在 2 个方面要优于传统的彩色图像四元数表示方法: ① 它不仅考虑了色彩信息而且同时考虑了在目标识别中具有重要作用的深度信息; ② 它将深度信息作为实部引入四元数中, 避免了存储冗余.

1.2 2DKQPCA

2DKQPCA 既可从行向也可从列向进行. 下面以文献[7]中定义的行向 2DKQPCA 为例, 列向类似. 行向 2DKQPCA 的主要思想就是先将四元数样本图像的行向量通过非线性映射函数映射到高维空间, 然后再进行行向二维 QPCA. 由于不需要映射后样本的显示形式以及为了避免高维空间的高复杂度计算, Chen 等^[7]采用了核技术: 用核矩阵代替映射后样本之间的协方差矩阵.

对大小为 $M \times N$ 的 2D 训练四元数数据样本图像 $X_s (s = 1, 2, \dots, S)$, 定义其四元数核矩阵 $\mathbf{K} = (k_{(s-1)M+m, (t-1)M+n})_{SM \times SM}$ 为

$$k_{(s-1)M+m, (t-1)M+n} = k(\mathbf{X}_{s,m}, \mathbf{X}_{t,n}), \quad 1 \leq s, t \leq S, 1 \leq m, n \leq M \quad (4)$$

其中: $\mathbf{X}_{s,m}$ 和 $\mathbf{X}_{t,n}$ 分别为第 s 个样本 \mathbf{X}_s 的第 m 个行向量和第 t 个样本 \mathbf{X}_t 的第 n 个行向量, $k(\mathbf{X}_{s,m}, \mathbf{X}_{t,n})$ 是四元数形式的核函数. 文献[7]中采用了如下所示的四元数多项式核函数:

$$k(\mathbf{X}_{s,m}, \mathbf{X}_{t,n}) = (\mathbf{X}_{s,m} \mathbf{X}_{t,n} + c)^b \quad (5)$$

其中 b 和 c 为实数参数.

在求解四元数矩阵 \mathbf{K} 的特征值之前, 对 \mathbf{K} 进行中心化:

$$\bar{\mathbf{K}} = \mathbf{K} - \mathbf{E}_{SM} \mathbf{K} - \mathbf{K} \mathbf{E}_{SM} + \mathbf{E}_{SM} \mathbf{K} \mathbf{E}_{SM} \quad (6)$$

其中 \mathbf{E}_{SM} 表示大小为 $SM \times SM$ 的所有元素值均为 $1/(SM)$ 的矩阵.

由于 $\bar{\mathbf{K}}$ 为四元数 Hermitian 矩阵, 其特征值和特征向量的求解可以采用文献[5-6]中介绍的方法. 限于篇幅, 这里不再赘述. 假设 $\mathbf{w}_l (l = 1, 2, \dots, r)$ 为 $\bar{\mathbf{K}}$ 的最大 r 个特征值所对应的特征向量, 则 2DKQPCA 投影矩阵为

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r] \quad (7)$$

对于一个新的样本 \mathbf{X} , 其投影特征为

$$\mathbf{Y} = \mathbf{W}^T \bar{\mathbf{K}}_X = (\mathbf{Y}_l)_{r \times 1} = \left(\sum_{\substack{1 \leq s \leq S \\ 1 \leq m \leq M}} w_{l, sM+m} \bar{k}(\mathbf{X}, \mathbf{X}_{s,m}) \right)_{r \times 1} \quad (8)$$

其中: $w_{l, sM+m}$ 为特征向量 \mathbf{w}_l 的第 $(sM + m)$ 个系数, $\bar{\mathbf{K}}_X$ 为经式(6)中心化后的核矩阵, 即

$$\bar{\mathbf{K}}_X = [\bar{k}(\mathbf{X}, \mathbf{X}_{1,1}), \bar{k}(\mathbf{X}, \mathbf{X}_{1,2}), \dots, \bar{k}(\mathbf{X}, \mathbf{X}_{S,M})]^T \quad (9)$$

其中: $\bar{k}(\mathbf{X}, \mathbf{X}_{s,m}) (s = 1, 2, \dots, S; m = 1, 2, \dots, M)$ 为经式(6)中心化后的新样本 \mathbf{X} 和所有 SM 个行向量训练样本 $\mathbf{X}_{s,m}$ 之间的四元数核系数, $(\cdot)^T$ 表示矩阵的转置运算.

