

基于对称正定流形潜在稀疏表示分类算法*

陈凯旋, 吴小俊

(江南大学 物联网工程学院, 江苏 无锡 214122)

通讯作者: 吴小俊, E-mail: wu_xiaojun@jiangnan.edu.cn



摘要: 使用对称正定(symmetric positive definite, 简称 SPD)矩阵将视觉数据建模到黎曼流形(SPD 流形), 对于模式识别和机器学习中许多任务有较好的效果. 其中, 将基于稀疏表示的分类算法扩展到 SPD 流形上样本的分类任务得到了广泛的关注. 本文综合考虑了稀疏表示分类算法的特点以及 SPD 流形的黎曼几何结构, 通过核函数将 SPD 流形嵌入到再生核希伯特空间(reproducing kernel Hilbert space, 简称 RKHS), 分别提出了核空间潜在稀疏表示模型和潜在分类方法. 但是, 原始的视觉数据在核空间中并没有明确的表示形式, 这给核空间中的潜在字典更新带来了不便. Nyström 是一种可以近似表征核特征的方法. 因此, 我们利用该方法得到训练样本在 RKHS 中的近似表示, 以更新潜在字典和潜在矩阵. 最后, 通过在 5 个标准数据集上的分类实验, 验证了该方法的有效性.

关键词: 对称正定矩阵; 黎曼流形; 再生核希伯特空间; Nyström 方法; 潜在字典

中图分类号: TP181

中文引用格式: 陈凯旋, 吴小俊. 基于对称正定流形潜在稀疏表示分类算法. 软件学报, 2020, 31(8): 2530–2542. <http://www.jos.org.cn/1000-9825/5823.htm>

英文引用格式: Chen KX, Wu XJ. Latent sparse representation classification algorithm based on symmetric positive definite manifold. Ruan Jian Xue Bao/Journal of Software, 2020, 31(8): 2530–2542 (in Chinese). <http://www.jos.org.cn/1000-9825/5823.htm>

Latent Sparse Representation Classification Algorithm Based on Symmetric Positive Definite Manifold

CHEN Kai-Xuan, WU Xiao-Jun

(School of IoT Engineering, Jiangnan University, Wuxi 214122, China)

Abstract: Modeling visual data onto the SPD (symmetric positive definite) manifold using the SPD matrices has been proven to yield high discriminatory power for many visual classification tasks in the domain of pattern recognition and machine learning. Among them, generalising the sparse representation classification algorithm to the SPD manifold-based visual classification tasks has attracted extensive attention. This study first comprehensively reviews the characteristics of the sparse representation classification algorithm and the Riemannian geometrical structure of the SPD manifold. Then, embedding the SPD manifold into the Reproducing Kernel Hilbert Space (RKHS) via a kernel function. Afterwards, the latent sparse representation model and latent classification model in RKHS has been suggested, respectively. However, the original visual data in RKHS is implicitly described, which is impossible to perform the subsequent dictionary learning. To handle this issue, the Nyström method is utilized to obtain the approximate representations of the training samples in RKHS for the sake of updating the latent dictionary and latent matrix. Finally, the classification results obtained on five benchmarking datasets show the effectiveness of the proposed approach.

Key words: symmetric positive definite matrix; Riemannian manifold; reproducing kernel Hilbert space; Nyström method; latent dictionary

稀疏表示因其具有鲁棒性, 在数字图像处理和模式识别领域中有着广泛的应用. 它使用字典中原子的线性组合

* 基金项目: 国家自然科学基金(61672265, 61373055)

Foundation item: National Natural Science Foundation of China (61672265, 61373055)

收稿时间: 2018-02-04; 修改时间: 2018-08-06; 采用时间: 2018-11-16

来重构样本^[1-4],对于一个过完备字典,每一个测试样本有一个稀疏系数.基于稀疏表示的样本分类方法将样本的分类问题转化为基于稀疏系数的多个线性回归模型的分类问题,且已有的稀疏表示及字典学习的方法大多作用于向量空间^[4,5].本文则是将基于稀疏表示的分类方法用于解决黎曼空间中样本的分类任务.

在数字图像处理和模式识别领域中,样本的描述方法是否具有鲁棒性和判别性,对最后的实验结果起到至关重要的作用.其中,最为简单的就是先将图像进行灰度化处理,然后将图像矩阵展开为列向量.但是该方法会对导致向量的维度过高,且忽略了样本中像素点的几何信息,影响了判别性.最近,协方差描述因其具有较强的鲁棒性,在计算机视觉领域得到了广泛的关注^[6-8].

$$C = \frac{1}{r-1} \sum_{i=1}^r (f_i - \bar{f})(f_i - \bar{f})^T \quad (1)$$

若是对单幅图像进行描述,公式(1)中的 f_i 为图像中第 i 个像素点的特征向量,这些特征可能是灰度值、梯度值或者 Gabor 滤波值等. r 是该图像中像素点的个数, \bar{f} 是所有像素点特征向量的平均向量.若是对图像集进行描述,公式(1)中 f_i 为图像集中第 i 幅图像的灰度值, r 是该图像集中图像的个数, \bar{f} 是所有的图像灰度值的平均向量.协方差描述融合了多种特征,且通过协方差描述得到的矩阵为对称正定(symmetric positive definite,简称 SPD)矩阵,该 SPD 矩阵的维度只与特征向量 f_i 的维度有关,与图像中像素点的个数和图像集中的图像的个数都无关.由于协方差描述在建模时有平均化操作,所以协方差描述可以减少样本中的噪声带来的干扰.该描述方法已经被广泛地应用于弥散张量成像(diffusion tensor imaging,简称 DTI)、行人检测^[9]、人脸识别^[1-3]和纹理识别^[10].由 SPD 矩阵所张成的空间是一个 SPD 流形,即非线性的黎曼空间.如果将欧氏空间的度量方法和分类方法直接应用于 SPD 矩阵,则忽略了黎曼空间的黎曼度量和破坏了 SPD 流形的黎曼几何结构,将无法得到较好的分类效果.通过以上分析,若将欧氏空间中基于稀疏表示的分类方法用于解决 SPD 流形^[6,7,10]中样本的分类任务,则将面临如下几个方面的挑战.

- (1) 在 SPD 流形这个非线性的黎曼空间中,如何用字典中的原子(SPD 矩阵)去表达一个测试样本(SPD 矩阵),将面临着巨大的挑战.
- (2) 在欧氏空间中,样本多以向量的形式描述,通过 l_2 范数度量原样本和重构样本之间的误差.而对于 SPD 矩阵, l_2 范数将不再适合.
- (3) 字典中的原子也是 SPD 矩阵,所以在字典的更新过程中,将会涉及在 SPD 流形上的约束优化问题.

SPD 流形存在于一个非线性的黎曼空间中,阻碍了将欧氏空间的稀疏表示和字典学习的方法直接应用于 SPD 矩阵.综合考虑以上几个方面的困难和 SPD 流形的特性,有如下的几组策略可以将基于稀疏表示的分类方法应用于 SPD 矩阵,以实现 SPD 流形上样本的分类任务.

