

## 黎曼核局部线性编码<sup>\*</sup>

姜 伟<sup>1</sup>, 毕婷婷<sup>1</sup>, 李克秋<sup>2</sup>, 杨炳儒<sup>3</sup>

<sup>1</sup>(辽宁师范大学 数学学院, 辽宁 大连 116029)

<sup>2</sup>(大连理工大学 计算机科学与技术学院, 辽宁 大连 116024)

<sup>3</sup>(北京科技大学 计算机与通信工程学院, 北京 100083)

通讯作者: 姜伟, E-mail: swxxjw@aliyun.com, http://www.lnnu.edu.cn

**摘 要:** 最近的研究表明:在许多计算机视觉任务中,将对称正定矩阵表示为黎曼流形上的点能够获得更好的识别性能.然而,已有大多数算法仅由切空间局部逼近黎曼流形,不能有效地刻画样本分布.受核方法的启发,提出了一种新的黎曼核局部线性编码方法,并成功地应用于视觉分类问题.首先,借助于最近所提出的黎曼核,把对称正定矩阵映射到再生核希尔伯特空间中,通过局部线性编码理论建立稀疏编码和黎曼字典学习数学模型;其次,结合凸优化方法,给出了黎曼核局部线性编码的字典学习算法;最后,构造一个迭代更新算法优化目标函数,并且利用最近邻分类器完成测试样本的鉴别.在3个视觉分类数据集上的实验结果表明,该算法在分类精度上获得了相当大的提升.

**关键词:** 黎曼流形;对称正定矩阵;切空间;局部约束线性编码;稀疏表示

**中图法分类号:** TP391

中文引用格式: 姜伟, 毕婷婷, 李克秋, 杨炳儒. 黎曼核局部线性编码. 软件学报, 2015, 26(7): 1812–1823. <http://www.jos.org.cn/1000-9825/4714.htm>

英文引用格式: Jiang W, Bi TT, Li KQ, Yang BR. Local linear coding based on Riemannian kernel. Ruan Jian Xue Bao/Journal of Software, 2015, 26(7): 1812–1823 (in Chinese). <http://www.jos.org.cn/1000-9825/4714.htm>

## Local Linear Coding Based on Riemannian Kernel

JIANG Wei<sup>1</sup>, BI Ting-Ting<sup>1</sup>, LI Ke-Qiu<sup>2</sup>, YANG Bing-Ru<sup>3</sup>

<sup>1</sup>(School of Mathematics, Liaoning Normal University, Dalian 116029, China)

<sup>2</sup>(School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China)

<sup>3</sup>(School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract:** Recent research has shown that better recognition performance can be attained through representing symmetric positive definite matrices as points on Riemannian manifolds for many computer vision tasks. However, most existing algorithms only approximate the Riemannian manifold locally by its tangent space and are incapable of scaling effectively distribution of samples. Inspired by kernel methods, a novel method, called local linear coding based on Riemannian kernel (LLCRK), is proposed and applied successfully to vision classification issues. Firstly, with the aid of recently introduced Riemannian kernel, symmetric positive definite matrices are mapped into the reproducing kernel Hilbert space by kernel method and a mathematical model of sparse coding and Riemannian dictionary learning is constructed by local linear coding theory. Secondly, an efficient algorithm of LLCRK is presented for dictionary learning according to the convex optimization methods. Finally, an iterative updating algorithm is constructed to optimize the objective function, and the test samples are classified by nearest neighbor classifier. Experimental results on three visual classification data sets demonstrate that the proposed algorithm achieves considerable improvement in discrimination accuracy.

**Key words:** Riemannian manifold; symmetric positive definite matrix; tangent space; locality-constrained linear coding; sparse representation

\* 基金项目: 国家自然科学基金(61175048); 辽宁省教育厅科学研究项目(L2013408)

收稿时间: 2014-05-06; 修改时间: 2014-06-28; 定稿时间: 2014-09-01

稀疏表示(sparse representation,简称 SR)源于对经典信号表示的扩展,已广泛应用于机器学习、模式识别以及信号处理等许多领域.SR 的核心思想是:信号在适当选取一组过完备基(overcomplete bases)或称字典(dictionary)下,使用尽可能少的基向量来简洁地表达输入信号,以此来揭示信号的内在结构和本质属性.2006 年,Huang 等人<sup>[1]</sup>首次将稀疏表示方法应用于分类研究,开启了稀疏表示理论与应用研究的热潮;Cai 等人<sup>[2]</sup>通过对数据投影方向施加稀疏性限制,构造了统一的稀疏子空间学习模型;Raina 等人<sup>[3]</sup>利用稀疏表示设计了自教(self-taught)学习方法,基于此,从大量无标签数据中获取高层语义特征;Mairal 等人<sup>[4,5]</sup>提出许多应用于纹理分类、手写数字识别等任务的稀疏表示算法;Wright 等人<sup>[6]</sup>基于压缩感知<sup>[7]</sup>理论与稀疏表示提出的 SRC(sparse representation-based classification)分类器在某些标准人脸数据集上获得了非常好的效果.稀疏编码研究表明:非零系数的基更靠近编码数据,基于此,Yu 等人<sup>[8]</sup>在研究高维非线性 Lipschitz 平滑函数进行学习时,基于流形假设,提出了一种在高维非线性流形上的半监督学习方法.该方法是将高维的非线性学习问题转化为简单全局线性学习问题,通过局部坐标编码(local coordinate coding,简称 LCC)学习得到基向量的一组锚点,构成一个局部坐标系,流形上的每个点仅由几个近邻锚点的线性组合来拟合,而线性组合的权重系数就是这个点的局部坐标编码,从而实现稀疏和局部一致性<sup>[9]</sup>.受 LCC 方法启发,局部约束线性编码(locality-constrained linear coding,简称 LLC)<sup>[10]</sup>方法同样是在编码时施以局部性约束,但优化问题的求解效率却很高.LLC 编码可以看作是 LCC 编码的一种快速实现.

然而,SR 和 LLC 是在欧几里德空间框架下,应用欧氏测度来执行字典学习.正如研究物理问题的时空不完全是欧氏空间一样,我们所面临的数据未必分布在欧氏空间.认知心理学研究表明:就感知数据而言,样本空间用弯曲的黎曼流形表达,能够对样本给出很好的解释.在计算机视觉领域,对称正定(symmetric positive definite,简称 SPD)矩阵所诱导的结构非常有用,形成一个非欧、弯曲的黎曼流形,提供了紧凑的目标模型表示方法,融合了图像多种特征,对目标大小、形状和光照变化等具有很强的鲁棒性,成功地应用在目标跟踪<sup>[11]</sup>、人脸识别<sup>[12]</sup>、纹理分类<sup>[13]</sup>、行人检测<sup>[14]</sup>和动作识别<sup>[15]</sup>等领域.

