

# 基于黎曼流形的平面目标识别

李广伟<sup>1,2,3</sup> 刘云鹏<sup>1,3</sup> 尹健<sup>4</sup> 史泽林<sup>1</sup>

**摘要** 平面目标识别中的几何形变可用射影变换群描述. 与紧致李群  $SO(n, \mathbf{R})$  不同, 正则化的射影变换群, 即非紧致李群  $SL(n, \mathbf{R})$  上由黎曼度量决定的黎曼指数映射不同于由单参数子群决定的李群指数映射. 基于黎曼流形优化算法得到取值于特殊线性群  $SL(3, \mathbf{R})$  的样本的内蕴均值和协方差矩阵, 并依此构建李群正态分布. 利用此先验知识, 根据贝叶斯定理进行简单背景下的平面目标的识别实验. 结果表明, 利用射影变换群的统计特性可有效提高平面目标识别的成功率.

**关键词** 黎曼流形, 目标识别, 射影变换, 李群, 流形优化

**DOI** 10.3724/SP.J.1004.2010.00465

## Planar Object Recognition Based on Riemannian Manifold

LI Guang-Wei<sup>1,2,3</sup> LIU Yun-Peng<sup>1,3</sup> YIN Jian<sup>4</sup> SHI Ze-Lin<sup>1</sup>

**Abstract** The geometric warps between planar objects can be represented by projective Lie groups. Compared with the compact Lie group  $SO(n, \mathbf{R})$ , the Riemannian exponential map on the noncompact Lie group  $SL(n, \mathbf{R})$  determined by a Riemannian metric is usually different from the Lie group exponential map determined by one-parameter subgroups. We compute the samples' intrinsic means on the special linear group  $SL(3, \mathbf{R})$  based on the Riemannian manifold optimum algorithm and propose the Lie group norm distribution. The test results of the planar object recognition in the simple background, which is based on the full Bayes statistical rule, have shown that the proposed algorithm with the intrinsic statistical property of the projective group may improve the rate of recognition effectively.

**Key words** Riemannian manifold, object recognition, projective transformation, Lie group, manifold optimization

目标识别和分类是计算机高层视觉研究的重要内容, 通常被看作基于某种度量准则的匹配问题, 譬如, 局部或全局的特征匹配、视觉不变量匹配、变形模板匹配等<sup>[1]</sup>. 从现实世界的概率性而言, 更应该从统计推理的角度理解和研究目标识别问题.

目标识别的挑战来自诸多方面, 其中, 目标的几何形变是影响目标识别率的重要因素. 同一个二维景物, 摄像机在不同的地点、从不同角度拍摄, 得到的图像的几何形状不同. 可以证明, 当物体具有一个平坦平面时, 射影变换精确地模型化了在摄像机或物体运动后, 同一物体的图像之间的关系<sup>[2]</sup>, 如图 1 所示. 与仿射变换不同, 射影变换不存在有限不变矩, 数学模型的高度非线性使得基于射影变换模型

的平面目标识别显得非常困难. 基于此, 研究者一方面提出了许多基于仿射变形的目标识别方法<sup>[3]</sup>, 这些方法对于一定程度的非线性也表现出较好的鲁棒性; 另一方面提出射影无限不变矩<sup>[4]</sup>的概念, 其对于平面目标识别的深入研究具有重要意义. 然而, 这些方法都没有考虑变换群的统计分布对目标识别的影响.

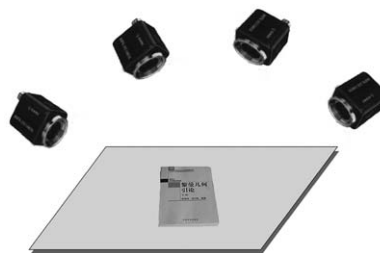


图 1 不同视角下平面目标的图像间关系用射影变换群描述

Fig. 1 A planar object imaged from different views can be described by the projective group

目标识别中的几何形变可用李变换群描述, 一般用特殊欧氏群  $SE(n, \mathbf{R})$  刻画目标的刚体变形, 用仿射变换群  $GA(n, \mathbf{R})$  刻画目标的仿射变形, 用射影变换群描述目标的射影变形, 而用拓扑群刻画更复杂的非刚体变形. 在统计模式识别实践中, 要获得变换群的经验概率分布, 一般首先求取有限样本的均值和方差. 均值是一个看似简单、实际却非常复杂

收稿日期 2009-02-16 录用日期 2009-11-06  
Manuscript received February 16, 2009; accepted November 6, 2009

国家自然科学基金 (60603097), 中国科学院国防科技创新基金 (CXJJ-65) 资助

Supported by National Natural Science Foundation of China (60603097) and National Defense Innovation Foundation of Chinese Academy Sciences (CXJJ-65)

1. 中国科学院沈阳自动化研究所 沈阳 110016 2. 青岛大学管理科学与工程系 青岛 266071 3. 中国科学院研究生院 北京 100049 4. 空军装备研究院总体所 北京 100076

1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016 2. Department of Management Science and Engineering, Qingdao University, Qingdao 266071 3. Graduate University of Chinese Academy of Sciences, Beijing 100049 4. Research Institute on General Development and Argumentation of Equipment of Air Force, Beijing 100076

的概念,有很多不同的定义方法,本文采用 Karcher 黎曼均值的概念<sup>[5]</sup>,其在统计推理、形状分析、数据平滑与插值、运动平均、图像融合和跟踪模板构建中都有重要应用。

一般采用流形优化的思想求取 Karcher 黎曼均值. 基于流形思想的几何优化算法的核心是将约束集视为流形,将传统的约束优化问题转化为流形上的无约束优化问题. 近年来,人们把这种算法应用于不同的流形以解决不同的问题. 特别是基于李群理论<sup>[6]</sup>和黎曼流形的优化算法<sup>[7]</sup>已开始逐步应用于信号处理、计算机视觉和模式识别等领域. 目前,关于黎曼均值的研究主要集中在某些具有特殊结构的流形,如球流形  $S(n)$ <sup>[8]</sup>,特殊正交群  $SO(n, \mathbf{R})$ <sup>[9]</sup>,特殊欧氏群  $SE(n, \mathbf{R})$ <sup>[10]</sup>,Grassman 流形  $G(n, k)$ <sup>[11]</sup>,Stiefel 流形<sup>[12]</sup>以及正定对称矩阵流形  $Sym^+(n, \mathbf{R})$ <sup>[13-14]</sup>等. 这些流形分别在动画设计、流形插值、刚体目标的分布设置、人脸识别、协方差特征目标跟踪和扩散张量核磁共振成像研究中有着广泛的应用. 由于这些流形都和连通紧致李群  $SO(n, \mathbf{R})$  直接相关,比较容易获取其黎曼指数映射和逆映射,因此,关于这些流形的黎曼均值已得到较为深入的分析. 其中,对球流形和特殊正交群的 Karcher 均值的理论研究比较成熟<sup>[15]</sup>,实践应用也逐渐深入. 但是  $SL(3, \mathbf{R})$  是非紧致群,没有双不变黎曼度量,单位元处的黎曼指数映射不再对应单参数子群,而且其逆映射没有封闭的解释式,这给射影变换群分布的相关研究带来很大的困难.