(1) 用统计度量计算重构误差

该策略主要是更换度量方法计算原样本和重构样本之间的误差,虽并没充分考虑 SPD 流形的几何特性,但是解决了重构误差的计算问题.在样本的重构过程中,将 SPD 矩阵视为欧氏空间的样本,用字典中原子的线性组合表示原样本,忽略了 SPD 流形的几何特性.在误差估计误差时,不是用 l_2 范数来度量原样本和重构样本,而是用一些经典的统计度量计算它们之间的误差,如 LogDet 散度(log-determinant divergence)^[11-14]和 Stein 散度(Stein divergence)^[6,7]等.文献[11,12]提出一种张量稀疏表示(tensor sparse representation,简称 TSR)方法,用字典中原子的线性组合重构样本,且用 LogDet 散度度量原样本和重构样本之间的误差.基于此框架,文献[13,14]提出了张量字典的更新方法(tensor dictionary learning,简称 TDL),文献[15]提出一种在线字典学习(generalized dictionary learning,简称 GDL)方法,在样本的重构过程中,用秩为 1 的矩阵的线性组合重构样本,用矩阵的 F 范数(Frobenius norm)计算重构样本与原样本之间的误差.以上两种方法——TSR 和 GDL,将稀疏表示方法与 SPD 流形相结合,实现了 SPD 流形上样本的分类任务.但在稀疏表示和字典学习的过程中,都是将 SPD 矩阵视为欧氏空间的样本,忽略了 SPD 流形的几何特性.在误差度量时,使用的并不是 SPD 流形上的黎曼度量,忽略了 SPD 流形的黎曼度量.文献[16]提出的黎曼稀疏表示(Riemannian sparse representation,简称 Rie_SR)方法考虑了 SPD 流形的黎曼度量.虽然也是用字典中原子的线性组合来重构样本,忽略了黎曼几何特性,但是在误差估计时,使用仿射不变黎

曼度量(affine invariant Riemannian metric,简称 AIRM)^[6,7,17,18]计算原样本和重构样本之间的误差.基于这一框架,文献[17]利用黎曼共轭梯度的方法实现了黎曼字典的更新(Riemannian dictionary learning,简称 Rie_DL).

(2) 考虑 SPD 流形的几何的特性

在 SPD 流形上,度量 SPD 流形两点之间的测地距离,除了仿射不变黎曼度,还有很多其他的黎曼度量方法,其中,LEM(log-Euclidean metric)^[8,10,19,20]是一个用于 SPD 流形的简洁有效的黎曼度量.该度量充分考虑 SPD 流形的几何特性,通过矩阵的对数映射(logarithm mapping),将 SPD 样本矩阵映射到其切空间中,切空间中的样本与原 SPD 流形上的样本是一一对应的关系.由于切空间是一个平面,可以视为欧氏空间,该度量用样本在切空间中的欧氏距离作为原 SPD 矩阵之间的测地距离,其中,对数映射是该度量方法的特别之处,文献[21,22]提出的方法就是首先将 SPD 矩阵映射到切空间中,对切空间中的样本进行字典学习和稀疏表示,然后通过切空间中字典原子的线性组合重构样本,用欧氏度量去计算重构误差.我们称这种方法为 LE_SR+LE_DL.该策略考虑了 SPD 流形的黎曼几何特性,将基于 SPD 流形上样本的分类问题转化为了基于切空间中样本的分类问题.

(3) 黎曼核稀疏表示

该策略与对数映射类似,首先将 SPD 矩阵映射到一个线性空间.但是该策略是通过核函数将样本映射到再生核希尔伯特空间(reproducing kernel Hilbert space,简称 RKHS),通过核空间字典原子的线性组合重构样本.文献[23]第一次提出将 SPD 矩阵映射到核空间以解决黎曼空间的稀疏表示问题,并提出了基于 Stein 散度^[6,7,24-26]的黎曼核,其中,Stein 散度为 $d_s = \log|(1/2)(X+Y)| - (1/2)\log|XY|$.我们称这种核空间稀疏表示的方法为 Ker_SR,并将基于这个框架的字典更新方法^[23,26]称为 Ker_DL.与该方法类似,文献[26,27]分别提出了基于 Jeffrey 核和 Log-Euclidean 核的核空间稀疏表示分类方法.

本文提出的方法与黎曼核稀疏表示方法类似,首先将 SPD 流形嵌入到 RKHS 中.但原黎曼核稀疏表示方法中字典原子的类别信息是预定义,所以我们在核空间的字典学习过程中嵌入潜在矩阵^[28],提出了核空间潜在稀疏表示分类算法,灵活地联系核空间中字典原子与类别信息,增加了核空间中字典的判别性,并最终实现了 SPD 流形上样本的分类任务.但由于核空间中样本没有具体的表示形式,给核空间中潜在字典和潜在向量的学习带来了不便.本文使用 Nyström 方法^[29,30]得到训练样本在核空间的向量表示,通过核空间样本的近似表示更新潜在字典和潜在向量.表 1 给出了以上方法在 SPD 流形上的稀疏表示和字典学习的对比.

Table 1 Comparisons of various methods on sparse representation and dictionary learning on SPD manifold

表 1 各方法在 SPD 流形上稀疏表示和字典学习的对比

方法	样本重构形式	是否黎曼度量	原子更新方式	是否满足 Mercer 条件
TSR+TDL ^[13,14]	线性表示(欧氏空间)	否(LogDet 散度)	欧氏空间	N/A
GDL ^[15]	线性表示(欧氏空间)	否(F-范数)	欧氏空间	N/A
Rie_SR+Rie_DL ^[17]	线性表示(欧氏空间)	是(AIRM)	黎曼空间	N/A
LE_SR+LE_DL ^[21]	线性表示(切空间)	是(LEM)	切空间	N/A
Ker_DLSR(Bregman) ^[26]	线性表示(核空间)	是(Stein/Jeffrey)	欧氏空间	一定条件下满足
Ker_DLSR(LogEK) ^[27]	线性表示(核空间)	是(LEM)	黎曼空间	满足
Proposed	线性表示(核空间)	是(LEM)	核空间	满足

本文第 1 节介绍与我们方法相关的工作:LEM 黎曼度量以及 Log-Euclidean 黎曼核、Nyström 方法和潜在矩阵.第 2 节介绍核空间潜在稀疏表示模型,并研究核空间的潜在字典和潜在矩阵学习方法.第 3 节给出各种算法在 5 个数据集上的实验对比及实验分析.第 4 节进行总结与展望.

1 相关工作

本节将简单介绍 LEM 黎曼度量以及相关的黎曼核、Nyström 方法和潜在矩阵.本文将使用以下的符号: S_n^+ 表示由 SPD 矩阵所张成的空间(SPD 流形),且每个 SPD 矩阵的维度是 $n \times n$; S_n 是由维度为 $n \times n$ 的对称矩阵所张成的空间,是在单位矩阵 $I_n \in R^{n \times n}$ 处的切平面;核映射 $\varphi: S_n^+ \rightarrow H$ 表示将 SPD 流形嵌入到 RKHS 中.

1.1 黎曼度量和黎曼核

在 SPD 流形上,较为常用的黎曼度量有 AIRM^[6,7,17,18]、Stein 散度^[6,7,24-26,31]、Jeffrey 散度^[6,7,26,31]和 LEM 度

量^[8,10,19,20,27].其中,LEM 度量是一种简洁有效的黎曼度量方法,对于任意给定的两个 SPD 矩阵 Sp_i, Sp_j ,通过 LEM 度量得到的距离 d_{LEM} 为

$$d_{LEM}(Sp_i, Sp_j) = \|\log(Sp_i) - \log(Sp_j)\|_F \tag{2}$$

公式(2)中的 $\log(\cdot)$ 为矩阵的对数操作, $\|\cdot\|_F$ 为矩阵的 F 范数.其中,矩阵的对数操作是将 SPD 矩阵映射到其所在 SPD 流形的切空间.

$$\varphi_{\log} = S_n^+ \rightarrow S_n, Sp \rightarrow \log(Sp), Sp \in S_n^+ \tag{3}$$

由 SPD 流形的这一几何特性,在 SPD 流形上,Log-Euclidean 核^[8]可以通过切空间中样本的内积来表示.对于任意给定的两个 SPD 矩阵 Sp_i, Sp_j :

$$k_{LogE}(Sp_i, Sp_j) = \langle Sp_i, Sp_j \rangle_{LogE} = \text{tr}(\log(Sp_i)\log(Sp_j)) \tag{4}$$

其中, tr 表示矩阵的迹.文献[5]证明了 Log-Euclidean 核满足 Mercer 定理.