作为一种新型、有效的特征表示模式,可以用矩阵在欧氏空间计算框架下处理 SPD 矩阵,最简单的方法是把  $n \times n$  维 SPD 矩阵看作是一个  $\mathbb{R}^{n(n+1)/2}$  向量,这样,应用欧氏空间相似测度来评价 SPD 矩阵之间的相似度.向量化对称正定矩阵,欧氏距离忽略特征空间的结构信息,降低了特征的有效性.为此,Pennec 等人<sup>[11]</sup>给 SPD 矩阵空间赋予一种仿射不变黎曼度量(affine invariant Riemannian metric,简称 AIRM),使 SPD 矩阵空间晋升为黎曼流形,但利用 AIRM 来度量两个对称正定矩阵之间距离是很耗时的.Tuzel 等人<sup>[15]</sup>提出了采用图像的区域协方差矩阵(即,对称正定矩阵)作为描述子,在所有正例样本的均值处构建切空间,并把所有特征映射到切空间上,训练多个弱分类器,利用 LogitBoosting 算法将多个弱分类器整合为一个强分类器,实现两类分类问题.切空间的分类问题是基于流形在局部上与欧氏空间存在着同胚(diffeomorphism)映射,这导致指数与对数映射只在一个局部邻域内是一一映射,因此,流形上的点没有全局坐标.针对以上问题,Arsigny 等人<sup>[16]</sup>通过在对称正定矩阵空间上赋予 Log-Euclidean 测度代替 AIRM,形成一种李群结构.这样,Log-Euclidean 框架定义了一个从黎曼流形到向量空间同构、同胚与等距映射.在此基础上,Tosato 等人<sup>[17]</sup>进一步将两类问题推广到多类.最近,Carreira 等人<sup>[18]</sup>基于 Log-Euclidean 测度,利用支持向量机对多类问题进行分类,较许多 state-of-the-art 方法更有效.

基于黎曼流形理论,可先通过对数映射将 SPD 矩阵映射到切空间进行相应的计算,再通过指数算子映射回原空间获得最终分析结果.该思想可以非常方便地将欧氏空间的学习方法推广到黎曼空间,但此类方法存在两个局限性:第一,需要频繁使用对数映射与指数映射,使算法的效率低下;第二,切空间上的欧氏距离只是流形上测地距离的近似.为了克服这种局限性,借助于黎曼核,将黎曼流形上的点映射到一个更高维甚至是无限维的再生核 Hilbert 空间.进而,将欧氏空间的数据分析方法推广到黎曼流形<sup>[19-23]</sup>.在许多情况下,这类方法比切空间方法具有更优的性能.受此启发,Li 等人<sup>[24]</sup>提出了基于黎曼核的稀疏表示与字典学习方法,在学习过程,考虑数据的几何结构,并在黎曼空间更新字典矩阵.然而,局部性具有比稀疏性更重要、更本质的特性,因为局部性必然导致稀疏性,反之则不成立.基于以上分析,本文结合局部坐标编码与黎曼核方法的思想,提出了黎曼核局部线性编码(local linear coding based on Riemannian kernel,简称 LLCRK)方法,以对称正定矩阵为输入单位,通过黎曼

核将其映射到一个更高维甚至是无限维的特征空间,由此,黎曼流形上的局部线性编码问题转化为特征空间的局部线性编码问题.通过局部线性编码理论,构建字典与编码学习的数学模型.结合最优化理论,给出了高效的优化算法.

### 1.1 黎曼流形

设  $M$  是一个  $d$  维黎曼流形,则对于任意一点  $X \in M$ ,都有  $X$  的一个开邻域与欧氏空间  $\mathbb{R}^d$  的一个开子集同胚.点  $X$  所有切向量全体张成的线性空间称为  $M$  在点  $X$  处的切空间,记为  $T_X M$ ;切空间矢量  $v \in T_X M$  的范数记为  $\|v\|_X^2 = \langle v, v \rangle_X$ ;  $X$  与  $Y$  两点之间的距离是连接  $X$  与  $Y$  的一条光滑曲线,测地线为所有在流形上连接  $X$  和  $Y$  的最短光滑曲线;对于完备流形来说,任意两点之间只存在一条测地线.切空间  $T_X M$  构成向量空间,而黎曼流形  $M$  是非向量空间.流形、测地线、切空间和切矢量的关系如图 1 所示.通过黎曼指数映射  $E_X: T_X M \rightarrow M$  映射切向量  $v$  到黎曼流形  $M$  上从点  $X$  出发等长同向测地线,  $E_X$  逆映射  $L_X: M \rightarrow T_X M$  将黎曼流形  $M$  上点  $X$  到点  $Y$  测地线映射为切空间  $T_X M$  中等长同向矢量  $v$ .对于  $X, Y \in M$ ,从  $X$  到  $Y$  测地线的切向量定义为  $v = \overline{XY} = L_X(Y)$ , 并且指数映射将  $v$  映射到流形上的点  $Y = E_X(v)$ .

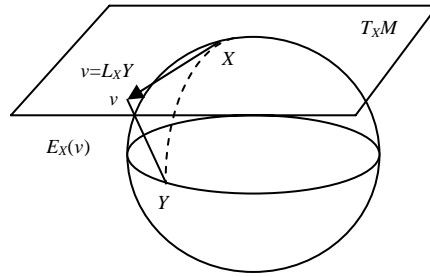


Fig.1 Geodesic and tangent vector

图 1 测地线与切向量

设  $S^+(n) = \{X | X \in S(n), u^T X u > 0, \forall u \in \mathbb{R}^n\}$  表示  $n \times n$  阶对称正定矩阵集合,则  $S^+(n)$  是一个可微的黎曼流形  $M$ .对于任意矩阵  $P \in S^+(n)$ ,  $P$  可被分解为  $P = U \Lambda U^T$ , 其中,  $U$  为  $P$  的特征向量构成的矩阵,  $\Lambda$  是对角矩阵,则相应指数与对数运算为

$$\exp(P) = \sum_{k=0}^{\infty} \frac{P^k}{k!} = U \exp(\Lambda) U^T,$$

$$\log(P) = \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{k} (P - I)^k = U \log(\Lambda) U^T.$$

对于黎曼流形  $M$  上的任意一点  $X$ ,均可做一切空间  $T_X M$  与流形  $M$  的微分同胚.对于切空间中每一个切向量  $S_i$ ,可通过指数映射将  $S_i$  映射为流形  $M$  上从点  $X$  出发的等长同向的测地线,则指数映射定义为

$$E_X(S_i) = P_i = X^{1/2} \exp(X^{-1/2} S_i X^{-1/2}) X^{1/2},$$

其中,  $\exp(\cdot)$  表示矩阵指数算子.

而指数映射逆映射,即对数映射定义为

$$L_X(P_i) = S_i = X^{1/2} \log(X^{-1/2} P_i X^{-1/2}) X^{1/2},$$

其中,  $\log(\cdot)$  表示矩阵对数算子.