流形上的多变量正态分布依赖参数空间的微分几何特性. 本文在黎曼流形的统计分析理论框架下<sup>[16]</sup>,首先通过赋予  $SL(3, \mathbf{R})$  黎曼度量得到黎曼流形,获取黎曼指数映射;然后,运用流形优化算法求解  $SL(3, \mathbf{R})$  的数字特征,并构造特殊线性李群正态分布;最后,理论分析了这种分布的合理性和优越性,同时进行了简单背景下的平面目标识别实验,佐证了获取变换群的先验知识有助于提高目标识别的成功率.

## 1 问题描述

### 1.1 基于贝叶斯统计的目标识别

本文的目标识别研究是在灰度恒常性假设下展开的,即在图像平面建立适当的坐标系后,同一平面目标在投影前、后的两幅图像的灰度值  $I(X)$  和  $I(T(X; \mathbf{r}))$  的关系可表示为

$$I(T(X; \mathbf{r})) = I(X) \quad (1)$$

这里

$$T(X; \mathbf{r}) = \frac{1}{r_7x + r_8y + 1} \begin{bmatrix} r_1x + r_2y + r_5 \\ r_3x + r_4y + r_6 \end{bmatrix} \quad (2)$$

表示第二幅图像相对于第一幅图像的投影变换,其中,  $X = (x, y)^T$  是图像位置坐标,  $\mathbf{r} = (r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8)^T$  是参数向量. 我们经常用

$$c = |r_7| + |r_8| \quad (3)$$

来度量式 (2) 所表示的投影变换的非线性程度.

这里,投影变换阵为

$$T(\mathbf{r}) = \begin{bmatrix} r_1 & r_2 & r_5 \\ r_3 & r_4 & r_6 \\ r_7 & r_8 & 1 \end{bmatrix} \quad (4)$$

对其正则化,即令其行列式之值为 1,得到特殊线性群  $SL(3, \mathbf{R})$ ,此处的正则化没有改变投影变换的自由度. 作为一般线性群  $GL(n, \mathbf{R})$  的子群,特殊线性群  $SL(3, \mathbf{R})$  既是一个连通(非单连通)的非紧致矩阵李群,又是一个 8 维的微分流形. 所有矩阵的迹为零的三阶矩阵构成  $SL(3, \mathbf{R})$  在单位元处的切空间,赋予李括号运算后的切空间称为  $SL(3, \mathbf{R})$  的李代数,记为  $sl(3, \mathbf{R})$ ,其一组基向量为

$$\begin{aligned} E_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & E_2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ E_3 &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & E_4 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\ E_5 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & E_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ E_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & E_8 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned} \quad (5)$$

现假设要识别  $N$  类平面刚体目标. 已知输入的是射影变换类型为  $T$  的待识别目标  $M$  的图像  $I$ ,由贝叶斯定理知,目标识别问题转化为求后验概率  $P(M, T|I)$ ,即

$$P(M, T|I) = \frac{P(I|M, T)P(T|M)P(M)}{P(I)} \quad (6)$$

其中,  $P(I)$  是归一化常数;  $P(I|M, T)$  是似然概率,表示图像和真实场景的拟合程度,应该这样理解,如果知道了具体的目标和射影变换类型,那么我们可以

以从相应的特殊视角重建图像,但由于场景的知识,如光源、目标表面的反射特性或图像系统的噪声和模型的精度等未知,这种重建会产生一个潜在的变化,概率  $P(I|M, T)$  就是对这种变化的度量;  $P(M)$  是待识别目标的类先验知识;而  $P(T|M)$  则表示特殊线性群的概率分布,是待识别目标关于变换的先验知识. 先验知识对于贝叶斯统计推理是非常重要的,但获取先验知识往往相当困难. 正如图像的统计模型是视觉知识和图像规律性的内在表示一样,变换群的概率分布也应该采用流形的内蕴统计特性. 本文就是要研究  $SL(3, \mathbf{R})$  上的样本点的正态分布,为目标识别提供先验知识,这必须首先求其试验样本的内蕴均值和方差,本文采用基于黎曼流形的优化方法求之.

## 1.2 分布不变性对于目标识别的意义

众所周知,欧氏空间是平直空间,正态分布满足平移变换不变性,即  $x \sim N(\mu, \Sigma)$ , 则  $x + t \sim N(\mu + t, \Sigma)$ . 平移不变性存在的根源在于欧氏空间的度量对平移的不变性. 样本点之间的距离的不变性对于目标识别是非常重要的,所以流形上的正态分布应该具有相应的对变换群作用的不变性.

图2示意了变换不变性的涵义. 假设有一幅手机图像  $I$ , 图像  $I_1$  和  $I_2$  分别是图像  $I$  经过变换  $f$  和  $g$  变换得到. 如果观众离屏幕的距离变为原来的一半,其所看的图像应该是尺度发生变化的图像  $J_1$  和  $J_2$ . 我们自然希望产生图像  $I_1$  和  $I_2$  的变换  $f$  和  $g$  之间的距离应该和产生图像  $J_1$  和  $J_2$  的变换  $hf$  和  $hg$  之间的距离相同. 在变换的空间,如果用  $h$  表示尺度变换,则不变性条件变为  $d(f, g) = d(hf, hg)$ , 其中  $d$  是由流形附加的度量结构所定义的距离. 因此,度量结构必须是内蕴的,和外在的欧氏空间无关.



图2  $I_1$  和  $I_2$  分别是手机图像  $I$  经  $f$  和  $g$  变换而得;  $J_1$  和  $J_2$  分别是图像  $I_1$  和  $I_2$  经  $h$  变换而得(不变性条件意味着图像  $I_1$  和  $I_2$  的距离应当与  $J_1$  和  $J_2$  的距离相同)

Fig.2  $I_1$  is transformed from the mobile telephone image  $I$  by  $f$ ,  $I_2$  by  $g$ ,  $J_1$  by  $hf$ ,  $J_2$  by  $hg$  (The invariance condition is that the distance between  $I_1$  and  $I_2$  should be the same as between  $J_1$  and  $J_2$ .)

对于  $SL(3, \mathbf{R})$  而言,其上可自然赋予左不变黎曼度量.

## 2 基于黎曼流形的李群正态分布

李群和黎曼几何是本文的理论基础,与本文有关的内容参见附录,详细内容参阅文献 [17–18].