2 B2DKQPCA

事实上, 2DKQPCA 核矩阵的维数通常很高, 尤其是当样本数量或者图像尺寸比较大时. 比如, 对于包括 S 个尺寸为 $M \times N$ 的二维四元数数据样本 \mathbf{X}_s 的训练集, 其行向量和列向 2DKQPCA 核矩阵的维数分别为 $SM \times SM$ 和 $SN \times SN$. 通常样本数量至少为 1 000, 图片尺寸至少为 100×100 , 这样核矩阵维

数至少为 $10^5 \times 10^5$. 如此高维数的核矩阵为接下来的特征值求解带来非常大的计算量. 因此, 为了解决高计算复杂度的问题, Chen 等^[7]采用平均行或平均列代表一幅图像的近似思想, 从而将核矩阵的维数减小为 $S \times S$, 计算量大大减少, 但是该思想严重破坏了图像的二维结构, 并没有实现真正意义上的 2DKQPCA.

为了尽可能地考虑图像的二维结构, 采用文献[14]针对二维 KPCA (2DKPCA) 的分块思想: 首先对核矩阵划分为一些小块, 然后对这些小块进行 2DKPCA, 最后将这些小块的投影特征整合作为原始核矩阵的投影特征. 但是, 如果考虑所有的分块, 将带来 2 个方面的问题: 一方面, 尽管分块思想使得每块的计算更简单, 但大量数量的分块导致时间复杂度还是很大; 另一方面, 由于核矩阵本身为四元数 Hermitian (或称自共轭) 矩阵, 所以除了主对角线小块为四元数 Hermitian 矩阵外, 其他分块都不是. 但是, 目前四元数矩阵的特征求解方法主要针对四元数 Hermitian 矩阵. 因此, 基于时间复杂度、应用性能和分块矩阵应为四元数 Hermitian 矩阵的综合考虑, 如图 1 所示, 满足自共轭性质的小块包括所有主对角线方向的小块 (纹理为右下方向斜线)、由主对角线旁方向的三角区域 (纹理为方格和散点) 构造出的四元数 Hermitian 矩阵小块以及由反对角线方向的小块 (纹理为竖线和左下方向斜线) 构造出的四元数 Hermitian 矩阵小块. 对于主对角线方向的小块, 已是四元数 Hermitian 矩阵, 因此直接处理即可. 对于主对角线旁方向, 以图 1 左上角第 1 个下三角块 (纹理为散点) 和上三角块 (纹理为方格) 为例, 将上三角块和下三角块沿对角线进行拼接得到四元数 Hermitian 矩阵小块, 以此类推, 可得主对角线旁方向的所有四元数 Hermitian 矩阵小块. 对于反对角线方向, 以一对小块 (右上角的第 1 个小块和左下角的最后 1 个小块) 为例, 将右上角小块中

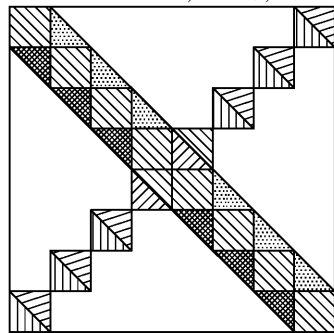


图 1 考虑图像块的示意图

斜线纹理的上三角部分(竖线纹理的下三角部分)与左下角小块中竖线纹理的下三角部分(斜线纹理的上三角部分)进行拼接得到四元数 Hermitian 矩阵小块,以此类推,可得反对角线方向的所有四元数 Hermitian 矩阵小块。

图1所示3个方向(主对角线、反对角线和主对角线旁)的小块进行2DKQPCA是可以完全独立的,因此为了进一步提升计算效率,考虑采用并行处理的思想。根据本课题组的计算机配置情况,采用单个CPU3个核并行计算,每个核处理一个方向上的所有小块。本课题计算机配置为ThinkStation P500, CPU为E5-2620v3,内存为16 GB。

3 B2DKQPCA 在 RGB-D 目标识别中的应用

作为PCA的一种改进算法,B2DKQPCA可以主要应用于包括人脸识别、虹膜识别和目标识别在内的模式识别领域,这些应用按照识别目标的数量可分为单目标识别和多目标识别。对于多目标识别,通常需要先进行多目标分割,随后对各个目标进行识别。也就是说,不管是单目标识别还是多目标识别,都需要对单个目标进行识别。因此,本节将详细介绍B2DKQPCA在RGB-D单目标识别中的应用,

