

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012
Vol.17 No.

4

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年4月 第17卷 第4期(总第192期)

目次

综述

图像分割中的模糊聚类方法 李旭超, 刘海宽, 王飞, 白春艳(447)

图像处理和编码

GPU 辅助的希尔伯特变换轮廓术 周波, 赵小敏, 王东平(459)

引入连续性强度和置信度因子的快速图像修复 李开宇, 孙玉刚(465)

自适应的有效非局部图像滤波 许光宇, 檀结庆, 钟金琴(471)

改进的 PMD 距离图像超分辨率重建算法 张旭东, 沈玉亮, 胡良梅, 陈菁菁(480)

压缩感知在 Micro-CT 图像超分辨重建中的应用 王丽艳, 韦志辉, 罗守华, 顾宁(487)

对偶四元数单片空间后方交会算法 姬亭, 盛庆红, 王惠南, 刘微微(494)

利用运动强度判据的高效自适应运动估计算法 郭晓珉, 姚睿, 刘智跃, 王友仁(504)

图像分析和识别

核空间散度阈值法 吴成茂(512)

近邻自适应局部尺度的谱聚类算法 孔万增, 孙昌思核, 张建海, 胡三清, 杨灿(523)

LUV 色彩空间中多层次化结构 Nyström 方法的自适应谱聚类算法 刘雅蓉, 汪西莉(530)

结合图像增强的心血管内超声中-外膜边缘检测 邱璇, 黄靖, 杨丰, 邢栋, 涂圣贤(537)

融合图像特征的一致点匹配方法及其应用 张久楼, 李春丽, 冯前进, 陈武凡, 阳维(546)

图像理解和计算机视觉

多蚁群动态协作优化的道路图像分割算法 林丽莉, 周文晖(553)

篮球比赛视频中持球队员行为预测	王千,夏利民,谭论正(560)
利用 Principal Warps 评估颅面几何相似度	朱新懿,耿国华,温超(568)

计算机图形学

图形处理器空间插值并行算法的实现	赵艳伟,程振林,董慧,方金云(575)
------------------------	---------------------

虚拟现实与增强现实

面向 GPU 的批 LOD 地形实时绘制	张兵强,张立民,张建廷(582)
----------------------------	------------------

遥感图像处理

光学遥感舰船目标识别方法	杜春,孙即祥,李智勇,滕书华(589)
自适应超完备字典学习的 SAR 图像降噪	杨萌,张弓(596)

第 18 届中国遥感大会征文通知	封 2
第 33 届亚洲遥感会议征文通知	封 2

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 4 期

2012 年 4 月 16 日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
 中国图象图形学学会
 北京应用物理与计算数学研究所
主 编 李小文
编辑出版 《中国图象图形学报》编辑出版委员会
 北京 9718 信箱 邮编 100101
 电子信箱:jig@irsa. ac. cn
 电话:010-68407995 010-82614429
 网 址:www. cjig. cn
印刷装订 北京北林印刷厂
广告经营许可证 京朝工商广字第 0346 号
总 发 行 北京报刊发行局
订 购 全国各地邮局
国外发行 中国国际图书贸易总公司
 (中国国际书店)
 (北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
 CAS China Society of Image and Graphics
 Institute of Applied Physics and Computational
 Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
 of Journal of Image and Graphics
 (P. O. Box 9718, Beijing 100101, China)
 E-mail: jig@irsa. ac. cn
Distributed by Beijing Bureau for Distribution of Newspapers
 and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
 (P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

第 18 届中国遥感大会征文通知

“第 18 届中国遥感大会”将于 2012 年 10 月 19 日-23 日在武汉召开。本届会议由中国遥感委员会主办,中国测绘学会摄影测量与遥感专业委员会和武汉大学承办。会议将围绕“遥感—全方位的社会服务”这一宗旨,以遥感学界院士与知名专家的特邀报告,分会场专题技术交流与技术讲座,重点项目研讨汇报、技术展览,新技术与新产品发布,专业委员会理事会等多种形式开展,同时举行“第 7 届中国青年遥感辩论会”和“第 2 届全国高分辨率遥感数据处理与应用研讨会”。