### 2.1 $SL(3, \mathbf{R})$ 上的不变度量和测地线

每个矩阵李群上都可以赋予左不变黎曼度量,这个不变的黎曼结构是由单位元处的切空间的内积决定.事实上,设李群  $M$  在单位元处的切空间为  $T_e M$ ,其上的内积为正定双线性映射  $g_e : T_e M \times T_e M \rightarrow \mathbf{R}$ . 记  $L_\sigma$  是  $M$  上的左移动:  $L_\sigma : m \rightarrow \sigma m$ ,  $\sigma, m \in M$ . 因为  $L_\sigma$  是一个微分同胚,所以其在单位元  $e$  处的微分  $T_e L_\sigma : T_e M \rightarrow T_\sigma M$  是向量空间同构. 所以,可用  $T_e L_\sigma$  在每一点的切空间定义内积

$$g_\sigma : T_\sigma M \times T_\sigma M \rightarrow \mathbf{R} \quad (7)$$

$$g_\sigma(\eta, u) = g_e(L_\sigma^{-1}(\eta), L_\sigma^{-1}(u))$$

其中,  $\eta, u \in T_\sigma M$ . 容易看出,  $g_\sigma$  在  $M$  上依赖于  $\sigma$  光滑变化,由此可定义一个黎曼度量

$$g(\eta, u) = g_\sigma(\eta, u) \quad (8)$$

其中,  $\sigma \in M$ ;  $\eta, u \in T_\sigma M$ , 这个度量是左不变黎曼度量. 对于  $SL(3, \mathbf{R})$  而言,我们定义其在单位元处的切空间的内积为附录 A 中的式 (A4).

紧致李群(如  $SO(3, \mathbf{R})$ ) 上存在双不变黎曼度量,由群结构所定义的李群指数映射对应着测地线.  $SL(3, \mathbf{R})$  上不存在双不变黎曼度量,关于左不变黎曼度量式 (A4) 的测地线  $\gamma(t)$  可通过求解如下的矩阵微分方程组得到

$$\begin{cases} \gamma'(t) = \gamma(t)X(t) \\ X'(t) = X^T(t)X(t) - X(t)X^T(t) \\ \gamma(0) = \gamma_0 \\ X(0) = X_0 \end{cases} \quad (9)$$

其中,  $X(t) \in sl(3, \mathbf{R})$ .

矩阵微分方程组 (9) 的一个解析解<sup>[19]</sup> 为

$$\gamma(t) = \gamma_0 \exp(-tX_0^T) \exp(t(X_0 + X_0^T)) \quad (10)$$

这里补充说明的是,中国数学家王宪钟在 1969 年基于嘉当分解得到连通半单李群上的测地线<sup>[20]</sup>,这里的特殊线性群是其特例.

于是,在单位元  $e$  处的黎曼指数映射可定义为

$$\text{Rexp}(X) = \exp(-X^T) \exp(X^T + X) \quad (11)$$

其中,  $X(t) \in sl(3, \mathbf{R})$ . 在单位元  $e$  处的逆黎曼指数映射, 即黎曼对数映射可自然定义为

$$\text{Rlogg}(x) = \text{Rexp}^{-1}(x), \quad x \in SL(3, \mathbf{R}) \quad (12)$$

其没有封闭的显式表达式.

## 2.2 $SL(3, \mathbf{R})$ 的黎曼均值和协方差矩阵算法

为了获得特殊线性李群  $SL(3, \mathbf{R})$  上的样本点的正态分布, 要求其黎曼均值和协方差矩阵.

完备的黎曼流形  $M$  上的点集  $\{x_1, x_2, \dots, x_n\}$  的 Karcher 黎曼均值  $\mu$  定义为

$$\mu = \text{E}[\{x_i\}] = \arg \min_{y \in M} \left( \frac{1}{n} \sum_{i=1}^n d^2(y, x_i) \right) \quad (13)$$

其中,  $d(y, x_i)$  为  $M$  上的黎曼测地距离函数.

在紧致李群上, 基于李群理论的一阶优化算法和基于黎曼流形的梯度下降算法是一致的, 因为由黎曼度量结构决定的黎曼指数映射和由群结构所定义的李群指数映射相同, 所以文献 [9] 求取特殊正交群  $SO(n, \mathbf{R})$  经验分布的数字特征的方法是合理的. 由于  $SL(3, \mathbf{R})$  的非紧致性, 基于黎曼优化求其经验分布的数字特征较为合理.

由定义在黎曼流形上的函数在法坐标系下的泰勒展开式 (B5) (见附录 B), 并结合梯度式<sup>[21]</sup>

$$\text{grad} \left( \sum_{i=1}^n d^2(y, x_i) \right) = -2 \sum_{i=1}^n \text{Rlog}_y x_i \quad (14)$$

可得到基于黎曼流形的求取 Karcher 黎曼均值的梯度下降算法为

$$\mu_{t+1} = \text{Rexp}_{\mu_t} \left( \sum_{i=1}^n \left( \frac{1}{n} \text{Rlog}_{\mu_t} x_i \right) \right) \quad (15)$$

当样本点  $x_i$  距离单位元  $e$  较远时, 所得结果的误差很大. 因此, 在求取  $SL(3, \mathbf{R})$  上的黎曼均值时, 应首先将每个样本左乘初始值之逆, 得到其残差样本, 并将残差样本映射到李代数空间; 然后在李代数空间将映射得到的样本点求均值后再映射回李群, 并左乘初始值得到新均值近似值; 再把所有样本左乘新均值之逆后, 映射至李代数空间; 最后在李代数空间求均值后映射回李群, 并左乘均值再次得到新均值. 反复迭代直至残差样本的均值满足终止条件. 其中, 残差均值的计算式为

$$\Delta \mu_{t+1} = \text{Rexp} \left( \frac{1}{n} \sum_{i=1}^n \text{Rlogg}(x_i) \right) \quad (16)$$

为获得协方差矩阵, 我们设计一个 (非唯一) 从  $sl(3, \mathbf{R})$  到  $\mathbf{R}^8$  的映射:  $f: X \in sl(3, \mathbf{R}) \rightarrow \mathbf{R}^8$ . 由

于  $sl(3, \mathbf{R})$  上的每一点  $X$  都可表示为其基 (6) 的线性组合, 注意到矩阵的迹为零, 可简单地选择映射为

$$f(X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}) = (x_{11}, x_{33}, x_{12}, x_{13}, x_{21}, x_{23}, x_{31}, x_{32})^T \quad (17)$$

于是, 经验协方差阵  $\Sigma_{ml}$  为

$$\Sigma_{ml} = \frac{1}{n-1} \sum_{i=1}^n f_m(X_i) f_i(X_i) \quad (18)$$

其中, 李代数空间的样本点  $X_i$  是由  $SL(3, \mathbf{R})$  上的样本点左乘均值之逆后通过黎曼对数映射而得到;  $f_k(X_i)$  表示  $X_i$  的线性表示的第  $k$  ( $k = 1, 2, \dots, 8$ ) 个分量.