### 1.2 稀疏表示与局部线性编码

#### 1.2.1 稀疏表示

稀疏表示<sup>[25]</sup>是在一定的误差范围内,从一组过完备基  $\{d_i \in \mathbb{R}^m\}_{i=1}^n$  组成的集合中选取少量基信号,线性表示输入信号  $x \in \mathbb{R}^m$ .通常,这些基被称之为原子,它们的集合称为字典,稀疏表示数学模型定义如下:

$$\begin{cases} \min_{\alpha} \|\alpha\|_0 \\ \text{s.t. } x = D\alpha \end{cases} \quad (1)$$

其中,  $D=[d_1, d_2, \dots, d_n] \in \mathbb{R}^{m \times n}$  为基矩阵,  $\alpha \in \mathbb{R}^n$  为系数向量, 向量  $\alpha$  的  $\ell_0$  范数表示向量中非零元素的数目. 但是, 公式(1)已被证明是一个 NP-hard 问题, 计算非常复杂. 知名学者 Donoho<sup>[7]</sup>, Candes 和 Tao<sup>[26]</sup> 已证明: 在足够稀疏的条件下,  $\ell_1$  范数等价  $\ell_0$  范数, 即:

$$\begin{cases} \min_{\alpha} \|\alpha\|_1 \\ \text{s.t. } x = D\alpha \end{cases} \quad (2)$$

由于  $\ell_1$  范数是凸的, 所以公式(2)是凸优化问题. 考虑到信号通常被噪声污染, 在公式(2)中引入噪声项, 则稀疏表示模型为

$$\min_x \|\alpha\|_1, \quad \text{s.t. } \|x - D\alpha\|_2 \leq \varepsilon \quad (3)$$

另外, 稀疏表示模型还可表示为如下正则化表达方式:

$$\min_x \frac{1}{2} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (4)$$

其中,  $\lambda$  为正则化参数, 用于均衡稀疏性与稀疏表达误差.

### 1.2.2 局部线性编码

局部线性编码是将数据信号  $x \in \mathbb{R}^m$  分解为锚点的线性组合. 也就是说, 每个数据信号是由数据点近邻的少数几个基信号或者字典原子的线性组合, 局部线性编码严格的数学模型如下所示:

$$\{\hat{D}, \{\hat{\alpha}_i\}_1^N\} = \min_{D \in \rho, \alpha} \sum_i \left( \frac{1}{2} \|x_i - D\alpha_i\|_2^2 + \mu \sum_j |\alpha_i^j| \|d_j - x_i\|_2^2 \right) \quad (5)$$

其中,  $D=[d_1, \dots, d_n] \in \mathbb{R}^{m \times n}$  为基矩阵,  $\rho = \{D \mid \|d_i\| \leq 1, i=1, \dots, n\}$  是  $D$  的凸可行集,  $\alpha_i \in \mathbb{R}^n$  是重构系数. 目标函数的第 1 项为重构误差; 第 2 项表示局部性, 物理意义是数据点  $x$  离字典原子  $d_j$  越近, 其重构系数越大. 参数  $\mu$  用于折中目标函数中两项的重要程度,  $\|\cdot\|_2$  表示  $\ell_2$  范数. 由于目标函数中有两个变量, 目标函数对于它们不能同时是凸的, 但如果一个固定, 则对另一个量, 其目标函数是凸的. 因此, 我们首先固定字典  $D$ , 求解稀疏编码  $\alpha$ , 那么目标函数(5)变换为如下形式:

$$\hat{\alpha} = \min_{\alpha} \frac{1}{2} \|x - D\alpha\|_2^2 + \mu \sum_j |\alpha^j| \|d_j - x\|_2^2 \quad (6)$$

令  $A_{ji} = \|d_j - x\|_2^2$ ,  $\beta = A\alpha$ , 并且假设  $d_j \neq x$ , 则目标函数(6)改写为如下形式:

$$\hat{\beta} = \min_{\beta} \frac{1}{2} \|x - DA^{-1}\beta\|_2^2 + \mu \|\beta\|_1 \quad (7)$$

其中,  $\|\beta\|_1 = \sum_j |\beta^j|$  表示  $\beta$  的  $\ell_1$  范数.

令  $\Gamma_i = DA^{-1}$ , 公式(7)被重写为

$$\hat{\beta} = \min_{\beta} \frac{1}{2} \|x - \Gamma\beta\|_2^2 + \mu \|\beta\|_1 \quad (8)$$

这等价于标准的稀疏编码, 我们可以利用 feature-sign-search<sup>[27]</sup> 算法优化目标函数求得稀疏编码. 获得数据信号的稀疏编码以后, 再把字典  $D$  看作未知变量, 对其进行优化, 重写目标函数(5)如下:

$$\hat{D} = \min_{D \in \rho} \frac{1}{2} \|x - D\alpha\|_2^2 + \mu \sum_j |\alpha^j| \|d_j - x\|_2^2 \quad (9)$$

显然, 公式(9)是一个二次规划问题, 我们可以利用块坐标下降法和随机梯度下降法求解.

## 2 黎曼核局部线性编码

$S^+(n)$  不是欧氏空间, 而是一个可微的黎曼流形, 没有点积运算, 因此, 欧式空间的核函数不能用于对称正定

矩阵样本集.基于此,最近几年,黎曼核得到广泛的研究,文献[14]给出了基于测地距离的伪核.

$$k_R(X,Y)=\exp\{-\sigma^{-1}d_G(X,Y)\} \quad (10)$$

其中,  $d_G(X,Y)=\text{trace}\left(\log^2\left(X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\right)\right)$ ,  $\text{trace}(\cdot)$ 表示矩阵的迹.最近,Sra 等人<sup>[19]</sup>使用 Bregman 测度引入半正定

Stein 核,即:

$$k(X,Y)=e^{-\sigma S(X,Y)}=2^{d\sigma}\frac{\sqrt{\det(X)^\sigma\det(Y)^\sigma}}{\det(X+Y)^\sigma} \quad (11)$$

其中,  $\sigma \in \left\{\frac{1}{2}, \frac{2}{2}, \dots, \frac{d-1}{2}\right\} \cup \left\{\tau \in \mathbb{R} : \tau > \frac{d-1}{2}\right\}$ ,  $d$  为  $X$  的行数,或者列数.Li 等人<sup>[24]</sup>研究了 Log-Euclidean 黎曼度量下协方差矩阵的李群结构,给出了黎曼正定核函数表达:

$$k_R(X,Y)=\exp(-\gamma\|\log(X)-\log(Y)\|^2) \quad (12)$$

第 1 种把欧式空间的高斯核函数引入黎曼流形的方式是,仅将欧式空间距离函数替换仿射不变黎曼度量 (affine invariant Riemannian metric,简称 AIRM)<sup>[28]</sup>,这样并不能确保替换后一定是正定核;第 2 种核函数使用 Stein 度量近似测地距离,只有在核宽度  $\sigma$  取某些值的情况下,才能保证正定性;第 3 种克服了前两种的缺点,所提出的核函数能够有效地刻画测地距离,这样可以精确地测量重建误差,并且也满足 Mercer 条件.因此,本文采用第 3 种核进行数据分析.

## 2.1 目标函数

给定  $d$  维黎曼流形  $M$  中的一个样本集  $X=[X_1, X_2, \dots, X_N]$ ,  $X$  是对称正定矩阵集合,  $N$  为样本总数.黎曼流形框架上局部线性编码的基本思想是:首先,通过非线性映射  $\phi$  将分布在黎曼流形  $M$  上的样本映射到再生核 Hilbert 空间  $F$  中,即,  $\phi(X)=[\phi(X_1), \dots, \phi(X_N)]$ ;然后,在空间  $F$  中实现局部线性编码.这样,黎曼流形框架上局部线性编码准则函数可表示为

$$\{\hat{D}, \{\hat{\alpha}_i\}_1^N\} = \min_{D \in \mathbb{R}^n, \alpha_i \in \mathbb{R}^n} \sum_{i=1}^N \left( \|\phi(X_i) - \phi(D)\alpha_i\|_2^2 + \mu \sum_{k=1}^n |\alpha_i^k| \|\phi(D_k) - \phi(X_i)\|_2^2 \right) \quad (13)$$

其中,  $X_i$  是样本集  $X$  中的第  $i$  个样本;  $\alpha_i = [\alpha_i^1, \dots, \alpha_i^n]^T$  是  $X_i$  的稀疏表示向量;  $D_k$  是字典原子,并且  $D=[D_1, D_2, \dots, D_n]$ ; 参数  $\mu$  用于折中目标函数两项的重要程度.