提出一种基于B2DKQPCA的RGB-D单目标识别算法。在其他识别方面的应用可以参照该算法。

2DKQPCA可以用于处理四维四元数信号。而且,如前文所述,传统的QR在使用四维四元数表示只拥有3个分量的彩色图像时会导致存储冗余及额外计算负担。因此,采用Chen等^[7]提出的如式(3)所示的RGBD_QR,并将其与B2DKQPCA相结合,提出一种新的RGB-D目标识别算法,该算法的流程如图2所示。训练阶段和测试阶段的具体步骤如下所述(注意,这里是以行向2DKQPCA为例,列向2DKQPCA与之类似)。

1) 训练阶段

① RGB-D图像四元数表示。首先采用式(3)所示的RGBD_QR将 S 个尺寸为 $M \times N$ 的RGB-D训练样本 $\mathbf{X}_s (s=1, 2, \dots, S)$ 表示为四元数矩阵 \mathbf{X}_s ,然后视 \mathbf{X}_s 的每一行 $\mathbf{X}_{s,m} (m=1, 2, \dots, M)$ 为新的训练样本。

② 分核分块核矩阵计算。由图1所示,不考虑所有的小块,因此只需计算所考虑小块的核矩阵。而且,不同方向的小块采用不同CPU核并行计算。假设分块大小为 $L \times L$,对于第 n 个小块中的某个元素 $(x, y), 1 \leq x, y \leq L, 1 \leq n \leq \lceil SM/L \rceil, [\cdot]$ 表示取整,通过式(4)计算 $k_{x,y} = k(\mathbf{X}_{s',nL+x+M-s'M},$

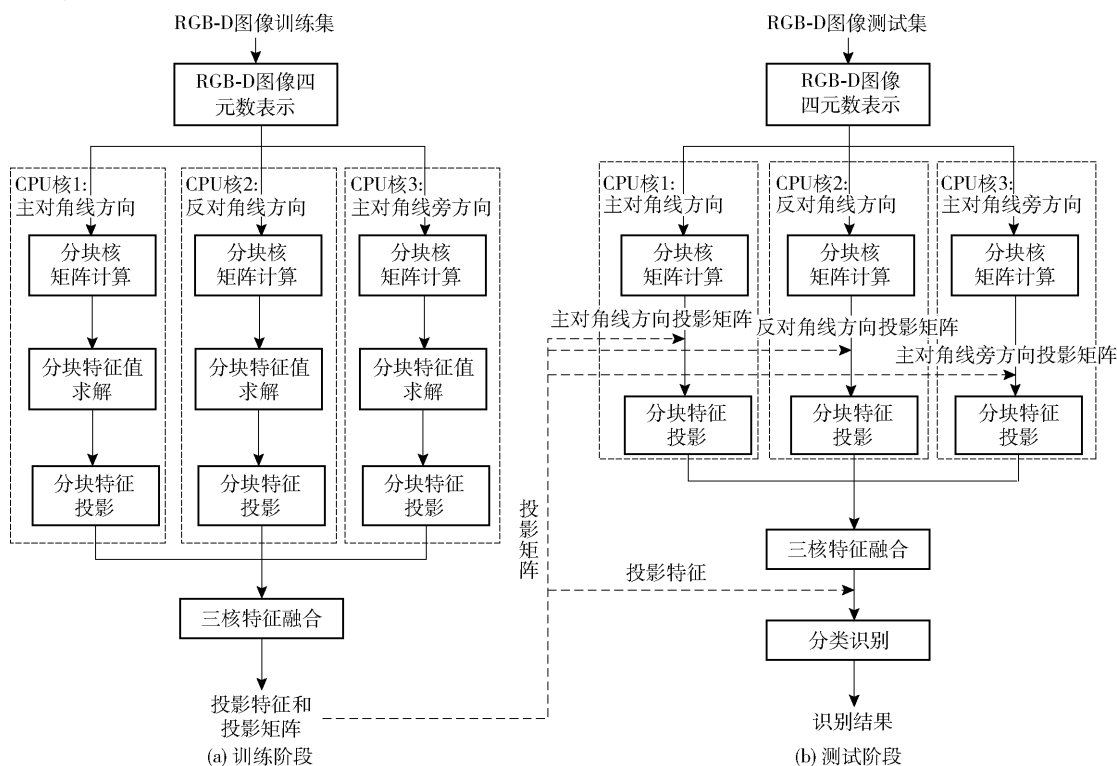


图2 提出的基于B2DKQPCA的RGB-D目标识别算法流程

$\mathbf{X}_{s', nL+x+M-s'M}^{'}$, 这里 $s' = \lceil (nL+x)/M \rceil$, $\lceil \cdot \rceil$ 表示往上取整, 从而得到所有主对角线上小块的核矩阵 $\mathbf{K}_n^{\text{dir1}}$. 同理, 采用 1.2 小节中的构造方法可得另 2 个方向的所有核矩阵块 $\mathbf{K}_n^{\text{dir2}}$ 和 $\mathbf{K}_n^{\text{dir3}}$.