综合以上所述, 求取  $SL(3, \mathbf{R})$  上的样本点的经验均值和协方差阵的算法如下:

**输入.**  $x_1, x_2, \dots, x_n \in SL(3, \mathbf{R})$  和精度  $\varepsilon > 0$ .

**输出.** 黎曼均值  $\mu$  和协方差阵  $\Sigma$ .

**初始化.**  $\mu = x_1$

**while** ( $\|\text{Rlogg} \Delta \mu\| \geq \varepsilon$ )

**do**

$$\Delta x_i = \mu^{-1} x_i$$

$$\Delta \mu = \text{Rexp} \left( \frac{1}{n} \sum_{i=1}^n \text{Rlogg} \Delta x_i \right)$$

$$\mu = \mu \cdot \Delta \mu$$

**end**

$$X_i = \text{Rlogg}(\mu^{-1} x_i)$$

$$\Sigma_{ml} = \frac{1}{n-1} \sum_{i=1}^n f_m(X_i) f_i(X_i)$$

这里, 有几个细节问题需要说明: 1) 因无法获取黎曼对数映射的解析表达式, 本文采用最小二乘法求之, Matlab 代码为:  $\text{Rlogg}(x) = \text{lsqnonlin}(@\text{Rexp}, 0, [], [], [], x)$ . 2) 由于黎曼指数映射和逆映射是局部同胚, 所以随机获取的样本点应比较集中, 不能过于分散, 否则算法容易发散. 一般而言, 黎曼均值是存在的, 但并不唯一. 我们未能对所设计的黎曼均值算法的收敛域作定量描述, 因为这需要深入研究黎曼指数映射的内射半径、凸半径、 $SL(3, \mathbf{R})$  的截面曲率以及黎曼度量的选择对算法的影响等较为复杂的课题, 留待以后探讨. 3) 同欧氏空间相似, 也存在黎曼流形上的大数定律, 即随着有效样本数的增多, 经验均值应该无限接近于真

实均值<sup>[22]</sup>, 这说明在目标识别获取先验分布知识时, 应采集足够数量的训练样本.

### 2.3 由 $GL(3, \mathbf{R})$ 投影回 $SL(3, \mathbf{R})$ 算法

由于计算的数值误差, 常导致迭代过程中的  $SL(3, \mathbf{R})$  的点不能精确地落至原来的流形, 而落在一般线性群  $GL(3, \mathbf{R})$  上, 因此必须把每次迭代的结果投影回  $SL(3, \mathbf{R})$ , 即寻找  $y \in SL(3, \mathbf{R})$ , 使得

$$\min_{y \in SL(3, \mathbf{R})} \|x - y\|_F^2 \quad (19)$$

其中,  $x$  是在迭代计算过程中得到的一般线性群.

设矩阵  $x$  和  $y$  的奇异值分解分别为  $x = u\Omega v$ ,  $y = u_1\Omega_1 v_1$ . 这里  $u, u_1, v, v_1$  均为正交矩阵;  $\Omega = \text{diag}\{a_1, a_2, a_3\}$ ;  $\Omega_1 = \text{diag}\{x_1, x_2, x_3\}$ , 其中,  $a_1 < a_2 < a_3$ ,  $x_1 < x_2 < x_3$ . 若令  $u_2 = u^T u_1$ ,  $v_2 = v^T v_1$ ,  $s = \begin{cases} \text{diag}\{1, 1, 1\}, & \text{若 } \det(x) > 0, \\ \text{diag}\{1, 1, -1\}, & \text{若 } \det(x) < 0. \end{cases}$  则有

$$\begin{aligned} \|a - b\|_F^2 &= \text{tr}(a - b)(a - b)^T = \\ &\|\Omega\|^2 + \|\Omega_1\|^2 - 2\text{tr}(u_2\Omega_1 v_2^T \Omega) \geq \\ &\|\Omega\|^2 + \|\Omega_1\|^2 - 2(\text{sgn}(\det(a)))a_3x_3 - \\ &2 \sum_{i=1}^2 a_i x_i = \|\Omega - s\Omega_1\|_F^2 \end{aligned} \quad (20)$$

在实验中发现,  $x$  的行列式在绝大多数情形下均为正值, 此时式 (20) 等价于

$$\min_{\{x_1, x_2, x_3\}} \|\Omega - \Omega_1\|_F^2 \quad (21)$$

且  $y$  的最优解为  $y^* = u\Omega_1 v$ . 考虑到  $x_1 x_2 x_3 = 1$ , 则式 (21) 成立的必要条件为

$$\begin{cases} (x_1)^2 - (x_2)^2 + a_2 x_2 = 0 \\ (x_1)^4 (x_2)^2 - a_1 (x_1)^3 (x_2)^2 + a_3 x_1 x_2 - 1 = 0 \end{cases} \quad (22)$$

由式 (22) 可以求出  $x_i, i = 1, 2, 3$ .

本文实验时直接采用式 (23) 作为投影结果, 实验证明其也可取得和最优解相似的算法性能, 但其结构更简单.

$$y = u \begin{bmatrix} \sqrt[3]{\frac{(a_3)^2}{a_1 a_2}} & 0 & 0 \\ 0 & \sqrt[3]{\frac{(a_2)^2}{a_1 a_3}} & 0 \\ 0 & 0 & \sqrt[3]{\frac{(a_1)^2}{a_2 a_3}} \end{bmatrix} v \quad (23)$$

在更严格意义上, 还需要研究由  $GL(3, \mathbf{R})$  投影回  $sl(3, \mathbf{R})$  算法, 本文从略.

### 2.4 李群正态分布

在求得  $SL(3, \mathbf{R})$  的样本的黎曼均值  $\mu$  和协方差阵  $\Sigma$  后, 基于黎曼流形的李群正态分布 (Riemannian Lie group norm distribution, RLN) 可表示为

$$\begin{aligned} \varphi(x) &\propto \exp -\frac{1}{2} f(\text{Rlog}(\mu^{-1}x))^T \times \\ &\Sigma^{-1} f(\text{Rlog}(\mu^{-1}x)) \end{aligned} \quad (24)$$

我们知道, 欧氏空间的正态分布满足平移不变性, 即  $x \sim N(\mu, \Sigma)$ , 则  $x + t \sim N(\mu + t, \Sigma)$ . 由黎曼均值和协方差阵的计算过程可知, 基于黎曼流形的李群正态分布 (RLN) 保持对变换作用的左不变, 即若随机变量  $x \sim \text{RLN}(\mu, \Sigma)$ , 则变换后的变量  $g \circ x \sim \text{RLN}(g \circ \mu, \Sigma)$  ( $\circ$  为群运算符号), 这是由在  $SL(3, \mathbf{R})$  上附加的左不变度量结构所决定的. 我们是通过把在  $\mu$  处定义的切空间上的正态分布映射至  $SL(3, \mathbf{R})$  而得到特殊线性群上的正态分布.