## 2.2 优化算法

目标函数是  $D$  和  $\alpha_i$  非凸函数,因此它没有闭式解(close-form solution).本文构建了一种迭代算法用以优化这个目标函数,其核心思想是:固定字典矩阵  $D$ , 求出一个最优的稀疏编码向量  $\alpha_i$ ; 然后固定  $\alpha_i$ , 求出一个最优的字典  $D$ , 当某个迭代终止条件满足时,输出字典  $D$  和编码向量  $\alpha_i$ . 对于每一个样本  $X_i$ , 在给定  $D$  的情况下,对稀疏编码  $\alpha_i$  学习,重写目标函数(13)如下:

$$\hat{\alpha}_i = \min_{\alpha_i \in \mathbb{R}^n} \left( \|\phi(X_i) - \phi(D)\alpha_i\|_2^2 + \mu \sum_{j=1}^n |\alpha_i^j| \|\phi(D_j) - \phi(X_i)\|_2^2 \right) \quad (14)$$

令对角矩阵  $A_i$  的对角元素为  $A_i(k,k)=\|\phi(X_i) - \phi(D_k)\|_2^2$ ,  $k=1, \dots, n$ ;  $\beta_i=A_i\alpha_i$ , 并且假设  $D_j \neq X_i$ , 则目标函数(14)改写为如下形式:

$$\hat{\beta}_i = \min_{\beta_i} \frac{1}{2} \|\phi(X_i) - \phi(D)A_i^{-1}\beta_i\|_2^2 + \mu \|\beta_i\|_1 \quad (15)$$

其中,  $\|\beta_i\|_1 = \sum_j |\beta_i^j|$  表示  $\beta_i$  的  $\ell_1$  范数.

展开目标函数(15),除去与  $\beta_i$  无关的项,利用核函数重写目标函数为

$$\hat{\beta}_i = \arg \min_{\beta_i} k(X_i, X_i) + \beta_i^T A_i^{-1} K(D, D) A_i^{-1} \beta_i - 2K(X_i, D) A_i^{-1} \beta_i + \mu \|\beta_i\|_1 \quad (16)$$

其中,

- $K(X_i, D) = [a_j]_{n \times 1}, a_j = k(X_i, D_j);$
- $K(D, D) = [a_{ij}]_{n \times n}, a_{ij} = k(D_i, D_j);$
- $A_i(k, k) = \|\phi(X_i) - \phi(D_k)\|_2^2 = K(X_i, X_i) + K(D_k, D_k) - 2K(X_i, D_k).$

我们可以利用 feature-sign-search 优化目标函数求得稀疏编码. 获得数据信号的稀疏编码后, 对字典  $D$  进行学习, 重写目标函数(13), 展开并除去与字典无关项, 即:

$$\begin{aligned} J &= \min_{D \in \rho} \sum_{i=1}^N \left( \left\| \phi(X_i) - \sum_{j=1}^n \alpha_i^j \phi(D_j) \right\|_2^2 + \mu \sum_{k=1}^n |\alpha_i^k| \|\phi(D_k) - \phi(X_i)\|_2^2 \right) \\ &= \min_{D \in \rho} \sum_{i=1}^N \left( \frac{1}{2} \sum_{j=1}^n \sum_{q=1}^n \alpha_i^j \alpha_i^q \phi(D_j)^T \phi(D_q) - \sum_{j=1}^n \alpha_i^j \phi(D_j)^T \phi(X_i) + \mu \sum_{k=1}^n |\alpha_i^k| (\phi(D_k)^T \phi(D_k) - 2\phi(D_k)^T \phi(X_i)) \right) \quad (17) \\ &= \min_{D \in \rho} \sum_{i=1}^N \left( \frac{1}{2} \sum_{j=1}^n \sum_{q=1}^n \alpha_i^j \alpha_i^q K(D_j, D_q) - \sum_{j=1}^n \alpha_i^j K(D_j, X_i) + \mu \sum_{k=1}^n |\alpha_i^k| (K(D_k, D_k) - 2K(D_k, X_i)) \right) \end{aligned}$$

采用迭代优化字典每一原子的方法来进行字典学习. 设当前需要优化的字典第  $m$  原子为  $D_m, m \in (1, 2, \dots, n)$ , 对目标函数(17)求导, 可得:

$$\frac{\partial J}{\partial D_m} = -\frac{1}{D_m} \gamma \sum_{i=1}^N \left( \frac{1}{2} \sum_{q=1}^n \alpha_i^m \alpha_i^q K(D_m, D_q) (\log D_m - \log D_q) - K(D_m, X_i) (\log D_m - \log X_i) (\alpha_i^m + |\alpha_i^m|) \right) \quad (18)$$

为了获得优化的  $D_m$ , 令  $\partial J / \partial D_m = 0$ . 这样, 我们很容易得到  $\log D_m$ . 然而, 由于目标函数中存在核函数项  $K(D_m, \cdot)$  和  $\log D_q$ , 求解仍很困难. 因此, 采用迭代算法来获得近似解. 在第  $k$  次迭代中, 我们使用第  $k-1$  次结果计算核函数项  $K(D_m, \cdot)$  和  $\log D_q$ . 采用文献[24]中提供的方法获得字典更新的表达式, 即:

$$D_m = \exp(\log(D_m) + d_{D_m} \log(-\alpha \partial f / \partial D_m)) \quad (19)$$

其中,  $d_{D_m}(U)$  表示在  $D_m$  点的矩阵对数与切矩阵  $U$  位移之差.

### 2.3 算法描述

输入: 来自黎曼流形上的训练集  $X = \{X_i\}_{i=1}^N$ ,  $X_i$  为对称正定矩阵; 测试样本  $Y$ ; 核函数;

输出: 黎曼字典  $D = \{D_i\}_{i=1}^n$ .

步骤 1. 从  $X$  中随机选取  $n$  个样本初始化字典  $D = \{D_i\}$ ;

步骤 2. 根据当前  $D$ , 由公式(16)求解稀疏表示  $\beta_i$ , 再由  $\alpha_i = A_i^{-1} \beta_i$ ;

步骤 3. 固定  $\alpha_i$ , 由公式(17)求得  $D$ ;

步骤 4. 返回步骤 2, 直至目标函数的值在连续两次迭代中足够接近或达到最大迭代次数.