1.2 Nyström方法

Nyström 方法^[29,30,32,33]可通过采样的方式构建一个低秩矩阵,去近似表示原核矩阵: $K \in R^{n \times n}$,降低核矩阵在计算中的运算代价;同时,也可以得到核空间中样本的向量表示.

- 首先,通过采样得到含有 m 个样本的集合 $\{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m\}$;
- 然后,构建一个低秩矩阵 $\hat{K}_r = K_b \hat{K}^\dagger K_b^k$,近似表示原核矩阵 K ,其中, $K_b = [k(x_i, \hat{x}_j)]_{n \times m}$, $\hat{K} = [k(\hat{x}_i, \hat{x}_j)]_{m \times m}$, \hat{K}^\dagger 为 \hat{K} 的伪逆矩阵, r 是矩阵 \hat{K} 的秩;
- 最后,对采样矩阵进行 SVD 分解: $\hat{K} = \hat{V} \hat{D}^{1/2} \hat{D}^{1/2} \hat{V}^T$, 可得到样本集中任一样本 x 的向量表示形式^[31]:

$$z_n(x) = \hat{D}_r^{-1/2} \hat{V}_r^T (k(x, \hat{x}_1), k(x, \hat{x}_2), \dots, k(x, \hat{x}_m))^T \tag{5}$$

其中, $\hat{D}_r = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_r)$ 和 $\hat{V}_r = (\hat{v}_1, \hat{v}_2, \dots, \hat{v}_r)$ 为最大的 r 个奇异值及与其对应的特征向量所构成的矩阵, $z_n(x)$ 为样本 x 的 r 维近似表示.对于样本集中的任意两个样本 x_i 和 x_j , 总有: $z_n(x_i)^T z_n(x_j) = [\hat{K}_r]_{i,j}$.

由于本文处理的样本为 SPD 矩阵,我们需用公式(4)中 Log-Euclidean 核来计算核矩阵,然后用上述介绍的方法得到 SPD 矩阵的向量表示.

1.3 潜在矩阵

基于潜在字典的稀疏表示分类算法^[28],在字典学习的过程学习一个潜在矩阵,该潜在矩阵描述的是潜在字典中原子与不同类别之间的关系.对于 M 类样本的分类问题,字典 $D = [d_1, d_2, \dots, d_N]$,潜在矩阵 W 是一个 $N \times M$ 的矩阵: $W = [Wr_1, Wr_2, \dots, Wr_N]$ 或 $W = [Wc_1, Wc_2, \dots, Wc_M]$.其中, $Wr_i = [W_{i,1}, W_{i,2}, \dots, W_{i,M}]$ 是潜在矩阵 W 的第 i 行,表示字典中第 i 个原子与 M 个类样本之间的关系(如图 1 所示); $Wc_j = [W_{1,j}, W_{2,j}, \dots, W_{N,j}]$ 是潜在矩阵 W 的第 j 列,表示第 j 类样本与 N 个字典原子之间的关系(如图 2 所示).与共享字典和类别相关字典不同,潜在字典不需要对字典中原子的类别信息预定义,而是通过潜在矩阵中的潜在向量去灵活的联系字典中原子和类别之间的关系.通过学习得到的字典具有较小的类内散度,且属于同类样本的潜在稀疏系数也更加的相似^[28].

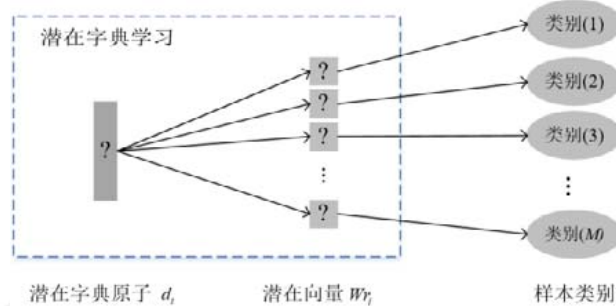


Fig.1 Relationship between the i -th dictionary atom and the sample categories

图 1 第 i 个字典原子和样本类别之间的关系

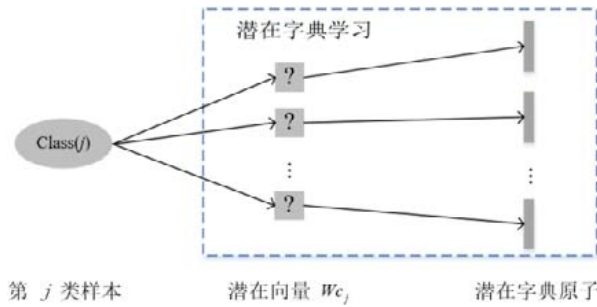


Fig.2 Relationship between the samples of j -th class and dictionary atoms
图 2 第 j 类样本和字典原子之间的关系

2 基于 SPD 流形潜在稀疏表示分类算法

本节将介绍基于 SPD 流形潜在稀疏表示分类算法.首先,将 SPD 矩阵映射到 RKHS 中,提出了核空间潜在稀疏表示模型及潜在分类模型;然后,使用 Nyström 方法得到 SPD 矩阵在 RKHS 中的近似表示来学习核空间中的潜在字典和潜在矩阵.

2.1 核空间潜在稀疏表示模型

本小节介绍核空间潜在稀疏表示模型及潜在分类模型.现有 M 类样本,测试样本 $Y \in S_n^+$ 和潜在字典: $D = \{d_1, d_2, \dots, d_N\}$, $d_i \in S_n^+$.该字典在核空间中的表示形式可以被写为: $\psi(D) = [\varphi(d_1), \varphi(d_2), \dots, \varphi(d_N)]$,潜在矩阵^[28]: $W \in R^{N \times M}$.核空间的潜在稀疏表示模型如下:

$$\tilde{v}_j = \arg \min_{v_j \in R^N} (\|\varphi(Y) - \psi(D) \text{diag}(Wc_j)v_j\|_2^2 + \lambda \|v_j\|_1) \tag{6}$$

\tilde{v}_j 是样本 $\varphi(Y)$ 相对于第 j 列潜在向量 Wc_j 的稀疏系数,该潜在向量表示核空间中第 j 类样本与字典中原子的关系.则该测试样本与第 j 类样本的重构误差为

$$Err_j = \|\varphi(Y) - \psi(D) \text{diag}(Wc_j)\tilde{v}_j\|_2^2 \tag{7}$$

通过公式(7),可以求得测试样本对于每一个类样本的重构误差.那么我们可以根据重构误差的大小来给测试样本进行分类,并将测试样本划分到重构误差最小的一类:

$$\text{identity}(Y) = \arg \min_j (Err_j), j = 1, 2, \dots, M \tag{8}$$

由于核映射 φ 没有具体的表达形式,所以公式(6)的第 1 项和公式(7)都需要利用核技巧改写.其中,公式(6)中的第 1 项可以改写为

$$\begin{aligned} \|\varphi(Y) - \psi(D) \text{diag}(Wc_j)v_j\|_2^2 &= \|\varphi(Y) - \sum_{i=1}^N Wc_{j,i}v_{j,i}\varphi(d_i)\|_2^2 \\ &= k(Y, Y) - 2v_j^T \text{diag}(Wc_j)K(Y, D) + v_j^T \text{diag}(Wc_j)K(D, D)\text{diag}(Wc_j)v_j \end{aligned} \tag{9}$$

公式(9)中的 $K(Y, D) = [k(Y, d_1), k(Y, d_2), \dots, k(Y, d_N)]^T$, $K = [k_{ij}]_{N \times N}$; $k_{ij} = k(d_i, d_j)$.其中,经过改写后的公式(6)的优化问题是核化 Lasso 问题^[23,26,27,34],也可以用 CVX 工具包^[23,35]进行优化求解.同理,公式(7)也可以通过利用核技巧改写求得重构误差.由公式(9)可知,核空间潜在稀疏表示分类方法无须知道核映射 φ 的具体形式,只需要知道测试样本与字典原子通过核函数所得到的向量 K 和字典原子之间通过黎曼核函数得到的核矩阵 K ,就可以得到测试样本 Y 的稀疏系数及对于每一个类样本的重构误差,然后根据重构误差的大小将测试样本分类.