③ 分核分块求解投影矩阵. 对②中获得的 3 个方向的所有核矩阵块按方向分核处理, 首先通过式 (6) 对其进行中心化, 然后对中心化后的所有核矩阵块求解特征值和特征向量, 取最大 r 个特征值所对应的特征向量, 进而由式 (7) 得到 2DKQPCA 投影矩阵 $\mathbf{W}^{\text{dir}}, \text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\}$.

④ 分核分块特征投影. 基于③中获得的 3 个方向上的投影矩阵 \mathbf{W}^{dir} , 分核分别通过式 (8) 对每个小块所对应的训练样本 $\mathbf{X}_{s,m}$ 进行投影, 得到 3 个方向的投影特征 $\mathbf{Y}_{s,m}^{\text{dir}} = \{\mathbf{Y}_{sM+m,h}^{\text{dir}} | h = 1, 2, \dots, r\}, \text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\}, s = 1, 2, \dots, S, m = 1, 2, \dots, M$.

⑤ 三核特征整合. 将 3 个核分别获取的投影特征进行整合. 比如, 对训练样本 \mathbf{X}_s , 其 M 行 3 个方向整合在一起的特征为 $\mathbf{Y}_s = \{\mathbf{Y}_{s,1}^{\text{dir1}}, \mathbf{Y}_{s,2}^{\text{dir1}}, \dots, \mathbf{Y}_{s,M}^{\text{dir1}}, \mathbf{Y}_{s,1}^{\text{dir2}}, \dots, \mathbf{Y}_{s,M}^{\text{dir3}}\}$. 然后, 将所有训练样本的投影特征 $\mathbf{Y}_s (s = 1, 2, \dots, S)$ 整合作为整个训练集的特征.

2) 测试阶段

① RGB-D 图像四元数表示. 首先采用式 (3) 所示的 RGBD_QR 将 T 个尺寸为 $M \times N$ 的 RGB-D 测试样本 $Z_t (t = 1, 2, \dots, T)$ 表示为四元数矩阵 \mathbf{Z}_t , 然后视 \mathbf{Z}_t 的每一行 $Z_{t,m} (m = 1, 2, \dots, M)$ 为新的测试样本.

② 分核分块核矩阵计算. 由于需要采用训练阶段获取的投影矩阵对测试样本进行投影, 因此这里所分小块的行数必须与训练阶段小块的行数一致, 即也为 L . 与训练阶段一样, 对于不同方向采用不同的 CPU 核. 假设分块大小为 $L \times P$, 对于第 n 个小块中的某个元素 $(x, y), 1 \leq x \leq L, 1 \leq y \leq P, 1 \leq n \leq \lceil SM/L \rceil$, 则可以通过式 (4) 计算 $\hat{k}_{x,y} = k(\mathbf{X}_{s', nL+x+M-s'M}^{'}, \mathbf{Z}_{t', nP+y+M-t'M}^{'})$, 这里 $s' = \lceil (nL+x)/M \rceil, t' = \lceil (nP+y)/M \rceil$, 从而得到所有主对角线上小块的核矩阵 $\hat{\mathbf{K}}_n^{\text{dir1}}$. 同理, 采用 1.2 小节中的构造方法可得另 2 个方向的所有核矩阵块 $\hat{\mathbf{K}}_n^{\text{dir2}}$ 和 $\hat{\mathbf{K}}_n^{\text{dir3}}$.