## 3 实验

为了保证实验的充分性与广泛性, 选用了 3 个具有不同特点的数据集, 分别是 Brodatz 纹理数据集、ETHZ 行人数据集和 FERET 人脸数据集, 以验证 LLCRK 算法的有效性. 实验中采用协方差矩阵(即, 对称正定矩阵)表示图像, 以此构成 Riemannian 流形上的数据. 设  $I$  表示灰度图像灰度值或三维彩色图像分量集合,  $F$  是从图像  $I$  中提取的图像特征, 则有:

$$F(\mathbf{x}, \mathbf{y}) = \phi(I, \mathbf{x}, \mathbf{y}),$$

其中,  $\phi$  表示与图像属性有关的映射, 比如灰度、颜色和梯度等.  $\phi(I, \mathbf{x}, \mathbf{y})$  的值域可以张成  $d$  维实空间, 所有像素对应的  $\phi(I, \mathbf{x}, \mathbf{y})$  值均为  $d$  维实空间中的一点. 对于给定目标区域, 区域内所有像素对应的  $\phi$  值可以表示为  $\{z_i\}_{i=1}^n$ ,  $z_i$  为  $d$  维实向量,  $n$  为区域内像素个数, 则目标区域协方差矩阵  $C_R$  定义为

$$C_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - \mu)(z_k - \mu)^T,$$

其中,  $\mu$  为向量集  $\{z_i\}_{i=1}^n$  的均值.

### 3.1 数据集描述

在我们的实验中用到了 3 个数据集,数据集的信息汇总如下.

- Brodatz 纹理数据集

该数据集包括 111 幅 640×640 的纹理图像,也即有 111 个不同纹理类,具有 256 个灰度级.图 2 显示了来自 Brodatz 纹理数据集的部分纹理图像.首先把每张图像不重叠地分割成 16 张 160×160 的子图,即,每类含有 16 个数据,共 1 776 个数据.实验中取  $\phi(I, \mathbf{x}, \mathbf{y}) = \left( \mathbf{x}, \mathbf{y}, I(\mathbf{x}, \mathbf{y}), \left| \frac{\partial I}{\partial \mathbf{x}} \right|, \left| \frac{\partial I}{\partial \mathbf{y}} \right|, \left| \frac{\partial^2 I}{\partial \mathbf{x}^2} \right|, \left| \frac{\partial^2 I}{\partial \mathbf{y}^2} \right| \right)^T$ , 其中,  $I(\mathbf{x}, \mathbf{y})$  坐标为  $(\mathbf{x}, \mathbf{y})$  处灰度值,  $\frac{\partial I}{\partial \mathbf{x}}$  为  $(\mathbf{x}, \mathbf{y})$  处横向一阶梯度,  $\frac{\partial I}{\partial \mathbf{y}}$  为  $(\mathbf{x}, \mathbf{y})$  处纵向一阶梯度,  $\frac{\partial^2 I}{\partial \mathbf{x}^2}$  为  $(\mathbf{x}, \mathbf{y})$  处横向二阶梯度,  $\frac{\partial^2 I}{\partial \mathbf{y}^2}$  为  $(\mathbf{x}, \mathbf{y})$  处纵向二阶梯度.这样,得到  $\mathbf{z}_i$  的维数是 7 维,计算得到  $C_R$  为 7×7 的对称正定矩阵.

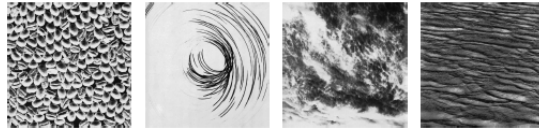


Fig.2 Some texture images from Brodatz texture data set

图 2 来自 Brodatz 纹理数据集的部分纹理图像

- ETHZ 行人数据集

该数据集广泛用于视频跟踪和人体目标再识别算法的实验评价,由移动像机从不同角度拍摄的 3 个图像序列组成,提供了一系列行人外观变化.第 1 个图像序列包括 83 个行人的 4 857 幅图像;第 2 个图像序列包括 35 个行人的 1 936 幅图像;第 3 个图像序列包括 28 个行人的 1 762 幅图像.这些跟踪序列包含部分遮挡、严重遮挡以及剧烈光照变化的图像.图 3 显示了来自 ETHZ 数据集的部分行人图像.为了描述每幅图像,使用如下特征计算方差加以描述:  $\phi(I, \mathbf{x}, \mathbf{y}) = [\mathbf{x}, \mathbf{y}, R_{\mathbf{x}, \mathbf{y}}, G_{\mathbf{x}, \mathbf{y}}, B_{\mathbf{x}, \mathbf{y}}, R'_{\mathbf{x}, \mathbf{y}}, G'_{\mathbf{x}, \mathbf{y}}, B'_{\mathbf{x}, \mathbf{y}}, R''_{\mathbf{x}, \mathbf{y}}, G''_{\mathbf{x}, \mathbf{y}}, B''_{\mathbf{x}, \mathbf{y}}]$ , 其中,  $\mathbf{x}, \mathbf{y}$  表示像素的位置.  $R_{\mathbf{x}, \mathbf{y}}$ ,  $G_{\mathbf{x}, \mathbf{y}}$  和  $B_{\mathbf{x}, \mathbf{y}}$  分别表示相应的颜色信息,而且,  $C'_{\mathbf{x}, \mathbf{y}} = \left[ \left| \frac{\partial C}{\partial \mathbf{x}} \right|, \left| \frac{\partial C}{\partial \mathbf{y}} \right| \right]$  和  $C''_{\mathbf{x}, \mathbf{y}} = \left[ \left| \frac{\partial^2 C}{\partial \mathbf{x}^2} \right|, \left| \frac{\partial^2 C}{\partial \mathbf{y}^2} \right| \right]$  分别表示颜色  $C$  的梯度和拉普拉斯算子.这样得到  $\mathbf{z}_i$  的维数是 11 维,计算得到  $C_R$  为 11×11 的对称正定矩阵.



Fig.3 Some pedestrian images from ETHZ data set

图 3 来自 ETHZ 数据集的部分行人图像

- FERET 人脸数据集

该数据集由 1 199 人的 14 126 幅人脸图像组成,包括不同表情、光照条件、姿态以及不同时间拍摄的图像.本文选取其中带有‘b’的一个子集进行算法性能评价,该子集由 198 个人组成,每个人 7 幅图片.在实验中,所有人脸图像都经过了标准化处理,即:首先标定出人脸图像中两眼的位置,并在该位置将图像对齐;然后,剪切出图像中的面部区域,并将其缩放成统一的 64×64 大小.图 4 显示了来自 FERET 数据集的部分人脸图像,这些图像包含了 ba, bd, be, bf, bg, bj, bk 这 7 种标识符.实验中,取  $\mathbf{x}, \mathbf{y}$  坐标,  $(\mathbf{x}, \mathbf{y})$  处灰度值,选择 5 个尺度、8 个方向 Gabor 变换提取图像的局部特征,得到  $\mathbf{z}_i$  的维数是 43 维.这样,计算得到  $C_R$  为 43×43 的对称正定矩阵.



Fig.4 Some face images of one person from FERET data set

图 4 来自 FERET 数据集中某个人的部分人脸图像

### 3.2 实验环境设置

针对 LLCRK 算法与其他相关算法的关系,对比算法包括:

- 稀疏表示分类(sparse representation-based classification,简称 SRC)<sup>[6]</sup>;
- 局部限制线性编码(locality-constrained linear coding,简称 LLC)<sup>[10]</sup>;
- 张量稀疏编码(tensor sparse coding,简称 TSC)<sup>[29]</sup>;
- Log-Euclidean 稀疏表示(log-Euclidean sparse representation,简称 logE-SR)<sup>[30]</sup>;
- 基于 Stein 核的黎曼稀疏表示(Riemannian sparse representation,简称 RSR)<sup>[19]</sup>;
- Log-Euclidean 核稀疏表示(log-Euclidean sparse representation,简称 logE-KSR)<sup>[24]</sup>.