2.2 核空间潜在字典学习模型

核空间的潜在字典和潜在矩阵^[28]在我们提出的分类算法中起着至关重要的作用,但是核映射 φ 没有明确的形式,阻碍了核空间中潜在字典和潜在矩阵的更新.为了能够得到核空间的潜在字典和潜在矩阵,我们使用

Nyström 方法得到训练样本在核空间的近似表示,通过样本的近似表示去学习核空间的潜在字典和潜在矩阵^[28].

现有 M 类样本,训练样本集为: $X=\{X_1, X_2, \dots, X_M\}$, 其中, $X_j=\{X_{j,1}, X_{j,2}, \dots, X_{j,m}\}$ 为第 j 类样本的子集,该子集中的每个样本都是 SPD 矩阵且样本的个数为 m . 对于该训练集合,可以使用第 1.2 节中的 Nyström 方法得到训练样本在 RKHS 中的近似表示: $Z(X)=[Z(X_1), Z(X_2), \dots, Z(X_M)]$, 其中, $Z(X_j)$ 为第 j 类的 m 个训练样本集合 X_j 的核空间近似表示集合. 在核空间的潜在字典学习模型中,需要被学习的有潜在字典: $D=[d_1, d_2, \dots, d_N]$ 和潜在矩阵 $W=[Wc_1, Wc_2, \dots, Wc_M]$, 其中, N 是字典中原子的个数, W 中的第 j 列 Wc_j 表示核空间中的第 j 类样本与字典 D 中原子的关系, 稀疏系数矩阵 $A=[A_1, A_2, \dots, A_M]$, 其中, A_j 为第 j 类样本子集的近似表示 $Z(X_j)$ 在字典 D 上的稀疏系数矩阵. 对于第 j 类样本,学习到的字典 D 和第 j 类潜在向量 Wc_j 可以很好地表示第 j 类的训练样本 $Z(X_j)$, 即 $Z(X_j) \approx D \text{diag}(Wc_j) A_j$, 核空间的潜在字典学习模型^[28]为

$$\left. \begin{aligned} \min_{A, D, W} \sum_{j=1}^M \|Z(X_j) - D \text{diag}(Wc_j) A_j\|_2^2 + \lambda_1 \|A_j\|_1 + \lambda_2 \|A_j - M_j\|_F^2 + \lambda_3 \sum_{j=1}^M \sum_{l \neq j} \sum_{n=1}^N \sum_{r \neq n} Wc_{j,r} (d_r^T, d_n)^2 Wc_{l,n} \\ \text{s.t. } Wc_{j,r} \geq 0, \forall j, r; \sum_r Wc_{j,r} = \delta, \forall r \end{aligned} \right\} \quad (10)$$

公式(10)里的 δ 是一个常量,为了平衡不同类的潜在向量 Wc_j 表达数据的能力. 对于不同类的潜在向量,潜在向量中值的和应该相等,即 $\sum_r Wc_{j,r} = \sum_n Wc_{l,n} = \delta, \forall j \neq l$. M_j 是与 A_j 相同尺寸的矩阵, M_j 中的每一列是矩阵 A_j 的平均向量. 公式的第 2 项为稀疏系数惩罚项,第 3 项是稀疏系数判别项,第 4 项为字典原子的不相关项. 学习得到的潜在字典和潜在向量之间的关系是相辅相成的,且核空间潜在稀疏表示模型也依赖潜在字典和潜在向量才能实现分类的任务.

2.3 核空间潜在字典学习

在公式(10)的目标函数中,需要学习的有潜在字典、潜在向量和训练样本在潜在字典上的稀疏系数. 对于该优化问题,使用 ADMM(alternating direction method of multipliers)算法,首先固定潜在向量,学习潜在字典和潜在稀疏系数,公式(10)的潜在字典学习模型可以重新写为

$$\min_{A, D} \sum_{j=1}^M \|Z(X_j) - D \text{diag}(Wc_j) A_j\|_2^2 + \lambda_1 \|A_j\|_1 + \lambda_2 \|A_j - M_j\|_F^2 + \lambda_3 \sum_{j=1}^M \sum_{l \neq j} \sum_{n=1}^N \sum_{r \neq n} Wc_{j,r} (d_r^T, d_n)^2 Wc_{l,n} \quad (11)$$

对于该优化问题,使用 ADMM 算法对公式(11)中的字典 D 和稀疏系数 A 进行迭代更新. 当字典 D 固定时,逐类求解稀疏系数,第 j 类样本稀疏系数的优化问题为

$$\min_{A_j} \|Z(X_j) - D \text{diag}(Wc_j) A_j\|_2^2 + \lambda_1 \|A_j\|_1 + \lambda_2 \|A_j - M_j\|_F^2 \quad (12)$$

公式(12)的优化问题可以像文献[28]中那样使用 IPM(iterative projection method)^[28,36]求解. 当稀疏系数固定时,我们重新定义潜在稀疏系数 $L=[L_1, L_2, \dots, L_M]$, 其中, $L_j = \text{diag}(Wc_j) A_j$. 字典更新模型为

$$\min_D \|Z(X) - DL\|_F^2 + \lambda_3 \sum_{j=1}^M \sum_{l \neq j} \sum_{n=1}^N \sum_{r \neq n} Wc_{j,r} (d_r^T, d_n)^2 Wc_{l,n} \quad (13)$$

公式(13)是为了更新核空间中的潜在字典 D . 令 $\Gamma = LL^T, A = Z(X)L^T$, 若要更新核空间潜在字典 D 中的第 n 个原子,公式(13)可以重写为^[28]

$$\min_{d_n} \|D^T D \Gamma - 2D^T A\|_F^2 + 2\lambda_3 \sum_{r \neq n} (d_r^T d_n)^2 \sum_{j=1}^M \sum_{l \neq j} Wc_{j,r} Wc_{l,n} \quad (14)$$

该问题是一个收敛问题. 基于文献[37]提出的字典更新方法,公式(14)的解为

$$u = (\Gamma_{n,n} I + 2\lambda_3 I_Q)^{-1} (A_n - D \Gamma_n - 2\lambda_3 Q D_n) \quad (15)$$

$$d_n = u / \|u\|_2 \quad (16)$$

公式(15)中的 I 为单位矩阵, $Q = \sum_{r \neq n} d_n d_r^T \sum_{j=1}^M \sum_{l \neq j} Wc_{j,r} Wc_{l,n}$, I_Q 是一个对角矩阵,且对角元素与 Q 的对角元素相同; $\Gamma_{n,n}$ 是 Γ 的第 n 行 n 列元素; A_n, Γ_n 和 D_n 是 A, Γ 和 D 的第 n 个列向量. 公式(16)是字典原子的归一化操作. 利用公式(15)和公式(16)对字典中的原子进行逐一更新,通过迭代更新得到核空间的潜在字典 D .