③ 分核分块特征投影. 对②中获得的所有核矩阵块按方向分核处理, 首先对核矩阵 $\hat{\mathbf{K}}_n^{\text{dir}} (\text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\})$ 通过类似于式 (6) 的方式进行中心化, 即

$$\bar{\mathbf{K}}_n^{\text{dir}} = \hat{\mathbf{K}}_n^{\text{dir}} - \mathbf{E}_L \hat{\mathbf{K}}_n^{\text{dir}} - \hat{\mathbf{K}}_n^{\text{dir}} \mathbf{E}_P + \mathbf{E}_L \hat{\mathbf{K}}_n^{\text{dir}} \mathbf{E}_P \quad (10)$$

其中: \mathbf{E}_L 是大小为 $L \times L$ 所有元素均为 $1/L$ 的矩阵, \mathbf{E}_P 是大小为 $P \times P$ 所有元素均为 $1/P$ 的矩阵; 然后采用训练阶段获取的投影矩阵 $\mathbf{W}^{\text{dir}}, \text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\}$, 通过式 (8) 对每个小块所对应的测试样本 $Z_{t,m}$ 进行投影, 得到 3 个方向的投影特征 $\hat{\mathbf{Y}}_{t,m}^{\text{dir}} = \{\hat{\mathbf{Y}}_{tM+m,h}^{\text{dir}} | h = 1, 2, \dots, r\}, \text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\}, t = 1, 2, \dots, T, m = 1, 2, \dots, M$.

④ 三核特征整合. 将 3 个核获取的特征进行整合. 比如, 对测试样本 Z_t , 其 M 行 3 个方向整合在一起的特征为 $\hat{\mathbf{Y}}_t = \{\hat{\mathbf{Y}}_{t,1}^{\text{dir1}}, \hat{\mathbf{Y}}_{t,2}^{\text{dir1}}, \dots, \hat{\mathbf{Y}}_{t,M}^{\text{dir1}}, \hat{\mathbf{Y}}_{t,1}^{\text{dir2}}, \dots, \hat{\mathbf{Y}}_{t,M}^{\text{dir3}}\}$. 然后, 将所有测试样本的投影特征 $\hat{\mathbf{Y}}_t (t = 1, 2, \dots, T)$ 整合作为整个测试集的特征.

⑤ 分类识别. 基于④中获取的测试集特征以及训练阶段获取的训练集特征, 采用基于四元数欧氏距离的最邻近分类器进行分类识别. 这里需要强调的是, 为了更好地测试 KQPCA 投影特征的性能, 采用简单的最邻近分类器. 这里也可以采用其他分类器, 如支持向量机、神经网络、贝叶斯等.

众所周知, 特征距离是最邻近分类器的关键. 以测试样本 Z_t 为例, 为了确定 Z_t 所在的类, 计算 Z_t 与所有训练样本 $\mathbf{X}_s (s = 1, 2, \dots, S)$ 之间的四元数欧氏距离:

$$d_{t,s} = \|\hat{\mathbf{Y}}_t - \mathbf{Y}_s\|_2 = \sqrt{\sum_{\text{dir} \in \{\text{dir1}, \text{dir2}, \text{dir3}\}} \sum_{m=1}^M \sum_{h=1}^r |\hat{\mathbf{Y}}_{tM+m,h}^{\text{dir}} - \mathbf{Y}_{sM+m,h}^{\text{dir}}|^2} \quad (11)$$

其中: $\|\cdot\|_2$ 表示文献 [6] 中定义的四元数向量欧氏距离, $|\cdot|$ 则为式 (2) 所示的四元数模运算.

然后采用最邻近分类器进行识别. 如果 $s^* = \arg \min_{1 \leq s \leq S} (d_{t,s})$, 则测试样本 Z_t 归属于训练样本 \mathbf{X}_{s^*} 所在的类.

4 实验结果与分析

下面将通过 RGB-D 目标识别实验来测试提出的 B2DKQPCA 的性能. 为了更好地展示算法的性能, 提出算法与现有一些基于 PCA 的算法 (基于四元数的 BD2DQPCA [6] 及 BD2DKQPCA [7]、传统的基于块的 2DKPCA (B2DKPCA, block-wise 2DKPCA) [15]) 进行了对比. BD2DQPCA 在文献 [6]

中应用于彩色人脸识别,在这里结合如式(3)所示的 RGBD_QR 进行 RGB-D 目标识别. B2DKPCA 在文献[15]中用于灰度人脸识别,在这里将 B2DKPCA 先独立处理 RGB-D 目标图像的每个分量,然后进行整合,从而实现 RGB-D 目标识别.

为了便于实验对比,采用与文献[7]相同的实验数据(相同的训练和测试样本). 实验数据为 2 个分别针对人脸和一般物体的公开 RGB-D 数据库: III-T-D 库^[16]和 CIN 2D/3D 库^[17]. 对于 III-T-D 库: ① 包括了 106 个个体的总共 4 065 幅 RGB-D 人脸图像,每个个体的图像数从 11 ~ 254 不等;② 它的人脸图像是在正常光照条件下具有姿势、表情和眼镜遮盖等变化;③ 包括 5 个文件夹,每个文件夹可以单独进行一次实验^[16],即每个个体的 4 幅图像用于训练,其他图像用于测试;④ 类似于文献[16],所有图像的尺寸归一化为 100×100 .