其中, SRC 是经典稀疏表示学习算法, LLC 是由局部限制所诱导出的稀疏表示分类学习算法,这两种算法处理的数据用向量表示,是基于非黎曼框架的方法; TSC 扩展了基于向量形式的稀疏编码,编码的对象是对称正定矩阵,衡量两个对称正定矩阵的逼近程度是 LogDet 散度; logE-SR 利用黎曼测度度量对称正定矩阵之间的距离,并在对数域进行稀疏分解; RSR 利用半正定 Stein 核在高维希尔伯特空间度量黎曼流形两点之间的相似性,并在该空间求解稀疏表示; logE-KSR 利用正定高斯核度量黎曼流形点与点之间的相似性测度,基于此进行稀疏分解.

实验中涉及到的算法均需设置参数,为了公平起见,对每一种算法,我们都尝试了多种参数的设置,最后选择最佳的参数作为最终实验结果.对于 RSR 算法,Stein 核参数  $\sigma$  的搜索范围为  $\left\{\frac{1}{2}, \frac{2}{2}, \dots, \frac{m-1}{2}\right\}$ ,  $m$  为对称正定矩阵的行数或者列数.对于我们所提出的算法,核函数选择公式(12)给出的黎曼核,其中,  $\gamma$  遍历集合  $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\}$  的所有值.

对于 Brodatz 纹理数据集,随机地选取每类中的  $t(t=8, 10, 12)$  个图像作为训练数据,其余作为测试数据进行实验.对于 ETHZ 行人集,为了计算上方便起见,将图像统一缩放至  $64 \times 32$  分辨率,对于每个行人,随机地选择  $t=10$  张图像作为训练集,剩下的作为测试集;对于 FERET 人脸数据集,随机地选取每个人的  $t(t=3, 4, 5)$  个图像作为训练数据,剩下的作为测试数据进行实验.对每种数据集的训练/测试分割,重复 10 轮,得到训练数据集和测试数据集的不同划分,平均识别率作为最终性能指标.在实验中,首先基于训练集提取特征,构造一个超完备字典,分别利用以上算法获得数据紧致的稀疏表示,进而采用简单的最近邻分类器对得到的稀疏特征进行分类.

### 3.3 识别率

基于上述数据集和环境设置,表 1~表 3 列出了不同算法的最优识别率及其对应的标准差和字典原子个数.图 5~图 7 显示了不同算法的分类性能对字典原子个数变化曲线,横坐标表示字典原子个数,纵坐标为识别率.

Table 1 Performance of different algorithms on Brodatz dataset

表 1 Brodatz 数据集上不同算法的性能

	8训练	10训练	12训练
SRC	72.01%±2.08% (532)	73.45%±2.32% (529)	75.34%±2.83% (581)
LLC	74.01%±2.67% (648)	77.45%±2.45% (684)	80.34%±2.43% (721)
TSC	88.02%±1.08% (638)	89.55%±1.52% (724)	91.34%±1.93% (521)
logE-SR	93.44%±1.35% (723)	93.55%±2.30% (699)	95.88%±1.87% (655)
RSR	94.40%±1.89% (484)	95.92%±0.98% (752)	97.24%±1.21% (659)
logE-KSR	97.16%±1.28% (692)	97.64%±1.11% (713)	97.24%±1.28% (889)
LLCRK	98.39%±0.87% (731)	99.72%±0.05% (621)	99.87%±0.03% (750)



Table 2 Performance of different algorithms on ETHZ dataset

表 2 ETHZ 数据集上不同算法的性能

	第1序列	第2序列	第3序列
SRC	43.23%±3.62% (427)	33.12%±3.97% (469)	37.67%±4.23% (222)
LLC	44.12%±4.13% (423)	35.00%±4.97% (234)	38.45%±3.45% (217)
TSC	80.65%±1.52% (535)	78.00%±1.97% (189)	88.02%±1.23% (189)
logE-SR	87.75%±1.73% (589)	89.05%±1.26% (235)	91.20%±1.64% (231)
RSR	89.82%±1.12% (623)	89.84%±1.46% (296)	92.52%±1.95% (218)
logE-KSR	92.44%±1.51% (553)	92.15%±1.03% (320)	94.44%±1.24% (189)
LLCRK	98.39%±0.84% (375)	94.21%±0.99% (299)	96.56%±0.78% (247)

Table 3 Performance of different algorithms on FERET dataset

表 3 FERET 数据集上不同算法的性能

	3训练	4训练	5训练
SRC	93.34%±1.53% (398)	94.49%±1.12% (358)	97.02%±1.79% (444)
LLC	94.23%±1.67% (423)	95.69%±1.32% (412)	98.13%±1.79% (432)
TSC	81.30%±2.34% (392)	88.49%±2.12% (358)	93.02%±2.09% (495)
logE-SR	92.71%±1.72% (434)	94.93%±1.67% (555)	97.05%±1.78% (576)
RSR	95.75%±1.14% (483)	96.54%±1.26% (589)	98.88%±1.09% (432)
logE-KSR	97.13%±1.01% (479)	98.03%±1.03% (389)	99.54%±1.11% (389)
LLCRK	98.56%±0.69% (584)	99.97%±0.04% (421)	99.98%±0.02% (501)

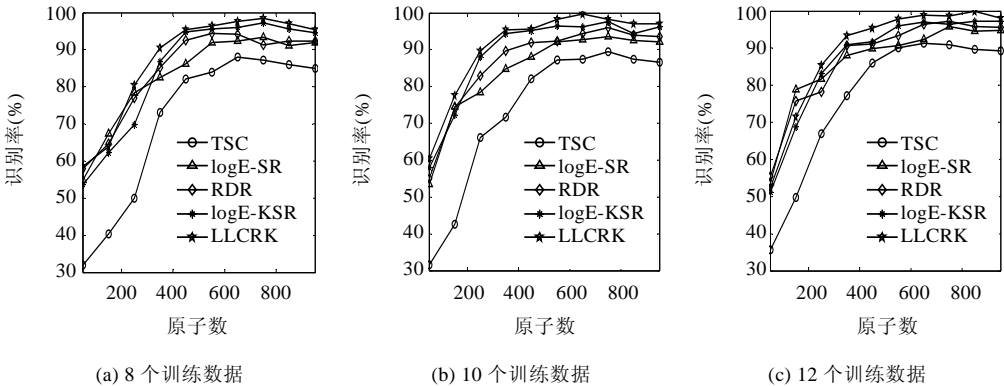


Fig.5 Recognition rate curves of different algorithms versus different numbers of atoms for Braodatz dataset