2.4 核空间潜在矩阵学习

对于公式(10)中核空间潜在向量的优化问题^[28],使用 ADMM 方法,固定潜在字典和潜在稀疏系数,更新潜在向量.潜在矩阵中,每一类的潜在向量都有约束条件,第 j 类的潜在向量的约束为 $\sum_r \mathcal{W}c_{j,r} = \delta$,所以在更新潜在矩阵时,需要逐类更新潜在向量,并将更新的潜在向量合并得到最终的潜在矩阵.优化第 j 类潜在向量的目标函数^[28]为

$$\left. \begin{aligned} \min_{\mathcal{W}c_j} \sum_{j=1}^M \|Z(X_j) - D \text{diag}(\mathcal{W}c_j) A_j\|_2^2 + 2\lambda_3 \sum_{n=1}^N \mathcal{W}c_{j,n} \sum_{r \neq n} (d_r^T, d_n)^2 \sum_{l \neq j}^M \mathcal{W}c_{l,r} \\ \text{s.t. } \mathcal{W}c_{j,r} \geq 0, \forall j, r; \sum_r \mathcal{W}c_{j,r} = \delta, \forall r \end{aligned} \right\} \quad (17)$$

公式(17)是一个受限二次规划问题.令 $A_j(n)$ 表示稀疏系数矩阵 A_j 的第 n 行; $B = [B_1, B_2, \dots, B_M]^T$, 其中, $B_n = \sum_{r \neq n} (d_r^T, d_n)^2 \sum_{l \neq j}^M \mathcal{W}c_{l,r}$; $a = \text{vec}(Z(X_j))$ 表示将矩阵 $Z(X_j)$ 展开成列向量; $R_n = d_n A_j(n)$ 和 $R = [\text{vec}(R_1), \text{vec}(R_2), \dots, \text{vec}(R_N)]$. 公式(17)可以被简化^[28,36]为

$$\left. \begin{aligned} \min_{\mathcal{W}c_j} \|a - R \mathcal{W}c_j\|_F^2 + 2\lambda_3 B^T \mathcal{W}c_j \\ \text{s.t. } \mathcal{W}c_{j,r} \geq 0, \forall j, r; \sum_r \mathcal{W}c_{j,r} = \delta, \forall r \end{aligned} \right\} \quad (18)$$

基于 IPM^[28,36]框架,公式(18)的优化问题可以通过迭代更新求解.其中,第 j 类的潜在向量的第 $t+1$ 次迭代的结果 $\mathcal{W}c_j^{t+1}$ 为^[28,36]

$$T_0 = \mathcal{W}c_j^t - (R^T (R \mathcal{W}c_j^t - a) + \lambda_3 B) / \sigma \quad (19)$$

$$T_1 = T_0 - \left(\sum_n T_{0,n} - \delta \right) / N \quad (20)$$

$$T_2 = \max(T_1, 0) \quad (21)$$

$$\mathcal{W}c_j^{t+1} = T_2 / \sum_n T_{2,n} \cdot \delta \quad (22)$$

公式(19)中的 σ 为 IPM 算法^[28,36]的步长, \max 为取大操作,且当 T_1 为负数时, $T_2 = 0$. 公式(22)得到了第 $t+1$ 次迭代的第 j 类潜在向量.算法 1 是核空间潜在字典和潜在矩阵学习的详细步骤.

算法 1. 核空间潜在字典和潜在矩阵学习.

输入:

- 初始化的潜在矩阵 W , 字典学习时的迭代次数;
- 训练样本集 $X = \{X_1, X_2, \dots, X_M\}$, 其中, $X_j = \{X_{j,1}, X_{j,2}, \dots, X_{j,m}\}$ 为第 j 类样本的子集.

输出:

- 更新后的潜在矩阵;
 - 核空间的潜在字典.
- 1) 使用 Nyström 方法求得样本在核空间的近似表示;
 - 2) 学习潜在字典(循环至收敛):
 - 使用公式(12)逐类更新稀疏系数;
 - 使用公式(13)更新潜在字典;
 - 3) 使用公式(17)逐类更新潜在矩阵中的潜在向量;
 - 4) 如果公式(10)中目标函数的值足够小或者迭代次数达到,输出潜在字典和潜在矩阵;否则返回第 2)步.

通过核空间训练样本的近似表示,我们学习到了核空间的潜在字典和潜在矩阵.在核空间潜在在稀疏表示模型中,核映射 ϕ 没有具体的形式,为了得到解析解,将公式(6)和公式(7)中带 ϕ 的项通过核技巧转换成公式(9)中核内积 $K(Y, D)$ 和 $K(D, D)$ 的形式,以求得对于不同类潜在向量的稀疏系数及重构误差.在我们的方法中,由于使用核空间的近似表示求得了字典 D 和潜在矩阵 \mathcal{W} ,且核空间是高维的线性空间,所以可用测试样本的近似表示 $Z(Y)$ 和学习到的字典 D 的内积去近似的表示公式(9)中的 K ,即 $K(Y, D) \cong D^T Z(Y)$,用学习到的字典 D 中原子的内积去

近似表示公式(9)中的 K ,即 $K(D,D)\cong D^T D$.算法 2 给出了基于 SPD 流形潜在稀疏表示分类算法的详细步骤.

算法 2. 基于 SPD 流形潜在稀疏表示分类算法.

输入:

- 训练样本集 $X=\{X_1,X_2,\dots,X_M\}$,其中, $X_j=\{X_{j,1},X_{j,2},\dots,X_{j,m}\}$ 为第 j 类样本的子集;
- 测试样本 $Y \in S_n^+$.

输出:测试样本 $Y \in S_n^+$ 的类别.

- 1) 利用算法 1 学习核空间的潜在字典和潜在矩阵;
- 2) 使用公式(6)和公式(9)计算测试样本在不同类潜在向量条件下的稀疏系数;
- 3) 使用公式(7)和公式(9)计算测试样本对于不同类样本的重构误差;
- 4) 将样本归为重构误差最小的一类.

3 实验结果与分析

为了验证我们的方法的有效性,我们在 5 个基准数据集上进行测试,他们分别应用于纹理识别、病毒细胞识别、材质识别、人脸识别和对象分类.这 5 个数据集分别是 Brodatz 纹理数据集^[16,17]、Virus 病毒细胞基准数据集^[28]、UIUCM 基准数据集^[38]、YouTube Celebrities(YTC)基准数据集^[8]和 ETH-80 基准数据集^[39],其中,后两个数据集常被用于基于图像集的分类任务中.在每个数据集上,我们分别做了 10 次迭代实验,取 10 次实验识别率的平均值和标准差作为最终的结果.

在 5 个数据集上,我们与已有的一些在 SPD 流形上使用稀疏表示的分类算法作对比,例如 Rie_SR+Rie_DL^[17]、LE_SR+LE_DL^[21]和 Ker_DLSR(LogEK)^[27].同时,在这 5 个数据集上,我们共同对比了以下几种方法.

- (1) Frob_SR+Frob_DL^[17]:将 SPD 矩阵视为欧氏空间的点,用字典原子的线性组合来重构样本,并用矩阵的 F 范数^[15]来度量重构样本的误差.
- (2) CDL-LDA(covariance discriminant learning)^[8]:该方法是通过黎曼核将 SPD 矩阵映射到核空间,然后利用 LDA(linear discriminant analysis)方法进行学习和分类.

对于 Brodatz,Virus 和 UIUCM 这 3 个数据集,我们还与文献[40]中给出的 3 种纹理识别的方法做对比,它们分别是:

- 1) COV-SVM:该方法将对象建模成 SPD 矩阵,并用 SVM 分类器进行分类.
- 2) Gau-SVM:该方法将对象建模成高斯模型,并用 SVM 分类器进行分类.
- 3) RoG-SVM:RoG 是文献[40]提出一种样本描述模型,该方法将样本建模成 RoG 模型^[40],并用 SVM 分类器进行分类.

对于上述的 SVM 方法,我们使用 LIBSVM 工具包^[41]实现“one vs all”分类器,该工具包中的参数“-c -p”与文献[40]中一致,分别设为 4 和 10^{-5} ,其余参数保持工具包的默认设置.