对于 CIN 2D/3D 库: ① 它包括 18 个类别的 163 个个体;② 对于每个个体,它包括 36 个不同视角的 RGB-D 图像,相邻 2 个视角相差 10° ;③ 实验中只考虑了每个类的前 2 个个体,而且没有考虑一些尺寸特别小(如 5×5)的图像,对于每个个体,随机选择 14 幅不同视角的图像进行训练,剩余的 22 幅图像进行测试;④ 所有图像的尺寸也归一化为 100×100 . 算法涉及的一些参数设置:式(5)定义的多项式核函数中的实数 b 和 c 分别取为 1 和 1.5, II-T-D 库尺寸为 $42\,400 \times 42\,400$ 的核矩阵的分块大小为 424×424 , CIN 2D/3D 库尺寸为 $50\,400 \times 50\,400$ 的核矩阵的分块大小为 504×504 .

图 3(a)给出各算法在这 2 个不同数据库上的平均识别率,其中 B2DKPCA^[15]和提出的 B2DKQPCA 都具有行向和列向 2 种算法: B2DKPCA_Row^[15]和 B2DKPCA_Col^[15]、B2DKQPCA_Row 和 B2DKQPCA_Col. 图 3(b)是为了更好地展示其中的 5 种算法随累积能量比变化的趋势. 由图 3 可以看出: ① 各算法的识别率一开始均随累积能量比的增大而提高,但当累积能量比达到 95% 时,识别率趋于稳定,这是因为图像特征之间通常是存在冗余的; ② 对于任意的累积能量比,提出的 B2DKQPCA_Col 算法优于其余对比算法,这主要归因于基于四元数的处理思想、非线性核方法以及图像二维结构的考虑等因素; ③ 对于 B2DKPCA^[15]和提出的 B2DKQPCA,列向算法均优于行向算法,这与文献[7]中取平均列算法优于取平均行算法的结果一

致. 由图 3 可知,当累积能量比达到 95% 时,算法趋于稳定. 因此,接下来上述一些基于 PCA 的算法采用最优累积能量比 95% 的结果与其余算法进行对比. 同时,为了测试深度信息的有效性,还与修改的提出算法 2DKQPCA_RGB_Row 和 2DKQPCA_R-GB_Col 进行了对比,这 2 种修改算法只考虑色彩信息而不考虑深度信息.

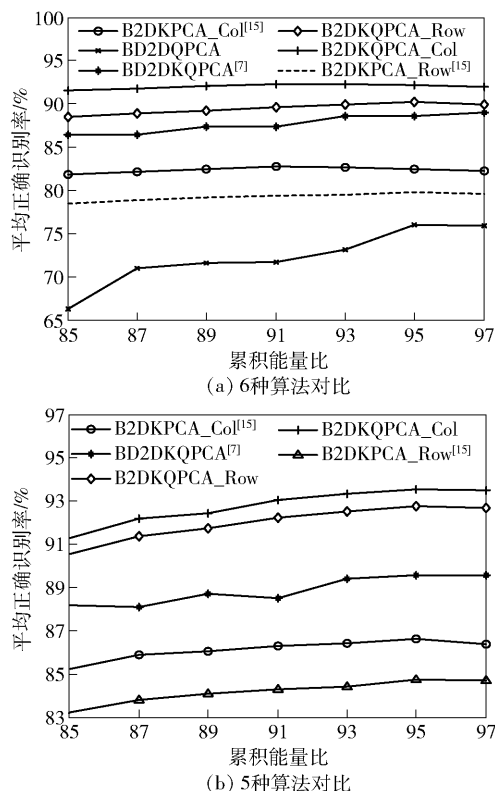


图 3 各算法在不同累积能量比下的平均正确识别率

针对 2 个数据库的结果分别如表 1 和表 2 所示. 表 1 同时提供了由 III-T-D 数据库创建者提出的 2 个算法(RISE 和 RISE + ADM)^[16]的结果;表 2 则同时给出了针对 CIN 2D/3D 库的现有 5 种算法(Ev2D3D^[17]、CNN-TRANSFER + DAE^[18]、CNN-SPM-RNN^[19]、Upgraded HMP^[20]和 Multi-Modal CNN^[21])的结果,这些结果均来源于相应文献. 由表 1 和表 2 可以看出: ① 提出的 B2DKQPCA_Col 算法在所有对比算法中效果最优,优于一些基于卷积神经网络(CNN, convolutional neural networks)的算法^[18-19,21]以及最近提出的同样基于 KQPCA 的 BD2DKQPCA 算法^[7]; ② 考虑深度信息较不考虑深度信息正确识别率能提高 1% 左右. 为了进一步分析深度信息对目标识别的有效性,并且 2 个数据库中深度图像的质量不高,因此通过对深度图像添加