如果在欧氏空间  $\mathbf{R}^8$  中建立  $SL(3, \mathbf{R})$  上样本的统计规律, 构建欧氏空间的多变量正态分布 (Norm distribution, N), 有两个缺点: 1) 由于没有准确利用 8 个独立参数所在流形的内在结构, 由此所得到的均值一般不会落在原来的流形上, 是不准确的; 2) 由此所定义的分具有对平移变换的不变性, 而对投影变换则不具备变换不变性, 在目标识别时必会增加误判率. 另外, 如果在以上求黎曼均值的算法中, 用李群对数映射代替黎曼对数映射, 也可以根据优化算法得出均值, 我们称据此构建的分布为基于李群指数映射的李群正态分布 (Lie group norm distribution, LN), 即

$$\varphi(x) \propto \exp -\frac{1}{2} f(\log(\mu^{-1}x))^T \Sigma^{-1} f(\log(\mu^{-1}x)) \quad (25)$$

因为李群指数映射和其逆映射的解析式相对简单, 所以构建基于李群指数映射的李群正态分布需要较少的计算量. 但是, 其迭代算法并非沿着最短测地线进行, 所以均值的计算也会受到影响; 另外, 基于李群理论的优化算法在求取数字特征时并不涉及度量概念, 也就无从谈及变换分布对变换群的不变性.

总之, 基于黎曼流形视角考虑目标识别问题是必要的, 因为先验正态分布的变换不变性依赖于流形的黎曼度量结构, 亦即依赖于切空间内积的不变性. 另外,  $SL(3, \mathbf{R})$  的非紧性使得优化算法的设计应该独立于群结构, 而分析其黎曼或联络结构.

### 3 实验与分析

从以上分析可以看出, 求取样本均值是构建李群正态分布最为关键的一步. 因此, 在实验 1 中, 我们分别给出了两组样本的均值图像. 根据图像变形是否显著, 实验 1 中分别采取了不同的方法求取两组样本图像相对于各自模板图像的投影变换阵. 在实验 2 中, 我们以三类工具钳的识别为例, 说明基于黎曼流形的平面目标识别算法, 同时综合比较了几种不同的识别方法的识别成功率. 编程采用 Matlab 语言.

#### 3.1 实验

**实验 1.** 分别求取 200 幅木箱图像和 100 幅钱夹图像的平均图像. 对前者而言, 因为图像间的变形不显著, 所以采用李群二阶最小化 (Efficient second-order minimization, ESM) 图像配准算法<sup>[23]</sup> 可快速获得每幅图像和模板图像间的最优单应阵; 对后者而言, 因为钱夹图像投影变形显著, 所以首先采用 SIFT 特征检测器<sup>[24]</sup> 检测出兴趣点, 再用 RANSAC 算法计算其和模板图像的最优的单应变换阵. 一般而言, 均值图像不会是样本图像. 图 3 给出 8 幅木箱的样本图像, 中间为基于第 2.2 节的黎曼均值算法所得到的均值图像. 图 4 给出 8 幅钱夹图像, 中间为基于第 2.2 节的黎曼均值算法所得到的均值图像.

**实验 2.** 识别张角分别为大、中、小的三类工具钳 (见图 5), 描述基于黎曼流形的平面目标识别算法, 验证获取基于黎曼流形的李群正态分布先验概率可有效提高平面目标的识别成功率.

实验过程:

1) 获取张角分别为大、中、小的三类铁钳图像各 300 幅, 并用 SIFT 特征检测器检测出每一幅图像的特征点. 实际实验时获取的图像应多于 300 幅,

因为并不是每一幅图像都可获取稳定的特征点.

2) 可任意指定三类工具钳图像的模板图像, 一般选择下视图 (在实际获取模板图像时, 一般通过卫星或其他途径获得下视图). 用 RANSAC 算法分别计算每类工具钳的其他 299 幅图像和各自模板图像的最优投影变换阵, 并将投影变换矩阵归一化为特殊线性矩阵.

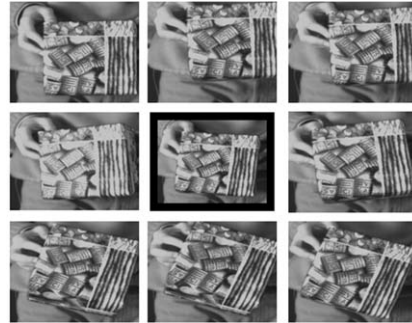


图 3 不同视角下的木箱图像 (中间为平均图像)

Fig. 3 Wooden box images from different views (The middle is their mean image.)



图 4 不同视角下的钱夹图像 (中间为平均图像)

Fig. 4 Wallet images from different views (The middle is their mean image.)



图 5 大、中、小三种不同张角的工具钳图像

Fig. 5 Tongs images with different angles

3) 采用 5 倍交叉验证法测试目标识别的成功率. 每次实验的训练集和测试集的设计如下: 将张角为大、中、小的三类铁钳各 300 幅图像分别随机分为五组, 其中, 每类铁钳的四组共计 240 幅图像用作训练, 即基于第 2.2 节的黎曼均值和协方差阵算法求出这 240 幅图像的黎曼均值和协方差矩阵, 从而得到每类工具钳的基于黎曼流形的李群正态分布 (RLN). 为了便于比较, 还需求出每类工具钳的基于李群指数映射的李群正态分布 (LN) 以及基于欧氏空间  $\mathbf{R}^8$  的多变量正态分布 (N). 如果每类工具钳剩下的一组共 60 幅图像中不包含模板图像, 便将其全部作为测试集, 即该次实验实际参加的有效测试样本数为 60; 如果包含模板图像, 鉴于模板图像的特殊性, 我们将其剔除, 在这种情况下, 该次实验实际参加计算的有效测试样本数为 59. 理论上, 三类工具钳的五倍交叉验证应该计算 125 次, 计算量很大. 在实际计算时, 我们随机选取 27 次, 即, 对每类工具钳而言, 选取五组中的三组依次作为测试集.

4) 对于每一次识别实验, 分别计算以下七种具有代表性的目标识别情形. 这里  $P(M_i)$  ( $i = 1, 2, 3$ ) 相等.

第一种情形, 以最小拟合匹配误差作为识别基准, 即只计算式 (6) 中的  $P(I|M, T)$ , 其表示图像和真实场景的度量. 我们采用如下的度量准则<sup>[25]</sup>

$$P(I|M, T) = K \exp(-E) \quad (26)$$

其中,  $K$  为归一化常数,  $E = \|I - M(T)\|^2$ . 这里  $I$  表示待识别的图像,  $M$  理解为待识别目标的模板,  $T$  理解为二者之间的变换.

第二种情形, 将特殊线性群看作是 8 个变量的欧氏空间, 求出其在欧氏空间  $\mathbf{R}^8$  的多变量正态分布 (N)  $P(T|M)$ , 仅用此先验知识进行目标识别.