对于 YTC 和 ETH-80 这两个图像集的数据集,我们与一些经典的基于图像集的分类算法进行对比,它们是:

- 1) PML(projection metric learning)^[42]:该算法首先将图像集建模为 Grassmann 流形上的点,然后通过学习得到一个正交的投影矩阵,使得投影空间的样本具有较大的类间散度和较小的类内散度.
- 2) GDA(Grassmann discriminant analysis)^[43]:该算法与 PML 一样首先将图像集建模为 Grassmann 流形上的点,然后将 Grassmann 流形嵌入到一个高维的希尔伯特空间,再通过 Fisher 判别准则去学习一个映射,从而使得原始的流形被映射到一个维度较低同时鉴别性更好的空间.
- 3) SPDML(SPD manifold leaning)^[6,7]:该方法首先将样本建模为 SPD 矩阵,然后通过学习得到一个投影矩阵,使得原始的 SPD 流形被投影到一个维度更低且更具有判别性的空间.

如第 2.1 节所述,潜在矩阵 $W=[W_{r_1},W_{r_2},\dots,W_{r_N}]$,潜在向量 W_{r_i} 表示第 i 个字典原子与样本类别之间的关系.在潜在字典的初始化过程中,我们通过直接从训练样本中抽取样本作为字典的原子,假设潜在向量中只有一个非 0 元素(例如等于 1)来初始化潜在向量 W_{r_i} ,若 $W_{r_i,j}=1$,那么字典中的第 i 个原子应用第 j 类样本中随机抽取来

初始化潜在字典.在所有的实验中,通过交叉验证来确定公式(10)中的 λ_1, λ_2 和 λ_3 的值.由于本文的实验和文献[27]的方法都是基于 Log-Euclidean 黎曼核,为了实验数据具有公正性,本文给出的这两个方法的实验数据,都是基于 LEM 度量的指数核^[27]所得出的实验结果.根据文献[33]中的描述,Nyström 方法的算法复杂度为 $O(rmn)$,这里的 r 为矩阵的秩, m 为采样样本的数量.潜在字典学习和潜在向量的复杂度为 $O(jn^2)$, j 为字典学习时的迭代次数;潜在稀疏表示模型的算法复杂度为 $O(n^2)$,因此,算法的整体复杂度为 $O((j+1)n^2)$.

3.1 图像分类

Brodatz 是一个纹理数据库^[16,17],在该数据集中包含 112 张纹理图片,选择前 20 张纹理图片作为该实验的样本集.在实验中,将每张图片的尺寸调整为 500×500 ,并将每张图片分成大小相等的 25 个方块,每个块的尺寸为 100×100 ,其中每张图片中,任意选择 10 块用作训练样本,剩下的 15 块用作测试样本.为了使得每个块可以用协方差描述表示,对块中的每个像素点的提取一个 5 维的特征描述.

$$F_{Brodatz}(x,y)=[x,y,I,abs(I_x),abs(I_y)] \quad (23)$$

公式(23)中,前 2 维的特征是该像素点在该图片块中的横、纵坐标,后面 3 维的特征分别是该像素点的灰度值、 x 轴和 y 轴方向的梯度.最后,每个图像块被描述为一个 5×5 的协方差矩阵.在该数据集上,通过交叉验证得: $\lambda_1=0.05, \lambda_2=0.5$ 和 $\lambda_3=0.05$.表 2 给出了各个算法在 Brodatz 数据集上的平均结果.

Virus 病毒数据集^[33]包含 15 个类别的病毒细胞图片,每个类别包含 100 张图片,每张图片的尺寸为 41×41 ,共 1 500 张图片.在每一类样本中,我们将 100 张样本随机分成 10 份,从每份中随机选择 8 张图片作为训练样本,剩下的 2 张作为测试样本.为了可以使用协方差矩阵描述图片样本,我们用 Gabor 滤波^[23,27]为每个像素点提取特征值.

$$F_{Virus}(x,y)=[G_{0,0}(x,y),\dots,G_{0,7}(x,y),\dots,G_{4,7}(x,y)] \quad (24)$$

公式(24)为每个像素点提取了 40 维(5 尺度,8 方向)的特征值,最后得到的协方差矩阵的维度是 40×40 .在该数据集上,通过交叉验证得: $\lambda_1=0.05, \lambda_2=0.01$ 和 $\lambda_3=0.01$.表 2 给出了各个算法在 Virus 数据集上的平均结果.

Table 2 Average results of algorithms on Brodatz, Virus and UIUCM datasets

表 2 各算法在 Brodatz, Virus 和 UIUCM 数据集上的平均结果

算法	Brodatz	Virus	UIUCM
COV-SVM ^[37]	65.23±2.14	51.00±7.71	23.70±2.40
Gau-SVM ^[37]	71.73±1.95	52.00±9.58	26.02±2.16
RoG-SVM ^[37]	73.07±1.96	59.33±8.72	28.98±2.83
CDL-LDA ^[8]	79.30±1.60	59.67±6.93	43.15±4.39
Frob_SR+Frob_DL ^[17]	70.67±1.98	54.00±7.83	42.41±5.25
LE_SR+LE_DL ^[21]	78.20±1.64	55.67±7.04	45.46±4.06
Ker_SR+Random_DL ^[27]	88.44±1.26	56.33±6.37	37.04±3.57
Proposed	90.22±1.07	62.67±7.83	47.69±3.06

UIUCM 数据集包含 18 个类别的材质图片,每个类包含 12 张样本图片,每张图片的尺寸大小不唯一.在每类样本中,我们随机抽取 6 张样本用于训练,剩余的 6 张样本用测试.为了使得每个样本可以用协方差描述表示,先将彩色图片转换为灰度图,对其中每个像素点提取 45 维的特征描述.

$$F_{UIUCM}(x,y)=[x,y,I,abs(I_x),abs(I_y),G_{0,0}(x,y),\dots,G_{0,7}(x,y),\dots,G_{4,7}(x,y)] \quad (25)$$

公式(25)中的特征为公式(23)和公式(24)特征的组合,最后得到的协方差矩阵维度为 45×45 .在该数据集上,通过交叉验证得: $\lambda_1=0.005, \lambda_2=0.5$ 和 $\lambda_3=0.5$.表 2 给出了各种算法在 UIUCM 数据集上的平均结果.

3.2 图像集分类

相比于单幅图像,图像集包含了更多的有效信息,例如光照、姿态和遮挡等.最近,基于图像集的分类问题在计算机视觉和模式识别领域获得了广泛的关注.所以,为了验证我们方法的泛化性,本文也将我们的方法应用于图像集的分类任务,分别在 ETH-80 和 YTC 两个数据集上与一些经典的基于图像集的分类算法作对比.

ETH-80 数据集^[36]中含有 8 个类别,每个类别有 10 个图像集,每个图像集中有 41 张从不同角度拍摄的照片.

我们把每个图片的尺寸调整为 20×20 , 并且在每一类图像集中, 任意选 5 个图像集用作训练, 剩下的用作测试. 在该数据集上, 通过交叉验证得: $\lambda_1=0.05, \lambda_2=0.001$ 和 $\lambda_3=0.001$. 表 3 为各算法在 ETH-80 上的平均结果.