图 5 不同算法在 Brodatz 数据集中随着字典原子个数变化的识别率曲线图

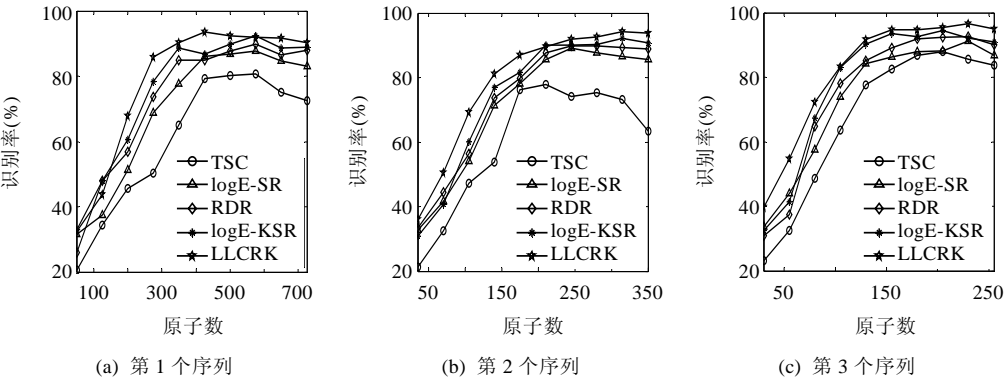


Fig.6 Recognition rate curves of different algorithms versus different numbers of atoms for ETHZ dataset

图 6 不同算法在 ETHZ 数据集中随着字典原子个数变化的识别率曲线图

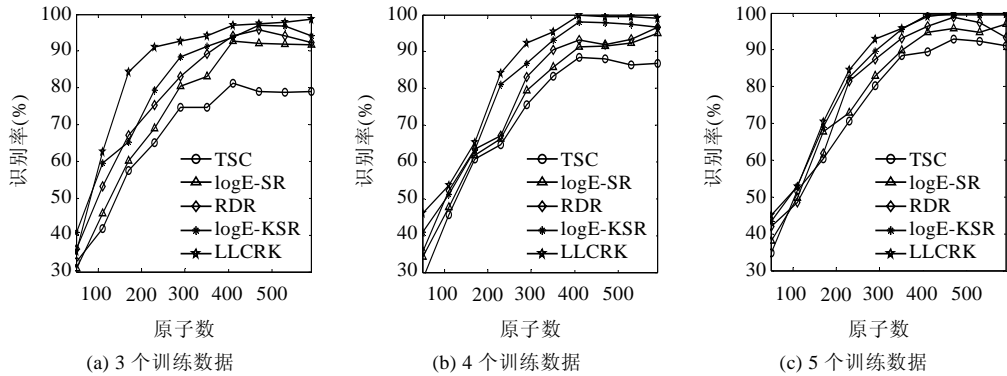
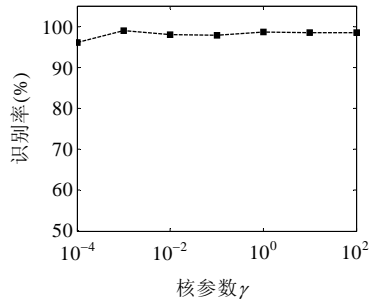
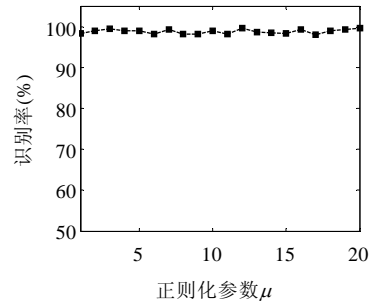


Fig.7 Recognition rate curves of different algorithms versus different numbers of atoms for FERET dataset

图 7 不同算法在 FERET 数据集中随着字典原子数变化的识别率曲线图

### 3.4 参数敏感性

参数选择是一个公开问题,没有可靠的理论作为依据,这部分主要考察正则化参数 $\mu$ 与黎曼核参数 $\gamma$ 对分类精度的影响.具体而言,我们在 FERET 数据集中随机地选取每类中的 3 个数据作为训练集,其余作为测试数据进行实验,通过最近邻分类器计算分类识别率.实验对每种训练\测试分割,重复 10 轮,得到训练数据集和测试数据集的不同划分,平均识别率作为最终性能指标.图 8 显示了固定字典原子个数为 500 和正则化参数 $\mu=10$ ,遍历不同核参数 $\gamma$ 所对应的识别率.图 9 显示了固定字典原子个数为 500 和核参数 $\gamma=1$ ,不同正则化参数 $\mu$ ,LLCRK 在 FERET 数据集上识别率的对比情况.

Fig.8 Classification accuracies using different kernel parameter  $\gamma$ 图 8 不同核参数 $\gamma$ 对应的识别率Fig.9 Classification accuracies using different regularization parameter  $\mu$ 图 9 不同正则化参数 $\mu$ 对应的识别率

### 3.5 实验结果总体讨论

上面我们在 3 个不同图像数据集上对所提算法进行了较为全面的评价,表 1~表 3 列出了各种算法对应的最好结果及其对应的标准差和字典原子数,图 5~图 7 给出了各种算法平均识别率对字典原子数变化曲线,图 8、图 9 分别对参数的敏感性进行了研究性实验.从上述实验结果可以得到如下主要结论:

- (1) 通过实验中在相同数据集下各种算法识别率的综合比较,TSC 算法的效果在矩阵表示方法中最差.这表明,黎曼测度较欧氏测度能够在一定程度上刻画数据的非线性结构,可更有效地保留数据判别信息;
- (2) 只要字典原子个数超过某个阈值,LLCRK 均一致地超过其余的 4 种算法,这说明,具有局部限制结构与正定黎曼核相融合的方法更有助于刻画数据的非线性流形结构,因而在实验中能够获得最优的识别性能;
- (3) 字典原子的个数并不是越多越好,当字典原子达到饱和时,再继续增加字典原子个数,识别率反而下降.这也印证了字典原子增多反而会破坏稀疏性条件,使识别率有所降低的结论;
- (4) 算法缺点:算法中使用了核函数,且核函数由于需要进行对数运算,其算法时间复杂度相对较高,很难推

广到大数据中.考虑使用 Nystrom 等算法来加速核函数的计算,是进一步的研究方向.

#### 4 总 结

本文提出一种新颖的黎曼核局部线性编码分析算法,在黎曼流形框架下,将数据从原始的黎曼流形上映射到一个更高维甚至是无穷维的特征空间中,进而把局部线性编码与黎曼核整合在单目标函数中,提取用对称正定矩阵所描述数据有效稀疏表示.在所有数据集上,LCCRK 算法在实验中表现得最好,主要是因为 LCCRK 同时考虑了数据的局部性和非线性性质,并且全面继承了 logE-KSR 算法的优点,从而得到最优的识别性能.进一步的研究方向:考虑使用 Nystrom 等算法来加速核函数的计算和探索高效的核函数,实现黎曼流形上局部线性编码.