不同程度的高斯噪声,然后采用提出的 B2DKQPCA_Col 算法进行识别,从而测试深度图像质量对识别的影响.

表 1 各算法针对 IIIT-D 库的正确识别率

算法	正确识别率/%
RISE ^[16]	82.78
RISE + ADM ^[16]	86.16
BD2DQPCA ^[6]	85.32
BD2DKQPCA ^[7]	88.59
B2DKPCA_Row ^[15]	79.83
B2DKPCA_Col ^[15]	82.50
B2DKQPCA_RGB_Row	89.07
B2DKQPCA_RGB_Col	91.26
B2DKQPCA_Row	90.23
B2DKQPCA_Col	92.18

表 2 各算法针对 CIN 2D/3D 库的正确识别率

算法	正确识别率/%
Ev2D3D ^[17]	82.80
Multi-Modal CNN ^[21]	88.40
Upgraded HMP ^[20]	91.00
CNN-TRANSFER + DAE ^[18]	91.30
CNN-SPM-RNN ^[19]	92.90
BD2DQPCA ^[6]	67.80
BD2DKQPCA ^[7]	91.04
B2DKPCA_Row ^[15]	89.98
B2DKPCA_Col ^[15]	90.80
B2DKQPCA_RGB_Row	91.58
B2DKQPCA_RGB_Col	92.64
B2DKQPCA_Row	92.77
B2DKQPCA_Col	93.56

图 4 展示了提出算法在 IIIT-D 库和 CIN 2D/3D 库上的结果. 可知,随着噪声强度不断增强,识别率不断降低,深度图像质量对识别具有一定的影响,但影响不大.

5 结束语

笔者采用分块和并行计算的思想,提出了 B2DKQPCA,实现了真正的 2DKQPCA. 同时,在 QR 中引入深度信息解决现有 QR 存在的存储冗余和额外计算开销问题,进而提出了结合 B2DKQPCA 与 RGBD_QR 的 RGB-D 目标识别算法. 实验表明,提

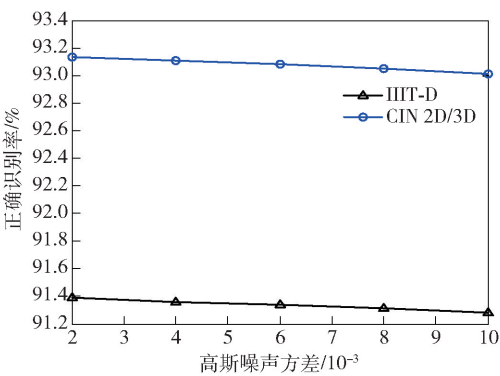


图 4 B2DKQPCA_Col 算法在不同程度噪声下的正确识别率

出的基于 B2DKQPCA_Col 的 RGB-D 目标识别算法优于现有的一些算法:① 优于一些非基于四元数的算法^[15-21],这主要归功于提出算法考虑了 RGB-D 图像 4 个分量的整体性;② 优于一些非核方法^[6,16-21],这主要是因为提出算法适用于处理高度非线性的现实场景中 RGB-D 目标识别问题^[14,22];③ 优于最近提出的基于 BD2DKQPCA 的算法^[7],这主要是因为提出算法充分考虑了图像的二维空间信息;④ 优于不考虑深度信息只考虑色彩信息的 B2DKQPCA_RGB_Row 和 B2DKQPCA_RGB_Col 算法,这主要归功于深度信息在识别中的有效性.

参考文献:

[1] Subakan O N, Vemuri B C. A quaternion framework for color image smoothing and segmentation[J]. International Journal of Computer Vision, 2011, 91(3): 233-250.

[2] Ell T A, Sangwine S J. Hypercomplex Fourier transforms of color images[J]. IEEE Transactions on Image Processing, 2007, 16(1): 22-35.