Table 3 Average results of algorithms on ETH-80 dataset

表 3 各算法在 ETH-80 数据集上的平均结果

算法	ETH-80
PML ^[38]	90.00±3.07
GDA ^[39]	93.25±4.80
SPDML ^[6,7]	90.50±3.87
CDL-LDA ^[8]	93.75±3.43
Frob_SR+Frob_DL ^[17]	93.25±4.45
Rie_SR+Rie_DL ^[17]	88.75±4.60
LE_SR+LE_DL ^[21]	95.12±4.17
Ker_SR+Random_DL ^[27]	94.88±4.17
Proposed	96.00±3.75

YouTube Celebrities(YTC)数据集^[8]包含 47 个类, 由于每一类的对象所包含的图像集的个数不一样, 所以我们从每一类图像中随机抽取图像制作成 9 个图像集, 任选 3 个图像集作训练, 剩下的 6 个图像集用作测试, 图像集中每个图片的尺寸调整为 20×20 . 在该数据集上, 通过交叉验证得: $\lambda_1=0.05, \lambda_2=0.05$ 和 $\lambda_3=0.005$. 表 4 为各算法在 YTC 上的平均结果.

Table 4 Average results of algorithms on YTC dataset

表 4 各算法在 YTC 数据集上的平均结果

算法	YTC
PML ^[38]	67.62±3.32
GDA ^[39]	65.78±3.34
SPDML ^[6,7]	61.57±3.43
CDL-LDA ^[8]	68.72±2.96
Frob_SR+Frob_DL ^[17]	63.16±2.72
Rie_SR+ Random_DL ^[16]	47.45±7.17
LE_SR+LE_DL ^[21]	74.68±2.71
Ker_SR+Random_DL ^[27]	72.55±2.77
Proposed	78.12±2.56

3.3 实验分析

上述表中的实验结果通过我们在每个数据集上经过 10 次迭代实验获得, 其中, 有个别稀疏表示方法中的字典原子是从样本中随机抽取得到的, 而不是通过学习得到, 我们称这种字典为 Random_DL. 通过所得 10 次实验结果的均值和标准差作为最终的结果, 从这 3 个表中的数据来看, 在 5 个基准数据集上, 我们提出的方法在分类正确率上都是最好的. 在已有的将稀疏表示应用于 SPD 流行上样本的分类算法中, 方法 LE_SR+LE_DL^[21]和 Ker_SR+Random_DL^[27]也可以获得相对较好的效果. 这两种方法都不是直接在 SPD 流形上进行分类的任务, 而是通过映射函数分别将 SPD 流形嵌入到切空间和核空间中, 然后再进行稀疏表示和样本分类. 文献[37]中的 3 种方法——COV-SVM, Gau-SVM 和 RoG-SVM 之间的识别率的差距并不大, 但由于 Gau-SVM 方法在对样本建模时考虑了平均值信息, 所以 Gau-SVM 的识别率要比 COV-SVM 高出一. 方法 RoG-SVM 在对样本建模时使用了更加鲁棒的描述方法, 所以 RoG-SVM 的识别率要比 Gau-SVM 高出一. 综上, 在这 3 个方法中, 方法 RoG-SVM 可以得到较高的识别率. 在表 2 的实验数据中, 确实是 RoG-SVM 的识别率高于 Gau-SVM, Gau-SVM 的识别率高于 COV-SVM. 在 Brodatz 数据集上, 除了方法 Ker_SR+Random_DL 以外^[27], 我们的方法的优势非常明显, 识别率可以达到 90.22%, 同时具有最小的标准差 1.07. 这说明在该数据集上, 我们的方法不仅具有较好的识别效果, 还具有很强的鲁棒性. 在 Virus 数据集上, 我们的方法具有最高的识别率 62.67%, 但是我们的方法的标准差并不是最小的, 但 7.83 的标准差在所有的方中也是相对较小的, 说明我们的方法在该数据上的鲁棒性仍然可以得到保证. 在 UIUCM 数据集, 所有方法的识别率普遍偏低, 但是我们的方法获得了最好的识别率, 且标准差较小.

在 ETH-80 数据集上,我们的方法标准差为 3.75,仅次于 PML^[38].但是我们方法的识别率为 96.00%,远高于 PML 的 90.00%.其中,方法 LE_SR+LE_DL^[21]和 Ker_SR+Random_DL^[27]的识别率和标准差都仅次于我们的方法.在 YTC 数据集上,除了方法 Ker_SR+Random_DL^[27]和 Ker_SR+Random_DL^[27],我们的方法的优势非常明显,识别率可以达到 78.12%,同时具有最小的标准差 2.56.综上,我们的方法在 5 个数据集上具有最好的识别率以及相对较好的鲁棒性.

4 总结与展望

用 SPD 矩阵描述图像或者图像集具有较强的鲁棒性,且由 SPD 矩阵所张成的 SPD 流形是一个非线性的黎曼空间.本文充分考虑 SPD 流形的黎曼度量和几何性质,提出了基于 SPD 流形潜在稀疏表示分类算法.首先,将 SPD 流形嵌入到核空间,且在核空间字典学习过程中学习一个潜在矩阵,灵活联系核空间中字典原子和类别之间的信息,提出了核空间潜在稀疏表示分类模型.但核空间的样本没有明确的形式,给核空间的字典的更新带来了不便,我们使用 Nyström 方法获得核空间训练样本的向量表示去学习更新核空间的潜在字典和潜在矩阵.至此,我们将欧氏空间分类算法与黎曼流形结合起来,实现 SPD 流形上样本的分类任务.接下来,如何在 SPD 流形上更新字典,将是我们下一步的研究重点.

References:

- [1] Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y. Robust face recognition via sparse representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(2):210–227. [doi: 10.1109/TPAMI.2008.79]
- [2] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: Which helps face recognition? In: *Proc. of the 2011 IEEE Int'l Conf. on Computer Vision (ICCV)*. IEEE, 2011. 471–478. [doi: 10.1109/ICCV.2011.6126277]
- [3] Yang M, Zhang L. Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary. In: *Proc. of the European Conf. on Computer Vision*. Berlin, Heidelberg: Springer-Verlag, 2010. 448–461. [doi: 10.1007/978-3-642-15567-3_33]
- [4] Chen Z, Wu XJ, Yin HF, Kittler J. Robust low-rank recovery with a distance-measure structure for face recognition. In: *Proc. of the Pacific Rim Int'l Conf. on Artificial Intelligence*. Cham: Springer-Verlag, 2018. 464–472. [doi: 10.1007/978-3-319-97310-4_53]
- [5] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. on Signal Processing*, 2006,54(11):4311–4322. [doi: 10.1109/TSP.2006.881199]
- [6] Harandi MT, Salzmann M, Hartley R. From manifold to manifold: Geometry-aware dimensionality reduction for SPD matrices. In: *Proc. of the European Conf. on Computer Vision*. Cham: Springer-Verlag, 2014. 17–32. [doi: 10.1007/978-3-319-10605-2_2]
- [7] Harandi M, Salzmann M, Hartley R. Dimensionality reduction on SPD manifolds: The emergence of geometry-aware methods. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2018,40(1):48–62. [doi: 10.1109/TPAMI.2017.2655048]
- [8] Wang R, Guo H, Davis LS, Dai Q. Covariance discriminative learning: A natural and efficient approach to image set classification. In: *Proc. of the 2012 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2012. 2496–2503. [doi: 10.1109/CVPR.2012.6247965]
- [9] Feng F, Wu XJ, Xu T. Object tracking with kernel correlation filters based on mean shift. In: *Proc. of the 2017 Int'l Smart Cities Conf. (ISC2)*. IEEE, 2017. 1–7. [doi: 10.1109/ISC2.2017.8090863]
- [10] Ren J, Wu X. Sparse coding for symmetric positive definite matrices with application to image set classification. In: *Proc. of the Int'l Conf. on Intelligent Science and Big Data Engineering*. Cham: Springer-Verlag, 2015. 637–646. [doi: 10.1007/978-3-319-23989-7_64]
- [11] Sivalingam R, Boley D, Morellas V, Papanikolopoulos N. Tensor sparse coding for region covariances. In: *Proc. of the European Conf. on Computer Vision*. Berlin, Heidelberg: Springer-Verlag, 2010. 722–735. [doi: 10.1007/978-3-642-15561-1_52]
- [12] Sivalingam R, Boley D, Morellas V, Papanikolopoulos N. Tensor sparse coding for positive definite matrices. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2014,36(3):592–605. [doi: 10.1109/TPAMI.2013.143]
- [13] Sivalingam R, Boley D, Morellas V, Papanikolopoulos N. Positive definite dictionary learning for region covariances. In: *Proc. of the 2011 IEEE Int'l Conf. on Computer Vision (ICCV)*. IEEE, 2011. 1013–1019. [doi: 10.1109/ICCV.2011.6126346]