**致谢** 在此,我们向对本文的工作给予支持和建议的同行,尤其是北京交通大学计算机科学技术系于剑教授来辽宁师范大学进行学术交流时给予的指导表示感谢.

#### References:

- [1] Huang K, Aviyente S. Sparse representation for signal classification. In: Proc. of the Advances in Neural Information Processing Systems. 2006. 609–616. <http://papers.nips.cc/paper/3130-sparse-representation-for-signal-classification.pdf>
- [2] Cai D, He XF, Han JW. Spectral regression: A unified approach for sparse subspace learning. In: Proc. of the IEEE Int'l Conf. on Data Mining. 2007. 73–82. [doi: 10.1109/ICDM.2007.89]
- [3] Raina R, Battle A, Lee H, Packer B, Ng AY. Self-Taught learning: Transfer learning from unlabeled data. In: Proc. of the Int'l Conf. on Machine Learning. Oregon: ACM Press, 2007. 759–766. [doi: 10.1145/1273496.1273592]
- [4] Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. Supervised dictionary learning. In: Proc. of the Advances in Neural Information Processing Systems. 2008. <http://papers.nips.cc/paper/3448-supervised-dictionary-learning.pdf>
- [5] Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. Discriminative learned dictionaries for local image analysis. In: Proc. of the 2008 IEEE Conf. on Computer Vision and Pattern Recognition. Anchorage: IEEE Computer Society Press, 2008. 1–8. [doi: 10.1109/cvpr.2008.4587652]
- [6] 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]
- [7] Donoho DL. Compressed sensing. IEEE Trans. on Information Theory, 2006,52(4):1289–1306. [doi: 10.1109/TIT.2006.871582]
- [8] Yu K, Zhang T, Gong Y. Nonlinear learning using local coordinate coding. In: Proc. of the Advances in Neural Information Processing Systems, Vol.22. 2009. 2223–2231. <http://papers.nips.cc/paper/3875-nonlinear-learning-using-local-coordinate-coding.pdf>
- [9] Wei CP, Chao YW, Yeh YR, Wang YCF. Locality-Sensitive dictionary learning for sparse representation based classification. Pattern Recognition, 2013,46(5):1277–1287. [doi: 10.1016/j.patcog.2012.11.014]
- [10] Wang JJ, Yang JC, Yu K, Lü F, Huang TS, Gong YH. Locality-Constrained linear coding for image classification. In: Proc. of the 2010 IEEE Int'l Conf. on Computer Vision. Washington: IEEE Computer Society Press, 2010. 3360–3367. [doi: 10.1109/CVPR.2010.5540018]
- [11] Hu WM, Li X, Luo WH, Zhang XQ, Maybank S, Zhang ZF. Single and multiple object tracking using log-Euclidean Riemannian subspace and block-division appearance model. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012,34(12):2420–2440. [doi: 10.1109/TPAMI.2012.42]
- [12] Pang Y, Yuan Y, Li X. Gabor-Based region covariance matrices for face recognition. IEEE Trans. on Circuits and Systems for Video Technology, 2008,18(7):989–993. [doi: 10.1109/TCSVT.2008.924108]
- [13] Bak S, Corvee E, Bremond F, Thonnat M. Boosted human re-identification using Riemannian manifolds. Journal Image and Vision Computing, 2012,30(6-7):443–452. [doi: 10.1016/j.imavis.2011.08.008]
- [14] Harandi MT, Sanderson C, Wiliem A, Lovell BC. Kernel analysis over Riemannian manifolds for visual recognition of actions, pedestrians and textures. In: Proc. of the 2012 IEEE Workshop on the Applications of Computer Vision. Washington: IEEE Computer Society Press, 2012. 433–439. [doi: 10.1109/WACV.2012.6163005]
- [15] Tuzel O, Porikli F, Meer P. Pedestrian detection via classification on Riemannian manifolds. IEEE Trans. on Pattern Analysis and Intelligence, 2008,30(10):1713–1727. [doi: 10.1109/TPAMI.2008.75]
- [16] Arsigny V, Fillard P, Pennec X, Ayache N. Geometric means in a novel vector space structure on symmetric positive-definite matrices. SIAM Journal on Matrix Analysis and Applications, 2006,29(1):328–347. [doi: 10.1137/050637996]

- [17] Tosato D, Farenzena M, Cristani M, Spera M, Murino V. Multi-Class classification on Riemannian manifolds for video surveillance. In: Proc. of the 11th European Conf. on Computer Vision. Heraklion: Eurographics Association Press, 2010. 378–391. [doi: 10.1007/978-3-642-15552-9\_28]
- [18] Carreira J, Caseiro R, Batista J, Sminchisescu C. Semantic segmentation with second-order pooling. In: Proc. of the 12th European Conf. on Computer Vision. Heraklion: Eurographics Association Press, 2012. 430–443. [doi: 10.1007/978-3-642-33786-4\_32]
- [19] Harandi MT, Sanderson C, Hartley R, Lovell BC. Sparse coding and dictionary learning for symmetric positive definite matrices: A kernel approach. In: Proc. of the 12th European Conf. on Computer Vision. Heraklion: Eurographics Association Press, 2012. 216–229. [doi: 10.1007/978-3-642-33709-3\_16]
- [20] Jayasumana S, Hartley R, Salzmann M, Li H, Harandi MT. Kernel methods on the Riemannian manifold of symmetric positive definite matrices. In: Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2013. 73–80. [doi: 10.1109/CVPR.2013.17]
- [21] Harandi MT, Sanderson C, Shen CH, Lovell B. Dictionary learning and sparse coding on Grassmann manifolds: An extrinsic solution. In: Proc. of the 2013 IEEE Int'l Conf. on Computer Vision. Washington: IEEE Computer Society Press, 2013. 3120–3127. [doi: 10.1109/ICCV.2013.387]
- [22] Sanin A, Sanderson C, Harandi MT, Lovell BC. Spatio-Temporal covariance descriptors for action and gesture recognition. In: Proc. of the IEEE Workshop on Applications of Computer Vision (WACV). Washington: IEEE Computer Society Press, 2013. 103–110. [doi: 10.1109/WACV.2013.6475006]
- [23] Jayasumana S, Hartley R, Salzmann M, Li HD, Harandi MT. Combining multiple manifold-alued descriptors for improved object recognition. In: Proc. of the Digital Image Computing: Techniques and Applications. 2013. 1–6. [doi: 10.1109/DICTA.2013.6691493]
- [24] Li PH, Wang QL, Zuo WM, Zhang L. Log-Euclidean kernels for sparse representation and dictionary learning. In: Proc. of the 2013 IEEE Int'l Conf. on Computer Vision. Washington: IEEE Computer Society Press, 2013. 1601–1608. [doi: 10.1109/ICCV.2013.202]
- [25] Chen SS, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit. Journal of SIAM Review, 2001,43(1):129–159. [doi: 10.1137/S003614450037906X]
- [26] Candes E, Tao T. Near optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans. on Information Theory, 2006,52(12):5406–5425. [doi: 10.1109/TIT.2006.885507]
- [27] Lee H, Battle A, Raina R, Andrew YN. Efficient sparse coding algorithms. In: Proc. of the Advances in Neural Information Processing Systems. 2006. 801–808. <http://papers.nips.cc/paper/2979-efficient-sparse-coding-algorithms.pdf>
- [28] 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]
- [29] Sivalingam R, Boley D, Morellas V, Papanikolopoulos N. Tensor sparse coding for region covariances. In: Proc. of the 11th European Conf. on Computer Vision. Berlin: Eurographics Association Press, 2010. 722–735. [doi: 10.1007/978-3-642-15561-1\_52]
- [30] 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. Washington: IEEE Computer Society Press, 2010. 188–195. [doi: 10.1109/AVSS.2010.71]



姜伟(1969—),男,辽宁大连人,博士,副教授,CCF 会员,主要研究领域为机器学习,模式识别,图像处理.



李克秋(1972—),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为云计算,图像处理.



毕婷婷(1990—),女,硕士,主要研究领域为机器学习,模式识别,图像处理.



杨炳儒(1943—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为知识发现,柔性建模.