第三种情形, 用李群优化算法得到基于指数映射的李群正态分布 (LN)  $P(T|M)$ , 仅用此先验知识进行目标识别.

第四种情形, 用本文算法求出基于黎曼流形的

李群正态分布 (RLN)  $P(T|M)$ , 仅用此先验知识进行目标识别.

第五种情形, 欧氏空间  $\mathbf{R}^8$  的多变量正态分布 (N)  $P(T|M)$  结合最小拟合匹配误差  $P(I|M, T)$ , 用贝叶斯式 (6) 进行目标识别.

第六种情形, 基于李群指数映射的李群正态分布 (LN)  $P(T|M)$  结合最小拟合匹配误差  $P(I|M, T)$ , 用贝叶斯式 (6) 进行目标识别.

第七种情形, 基于黎曼流形的李群正态分布 (RLN)  $P(T|M)$  结合最小拟合匹配误差  $P(I|M, T)$ , 用贝叶斯式 (6) 式进行目标识别.

5) 对于以上七种情形中的每一种情形, 均按以下方法统计数据. 首先, 分别求取张角为大、中、小的三类工具钳在 27 次实验的每次实验中的目标识别成功率 (成功识别的图像数目除以有效测试样本数目 59 或 60); 其次, 分别求取张角为大、中、小的三类工具钳的 27 次实验的目标识别成功率的代数平均值 (27 次实验的目标识别率之和除以 27), 并将其作为三类工具钳的最终的成功率; 最后, 七种情形中的每种情形的工具钳的识别成功率是三类工具钳的最终的成功率的代数平均值 (每类工具钳的最终成功率之和除以 3). 某 27 次随机实验的实验结果统计见表 1.

### 3.2 分析

由表 1 可以看出, 拟合匹配误差的识别结果很差, 三类工具钳的平均成功识别率约为 0.3278, 与随机识别的结果相差无几, 这是因为射影变换的参数变量之多和高度非线性之故. 事实上, 即使是不同的物体也可能有非常相似的图像. 如果平面目标仅有轮廓知识, 而没有其他的先验知识, 目标几乎是不可识别的. 可以证明, 定义在平面上的所有封闭曲线, 在投影变换下几乎是等价的<sup>[26]</sup>, 这也是基于投影不变矩识别的根本局限. 因此, 关于目标识别的先验知识是非常重要的.

由表 1 可以看出, 当结合正态分布先验知识后, 目标的识别率都有所提高. 其中, 结合欧氏空间正态

表 1 七种方法的工具钳目标识别实验结果比较

Table 1 Comparison of tong recognition results of seven methods

识别方法	成功率(大)	中	小	平均)
最小拟合匹配误差	0.3167	0.3325	0.3342	0.3278
基于欧氏空间 $\mathbf{R}^8$ 的正态分布 (N)	0.3015	0.3128	0.3021	0.3055
基于李群指数映射的李群正态分布 (LN)	0.4337	0.4618	0.4545	0.4500
基于黎曼流形的李群正态分布 (RLN)	0.4538	0.4685	0.4610	0.4611
最小拟合匹配误差 + 基于欧氏空间 $\mathbf{R}^8$ 正态分布 (N)	0.4152	0.4324	0.4358	0.4278
最小拟合匹配误差 + 基于李群指数映射的李群正态分布 (LN)	0.5151	0.5096	0.5086	0.5111
最小拟合匹配误差 + 基于黎曼流形的李群正态分布 (RLN)	0.5256	0.5490	0.5421	0.5389

分布 (N) 后, 三类工具钳的平均成功识别率可升为 0.4278; 结合基于李群指数映射的李群正态分布 (LN) 后, 目标识别的成功率升为 0.5111; 而结合基于黎曼流形的李群正态分布 (RLN) 后, 目标识别的成功率升为 0.5389. 这说明, 充分利用变换参数所在流形的内在特性, 可有效提高目标的识别率. 在流形上开展统计研究, 不仅是统计理论自身发展的需要, 也具有重要的应用意义.

另外, 从表 1 还可以看出, 如果仅利用先验分布进行识别, 则基于黎曼流形的李群正态分布 (RLN) 的目标识别的成功率与基于李群指数映射的李群正态分布 (LN) 的目标识别的成功率相差不大, 这是因为在单位变换附近, 黎曼指数映射可以用李群指数映射逼近的缘故; 虽然  $SL(3, \mathbf{R})$  是非紧致群, 但在样本比较集中的条件下, 可以看作是一个局部的紧致群. 但是, 两者都远大于基于欧氏空间正态分布 (N) 的目标识别成功率, 这是因为欧氏空间正态分布在特殊线性群流形上不具有对特殊线性变换的不变性. 遗憾的是, 实验中没有发现射影变换的非线性程度的度量系数  $c = |r_7| + |r_8|$  和识别成功率之间的内在规律, 在后续的研究中, 应该采用其他可操作的度量, 如文献 [27] 在研究 Affine-SIFT 时所提出的量化不同视角图像的方法.

最后指出的是, 拟合误差  $P(I|M, T)$  和先验知识  $P(T|M)$  对于目标识别都非常重要, 但对各类噪声或遮挡等的鲁棒性主要体现在拟合误差  $P(I|M, T)$ . 一般所用的普通的平面目标识别方法, 如基于不变矩的识别<sup>[4]</sup>、基于多尺度自卷积仿射不变识别<sup>[3]</sup> 和其他的基于局部特征的目标识别方法<sup>[24]</sup> 等, 本质都是对  $P(I|M, T)$  的不同描述. 本文所构建的  $P(T|M)$  的作用主要体现在辅助识别. 如果训练样本的图像质量很差或者背景非常复杂, 关于变换群分布的先验知识必然受到影响, 因为获取精确的投影变换矩阵将变得困难, 而这是另外的一个课题了.

## 4 结论

模式识别和计算机视觉领域中的许多问题都需要在适当的受限函数空间提出统计模型. 本文根据黎曼流形优化理论, 在内蕴统计分析框架下, 基于黎曼流形构建特殊线性群的正态分布, 并利用此先验知识, 辅助基于贝叶斯推理的平面目标识别. 简单背景下的平面目标识别实验说明, 如果能结合流形上的点的内蕴统计特性, 适当地引入变换群的概率统计分布, 将有助于提高目标识别的成功率.

研究射影变形下的自动目标识别和跟踪在导弹制导、无人机导航和图像拼接等领域有重要意义, 本文进行了初步的探索. 目前, 对射影变换非线性敏感度的度量 and 特殊线性群  $SL(3, \mathbf{R})$  的几何学研究还

有待深入, 更需要对不同度量的选择对函数优化或目标识别的影响问题进行深入探讨. 在此基础上, 我们将对本文所提出的李群正态分布的合理性问题做更进一步的理论分析和实验验证.