- [14] Sivalingam R, Boley D, Morellas V, Papanikolopoulos N. Tensor dictionary learning for positive definite matrices. *IEEE Trans. on Image Processing*, 2015,24(11):4592–4601. [doi: 10.1109/TIP.2015.2440766]
- [15] Sra S, Cherian A. Generalized dictionary learning for symmetric positive definite matrices with application to nearest neighbor retrieval. In: *Proc. of the Machine Learning and Knowledge Discovery in Databases*. 2011. 318–332. [doi: 10.1007/978-3-642-23808-6_21]
- [16] Cherian A, Sra S. Riemannian sparse coding for positive definite matrices. In: *Proc. of the European Conf. on Computer Vision*. Cham: Springer-Verlag, 2014. 299–314. [doi: 10.1007/978-3-319-10578-9_20]
- [17] Cherian A, Sra S. Riemannian dictionary learning and sparse coding for positive definite matrices. *IEEE Trans. on Neural Networks and Learning Systems*, 2017,28(12):2859–2871. [doi: 10.1109/TNNLS.2016.2601307]
- [18] Pennec X, Fillard P, Ayache N. A Riemannian framework for tensor computing. *Int'l Journal of Computer Vision*, 2006,66(1): 41–66. [doi: 10.1007/s11263-005-3222-z]
- [19] Quang MH, San Biagio M, Murino V. Log-Hilbert-Schmidt metric between positive definite operators on Hilbert spaces. In: *Proc. of the Advances in Neural Information Processing Systems*. 2014. 388–396.
- [20] Faraki M, Harandi MT, Porikli F. Image set classification by symmetric positive semi-definite matrices. In: *Proc. of the 2016 IEEE Winter Conf. on Applications of Computer Vision (WACV)*. IEEE, 2016. 1–8. [doi: 10.1109/WACV.2016.7477621]
- [21] Guo K, Ishwar P, Konrad J. Action recognition using sparse representation on covariance manifolds of optical flow. In: *Proc. of the 2010 7th IEEE Int'l Conf. on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2010. 188–195. [doi: 10.1109/AVSS.2010.71]
- [22] Arsigny V, Fillard P, Pennec X, *et al.* Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 2006,56(2):411–421. [doi: 10.1002/mrm.20965]
- [23] Harandi MT, Sanderson C, Hartley R, *et al.* Sparse coding and dictionary learning for symmetric positive definite matrices: A kernel approach. In: *Proc. of the Computer Vision (ECCV 2012)*. Berlin, Heidelberg: Springer-Verlag, 2012. 216–229. [doi: 10.1007/978-3-642-33709-3_16]
- [24] Sra S. Positive definite matrices and the S-divergence. *Proc. of the American Mathematical Society*, 2016,144(7):2787–2797. [doi: 10.1090/proc/12953]
- [25] Cichocki A, Cruces S, Amari S. Log-determinant divergences revisited: Alpha-beta and gamma log-det divergences. *Entropy*, 2015, 17(5):2988–3034. [doi: 10.3390/e17052988]
- [26] Harandi MT, Hartley R, Lovell B, Sanderson C. Sparse coding on symmetric positive definite manifolds using Bregman divergences. *IEEE Trans. on Neural Networks and Learning Systems*, 2016,27(6):1294–1306. [doi: 10.1109/TNNLS.2014.2387383]
- [27] Li P, Wang Q, Zuo W, Zhang L. Log-Euclidean kernels for sparse representation and dictionary learning. In: *Proc. of the IEEE Int'l Conf. on Computer Vision*. 2013. 1601–1608. [doi: 10.1109/ICCV.2013.202]
- [28] Yang M, Dai D, Shen L, *et al.* Latent dictionary learning for sparse representation based classification. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2014. 4138–4145. [doi: 10.1109/CVPR.2014.527]
- [29] Ren JY, Wu XJ. Vectorial approximations of infinite-dimensional covariance descriptors for image classification. *Computational Visual Media*, 2017,3(4):379–385. [doi: 10.1007/s41095-017-0094-4]
- [30] Faraki M, Harandi MT, Porikli F. Approximate infinite-dimensional region covariance descriptors for image classification. In: *Proc. of the 2015 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015. 1364–1368. [doi: 10.1109/ICASSP.2015.7178193]
- [31] Harandi M, Salzmann M. Riemannian coding and dictionary learning: Kernels to the rescue. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2015. 3926–3935. [doi: 10.1109/CVPR.2015.7299018]
- [32] Yang T, Li Y F, Mahdavi M, Jin R, Zhou ZH. Nyström method vs random Fourier features: A theoretical and empirical comparison. In: *Proc. of the Advances in Neural Information Processing Systems*. 2012. 476–484.
- [33] Kumar S, Mohri M, Talwalkar A. Sampling methods for the Nyström method. *The Journal of Machine Learning Research*, 2012, 13(1):981–1006.

- [34] Elad M. Sparse and redundant representation modeling—What next? *IEEE Signal Processing Letters*, 2012,19(12):922–928. [doi: 10.1109/LSP.2012.2224655]
- [35] Grant M, Boyd S, Ye Y. CVX: Matlab software for disciplined convex programming. 2009. <http://cvxr.com/cvx/>
- [36] Rosasco L, Verri A, Santoro M, *et al.* Iterative projection methods for structured sparsity regularization. Technical Report, MIT-CSAIL-TR-2009-050CBCL-282, 2009. <http://hdl.handle.net/1721.1/49428>
- [37] Mairal J, Bach F, Ponce J, *et al.* Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 2010,11(1):19–60. [doi: 10.1145/1756006.1756008]
- [38] Liao Z, Rock J, Wang Y, Forsyth D. Non-parametric filtering for geometric detail extraction and material representation. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2013. 963–970. [doi: 10.1109/CVPR.2013.129]
- [39] Huang Z, Wang R, Shan S, Li X, Chen X. Log-Euclidean metric learning on symmetric positive definite manifold with application to image set classification. In: Proc. of the Int'l Conf. on Machine Learning. 2015. 720–729.
- [40] Wang Q, Li P, Zuo W, Zhang L. RAID-G: Robust estimation of approximate infinite dimensional Gaussian with application to material recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 4433–4441. [doi: 10.1109/CVPR.2016.480]
- [41] Chang CC, Lin CJ. LIBSVM: A library for support vector machines. *ACM Trans. on Intelligent Systems and Technology (TIST)*, 2011,2(3):1–27. [doi: 10.1145/1961189.1961199]
- [42] Huang Z, Wang R, Shan S, Chen X. Projection metric learning on Grassmann manifold with application to video based face recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2015. 140–149. [doi: 10.1109/CVPR.2015.7298609]
- [43] Hamm J, Lee DD. Grassmann discriminant analysis: A unifying view on subspace-based learning. In: Proc. of the 25th Int'l Conf. on Machine Learning. ACM, 2008. 376–383. [doi: 10.1145/1390156.1390204]



陈凯旋(1993—),男,硕士,主要研究领域为模式识别,计算机视觉,黎曼流形学习.



吴小俊(1967—),男,博士,教授,博士生导师,CCF专业会员,主要研究领域为模式识别,计算机视觉,模糊系统,神经网络,智能系统.