[3] Gai S. New banknote defect detection algorithm using quaternion wavelet transform [J]. Neurocomputing, 2016, 196: 133-139.

[4] Bihan N L, Sangwine S J. Quaternion principal component analysis of color images[C]//2003 10th IEEE International Conference on Image Processing (ICIP 2003). Barcelona: [s. n.], 2003: 809-812.

[5] 郎方年,周激流,闫斌,等. 四元数矩阵正交特征向量系的求解方法及其在彩色人脸识别中的应用[J]. 自动化学报, 2008, 34(2): 121-129.

Lang Fangnian, Zhou Jiliu, Yan Bin, et al. Obtain method of quaternion matrix orthogonal eigenvector set and its application in color face recognition[J]. Acta Automatica Sinica, 2008, 34(2): 121-129.

- [6] Sun Y F, Chen S Y, Yin B C. Color face recognition based on quaternion matrix representation[J]. *Pattern Recognition Letters*, 2011, 32(4): 597-605.
- [7] Chen B J, Yang J H, Jeon B, et al. Kernel quaternion principal component analysis and its application in RGB-D object recognition[J]. *Neurocomputing*, 2017(266): 293-303.
- [8] Saoud L S, Ghorbani R, Rahmoume F. Cognitive quaternion valued neural network and some applications[J]. *Neurocomputing*, 2017(221): 85-93.
- [9] 王金伟, 周春飞, 王水平, 等. 基于分数阶四元数傅里叶变换的彩色图像自适应水印算法[J]. *电子与信息学报*, 2016, 38(11): 2832-2839.
Wang Jingwei, Zhou Chunfei, Wang Shuiping, et al. Color image adaptive watermarking algorithm using fractional quaternion Fourier transform[J]. *Journal of Electronics and Information Technology*, 2016, 38(11): 2832-2839.
- [10] Yang H Y, Liang L L, Li Y W, et al. Quaternion exponent moments and their invariants for color image[J]. *Fundamenta Informaticae*, 2016, 145(2): 189-205.
- [11] Chen B J, Qi X M, Sun X M, et al. Quaternion pseudo-Zernike moments combining both of RGB information and depth information for color image splicing detection[J]. *Journal of Visual Communication and Image Representation*, 2017(49): 283-290.
- [12] Assefa D, Mansinha L, Tiampo K F, et al. The trinion Fourier transform of color images[J]. *Signal Processing*, 2011, 91(8): 1887-1900.
- [13] Schölkopf B, Smola A, Müller K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, 10(5): 1299-1319.
- [14] Eftekhari A, Forouzanfar M, Moghaddam H A, et al. Block-wise 2D kernel PCA/LDA for face recognition[J]. *Information Processing Letters*, 2010, 110(17): 761-766.
- [15] Sun N, Wang H X, Ji Z H, et al. An efficient algorithm for kernel two-dimensional principal component analysis[J]. *Neural Computing Applications*, 2008, 17(1): 59-64.
- [16] Goswami G, Vatsa M, Singh R. RGB-D face recognition with texture and attribute features[J]. *IEEE Transactions on Information Forensic and Security*, 2014, 9(10): 1629-1640.
- [17] Browatzki B, Fischer J, Graf B, et al. Going into depth: evaluating 2D and 3D cues for object classification on a new, large-scale object dataset[C] // 2011 IEEE International Conference on Computer Vision Workshops (ICCV2011). Barcelona: [s. n.], 2011: 1189-1195.
- [18] Tang J H, Jin L, Li Z C, et al. RGB-D object recognition via incorporating latent data structure and prior knowledge[J]. *IEEE Transactions on Multimedia*, 2015, 17(11): 1899-1908.
- [19] Cheng Y, Zhao X, Huang K, et al. Semi-supervised learning and feature evaluation for RGB-D object recognition[J]. *Computer Vision and Image Understanding*, 2015, 139: 149-160.
- [20] Bo L F, Ren X F, Fox D. Unsupervised feature learning for RGB-D based object recognition[C] // 13th International Symposium on Experimental Robotics. Québec City: [s. n.], 2013: 387-402.
- [21] Wang A R, Lu J W, Cai J F, et al. Large-margin multi-modal deep learning for RGB-D object recognition[J]. *IEEE Transactions on Multimedia*, 2015, 17(11): 1887-1898.
- [22] Zhang H, Parker L E. CoDe4D: color-depth local spatio-temporal features for human activity recognition from RGB-D videos[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, 26(3): 541-555.