## 附录 A 李群上的两类特殊指数映射

### A 1 李群指数映射

赋予群以光滑流形结构得到李群, 其群运算记为  $\circ$ . 李群  $M$  在单位元  $e$  处的切空间记为  $T_e M$ ; 设  $m \in M$ ,  $X_e \in T_e M$ , 通过左移动  $X_m = (L_m)_* X_e$  得到左不变向量场  $X$ ; 过点  $e$  的  $X$  的积分曲线就是单参数子群, 记为  $\gamma_X(t)$ ; 赋予李括号运算后的  $(T_e M, [\cdot, \cdot])$  为  $M$  的李代数, 记作  $\Lambda(M)$ . 设  $X \in \Lambda(M)$ ,  $t \in [0, 1]$ , 则李群指数映射定义为

$$(X, t) \rightarrow \exp(tX) = \gamma_X(t) \quad (\text{A1})$$

李群  $M$  的李代数  $\Lambda(M)$  内一定存在一个包含 0 的开凸集  $W$ , 使得指数映射  $\exp$  是  $W$  到  $M$  内的开核  $U$  的一个微分同胚. 由此, 可定义李群指数映射的逆映射, 即李群对数映射为  $W \rightarrow U: \log(y) = \exp^{-1}(y)$ , 见图 A1.

对于矩阵李群, 李群指数映射是普通意义上的代数矩阵指数, 即

$$\exp X = \sum_{n=0}^{\infty} \frac{X^n}{n!} \quad (\text{A2})$$

当矩阵没有负实部的特征值时, 矩阵李群指数映射存在逆映射, 称之为李群主对数映射, 即

$$\log x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (x - I)^n \quad (\text{A3})$$

其中  $I$  表示单位矩阵.

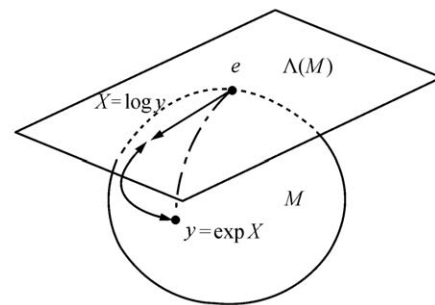


图 A1 李群指数映射和逆映射

Fig. A1 Lie group exponential map and its inverse map

### A 2 黎曼指数映射

若在矩阵李群  $M$  的每一点  $p$  的切空间  $T_p M$  定义内积

$$\langle A, B \rangle = \text{tr}(AB^T) \quad (\text{A4})$$

则赋予度量结构后的李群成为黎曼流形  $M$ . 设  $\gamma: [a, b] \rightarrow M$  是  $M$  中连接点  $p$  和  $q$  的一条光滑曲线, 其长度定义为

$$L(\gamma) = \int_a^b \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt \quad (\text{A5})$$

则连接  $M$  上两点  $p, q$  的最短距离  $d(p, q)$  称为黎曼测地距离, 所经过的路径称为最短测地线.



设曲线  $\gamma(t; p, v)$ ,  $t \in [0, 1]$  是满足条件  $\gamma(0) = p$ ,  $\gamma'(0) = v$  的黎曼流形  $M$  上的一条测地线;  $B_p(\varepsilon)$  是  $T_p M$  中以原点为中心、以  $\varepsilon$  为半径的开球, 定义  $M$  在  $p$  点的黎曼指数映射为

$$\text{Rexp}_p : B_p(\varepsilon) \rightarrow M, \quad \text{Rexp}_p v = \gamma(1; p, v) \quad (\text{A6})$$

显然,  $\text{Rexp}_p$  满足  $d(p, \text{Rexp}_p v) = |v|$ . 黎曼流形的每一点  $p$  的充分小的邻域和该点的切空间  $T_p M$  可通过黎曼指数映射建立微分同胚, 如图 A2 所示. 于是, 过点  $p$  的黎曼指数映射的逆映射可定义为  $\text{Rlog}_p(\cdot)$ . 本文记在单位元  $e$  处的黎曼指数映射及其逆映射分别为  $\text{Rexp}(\cdot)$  和  $\text{Rlog}(\cdot)$ .

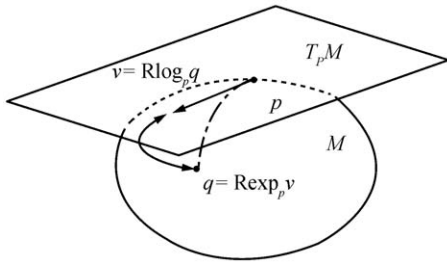


图 A2 黎曼指数映射和逆映射

Fig. A2 Riemannian exponential map and its inverse map

对于紧致李群而言, 其上存在双不变黎曼度量, 具有非负的截面曲率, 单参数子群 (或平移) 对应着流形上的测地线; 对于非紧致李群而言, 其上不存在双不变黎曼度量, 由群结构定义的李群指数映射和由黎曼度量定义的黎曼指数映射不同, 即单参数子群不再对应着测地线, 一般采用变分方法获得.

## 附录 B 李群上的几何优化

为简单起见, 定义在李群和黎曼流形的光滑函数的泰勒展开公式均在规范坐标下进行. 关于基于李群理论和黎曼流形的函数优化算法的收敛性分析, 可参见文献 [28–30].

### B1 基于李群理论的几何优化

设  $E_i$  ( $i = 1, \dots, n$ ) 是李群  $M$  的李代数  $\Lambda(M)$  的基向量, 设  $t$  是李群  $M$  上一点,  $x$  是  $t$  邻域内一点, 由李群的第一类规范坐标知, 存在  $\Phi = \sum_i^n v_i E_i \in \Lambda(M)$ , 使得

$$x = t \circ \exp(\Phi) = t \circ \exp\left(\sum_i^n v_i E_i\right) \quad (\text{B1})$$

其中  $v = (v_1, v_2, \dots, v_n)^T$ . 设  $\varphi(\cdot)$  是定义在李群  $M$  上的函数, 则其在  $t$  点的泰勒展开式为

$$\varphi(t \circ \exp(\Phi)) = \varphi(t) + J_t^\varphi v + \frac{1}{2} v^T H_t^\varphi v + O(\|v\|^3) \quad (\text{B2})$$

其中

$$\begin{aligned} [J_t^\varphi]_i &= \frac{\partial}{\partial v_i} \varphi(t \circ \exp(\Phi))|_{v=0} \\ [H_t^\varphi]_{ij} &= \frac{\partial^2}{\partial v_i \partial v_j} \varphi(t \circ \exp(\Phi))|_{v=0} \end{aligned} \quad (\text{B3})$$

类比欧氏空间中的优化算法, 可构造各种内蕴优化算法, 如梯度下降法、拟牛顿迭代等. 比如, Newton-Raphson 内蕴迭代为

$$t \leftarrow t \circ \exp(\Phi) \quad (\text{B4})$$

其中  $v$  可由方程  $H_t^\varphi v = -[J_t^\varphi]^T \varphi(t)$  解得.

### B2 基于黎曼流形的几何优化

设  $f(\cdot)$  是定义在黎曼流形  $M$  上的光滑函数, 则其在法坐标系下的函数泰勒展开式为

$$\begin{aligned} f(\text{Rexp}_\mu v) &= f(\mu) + \text{grad} f(v) + \\ &\quad \frac{1}{2} \text{Hess} f(v, v) + O(\|v\|^3) \end{aligned} \quad (\text{B5})$$

其中  $v \in T_\mu^M$ . 类比欧氏空间中的优化算法, 可构造各种内蕴优化算法. 比如, 当满足适当条件时, 牛顿算法的迭代公式为

$$u \leftarrow \text{Rexp}_u (H_f^{-1}(\text{grad} f)) \quad (\text{B6})$$

这里  $H_f^{-1}$  是  $H_f(v)(w) = \text{Hess} f(v, w)$  的逆映射.

## References

- Zitova B, Flusser J. Image registration methods: a survey. *Image and Vision Computing*, 2003, **21**(11): 977–1000
- Mann S, Picard R W. Video orbits of the projective group: a simple approach to featureless estimation of parameters. *IEEE Transactions on Image Processing*, 1997, **6**(9): 1281–1295
- Esa R, Mikko S, Janne H. Affine invariant pattern recognition using multiscale autoconvolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(6): 908–918
- Tomas S, Janne F. Projective moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(10): 1364–1367
- Karcher H. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 1977, **30**(5): 509–541
- Owren B, Welfert B. The Newton iterations on Lie groups. *BIT Numerical Mathematics*, 2000, **40**(1): 121–145
- Smith S T. Geometric Optimization Methods for Adaptive Filtering [Master dissertation], Harvard University, USA, 1993
- Buss S R, Fillmore J P. Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, 2001, **20**(2): 95–126
- Moakher M. Means and averaging in the group of rotations. *SIAM Journal of Matrix Analysis and Applications*, 2002, **24**(1): 1–16
- Tron R, Vidal R, Terzis A. Distributed pose averaging in camera networks via consensus on SE(3). In: Proceedings of the 2nd ACM and IEEE International Conference on Distributed Smart Cameras. Stanford, USA: IEEE, 2008. 1–10
- Chang J M. Classification on the Grassmannians: Theory and Applications [Ph. D. dissertation], Colorado State University, USA, 2008
- Turaga P, Veeraraghavan A, Chellappa R. Statistical analysis on Stiefel and Grassmann manifolds with applications in computer vision. In: Proceedings of Conference on Computer Vision and Pattern Recognition. Anchorage, USA: IEEE, 2008. 1–8

- 13 Pennec X, Fillard P, Ayache N. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 2006, **66**(1): 41–66
- 14 Tuzel O, Porikli F, Meer P. Pedestrian detection via classification on Riemannian manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(10): 1713–1727
- 15 Manton J H. A globally convergent numerical algorithm for computing the centre of mass on compact Lie group. In: Proceedings of the 8th International Conference on Control, Automation, Robotics and Vision. Kunming, China: IEEE, 2004. 2211–2216
- 16 Pennec X. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 2006, **25**(1): 127–154
- 17 Helgason S. *Differential Geometry, Lie Groups, and Symmetric Spaces*. New York: Academic Press, 1978
- 18 Chen Wei-Heng, Li Xing-Xiao. *An Introduction to Riemannian Geometry (I, II)*. Beijing: Peking University Press, 2002  
(陈维桓, 李兴校. 黎曼几何引论 (上、下册). 北京: 北京大学出版社, 2002)
- 19 Fletcher P, Whitaker R. Riemannian metrics on the space of solid shapes. In Proceedings of International Workshop on Mathematical Foundations of Computational Anatomy. Copenhagen, Denmark: Med Image Comput Assist Intev, 2006. 1–11
- 20 Wang H C. Discrete nilpotent subgroups of Lie groups. *Journal Differential Geometry*, 1969, **3**(3-4): 481–492
- 21 Subbarao R. Robust Statistics over Riemannian Manifolds for Computer Vision [Ph.D. dissertation], The State University of New Jersey, USA, 2008
- 22 Bhattacharya R, Patrangenaru V. Large sample theory of intrinsic and extrinsic sample means on manifolds (II). *Annals of Statistics*, 2005, **33**(3): 1225–1259
- 23 Benhimane S, Malis E. Homography-based 2D visual tracking and servoing. *International Journal of Robotics Research*, 2007, **26**(7): 661–676
- 24 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
- 25 Srivastava A, Jermyn I. Riemannian analysis of probability density functions with application in vision. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 234–243
- 26 Astrom K. Fundamental limitations on projective invariants of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, **17**(1): 77–81
- 27 Morel M, Yu G S. ASIFT: a new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2009, **2**(2): 438–469

- 28 Mahony R. Optimization Algorithms on Homogeneous Spaces: with Applications in Linear Systems Theory [Ph.D. dissertation], Australian National University, Australian, 1994
- 29 Mahony R, Manton J H. The geometry of the Newton method on non-compact Lie group. *Journal of Global Optimization*, 2002, **23**(3-4): 309–327
- 30 Yang Y. Globally convergent optimization algorithm on Riemannian manifolds: uniform framework for unconstrained and constrained optimization. *Journal of Optimization Theory and Applications*, 2007, **132**(2): 245–265



李广伟 中国科学院沈阳自动化研究所博士研究生. 主要研究方向为目标跟踪和识别, 鲁棒控制. 本文通信作者.

E-mail: liguangwei@sia.cn

(LI Guang-Wei Ph.D. candidate at Shenyang Institute of Automation, Chinese Academy of Sciences. His research interest covers object tracking and recognition, and robust control. Corresponding author of this paper.)



刘云鹏 中国科学院沈阳自动化研究所博士研究生. 主要研究方向为目标跟踪和识别. E-mail: ypliu@sia.cn

(LIU Yun-Peng Ph.D. candidate at Shenyang Institute of Automation, Chinese Academy of Sciences. His research interest covers object tracking and recognition.)

尹健 空军装备研究院总体所研究员. 主要研究方向为武器系统总体设计和评估、目标跟踪和识别.

E-mail: yinyin610@sohu.com

(YIN Jian Professor at Research Institute on General Development and Argumentation of Equipment of Air Force. His research interest covers general design and evaluation of weapon system, and object tracking and recognition.)



史泽林 中国科学院沈阳自动化研究所研究员. 主要研究方向为图像处理、模式识别和智能控制. E-mail: zlshi@sia.cn

(SHI Ze-Lin Professor at Shenyang Institute of Automation, Chinese Academy of Sciences. His research interest covers image processing, pattern recognition, and intelligent control.)