会议将全方位地展示遥感(RS)、全球定位系统(GPS)、地理信息系统(GIS)等方面的最新成果,为专家、学者和政府主管部门搭建联系纽带,为研发和用户提供技术交流平台,共同促进遥感科技的发展、遥感产业化的推进和大遥感体系的建立。

本届会议围绕大会主题将就遥感新理论、技术、方法和应用进行征文,范围包含但不限于以下方面:

- 1) 国家遥感中长期发展战略、国际遥感前沿与进展;
- 2) 航天、航空、低空、地面遥感技术及系统;
- 3) 光学、红外、高光谱及激光遥感技术;
- 4) 主、被动微波及雷达遥感技术;

- 5) 数字摄影测量与制图;
- 6) 高分辨率遥感数据处理与应用;
- 7) 地理空间数据处理技术与方法;
- 8) 地理国情监测(土地、农业、林业、矿产、环境、地质及水资源等);
- 9) 海洋、气象与全球变化;
- 10) 遥感、地理信息系统与导航定位系统(3S)集成与应用;
- 11) 智慧城市与数字地球;
- 12) 深空探测与行星测绘;
- 13) 教育、培训与社会公共事业。

征文采用在线方式投稿;

投稿要求:论文内容不涉密,且未在国内外学术刊物或正式学术会议上发表过;被录用的全文将收入大会论文集(送 ISTP 检索),并精选 70~90 篇口头报告论文编辑出版英文 SPIE 会议文集;大会将评选青年优秀论文(参加口头报告),论文将直接进入英文 SPIE 会议文集。

论文摘要截止日期为 2012 年 5 月 15 日,全文截稿日期为 2012 年 6 月 15 日。

会议相关信息,请查阅会议网址:<http://rsgis.whu.edu.cn/18ccrs/index.html>

“第 18 届中国遥感大会”组委会

第 33 届亚洲遥感会议征文通知

“第 33 届亚洲遥感会议”将由泰国地理信息和空间技术发展局(GISTDA)、科技部(MOST)和亚洲遥感协会(AARS)联合主办,于 2012 年 11 月 26-30 日,在泰国芭堤雅市宗滴恩酒店举行。这是亚洲遥感协会每年一届的系列学术会议。本届大会征文包括传感器与平台、算法和图像处理、GIS 与 Web GIS、全球导航卫星系统、灾害、自然资源、环境科学、教育和宣传、健康科学、制图、其他等方面。

会议重要日期:

- 论文摘要提交截止:2012 年 5 月 15 日;
 - 论文接收通知:2012 年 7 月 1 日;
 - 论文全文提交截止:2012 年 9 月 30 日;
 - 网上注册截止:2012 年 10 月 26 日;
 - 会议召开日期:2012 年 11 月 26-30 日。
- 会议还将组织学生专场和技术展览,其他信息请访问会议网站:<http://acrs2012.gistda.or.th>

与往年一样,中国遥感委员会仍将鼓励中国遥感科研人员和企事业单位参加会议,并组团参加学术交流和会议展览。

中国遥感委员会

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 4 April 2012

Contents

Review

The survey of fuzzy clustering method for image segmentation Li Xuchao, Liu Haikuan, Wang Fei, Bai Chunyan (447)

Image Processing and Coding

GPU assisted Hilbert transform profilometry Zhou Bo, Zhao Xiaomin, Wang Dongping (459)

Fast image inpainting algorithm introducing continuous strength and confidence factor Li Kaiyu, Sun Yugang (465)

Adaptive efficient non-local image filtering Xu Guangyu, Tan Jieqing, Zhong Jinqin (471)

Improved super-resolution reconstruction algorithm for PMD range image
..... Zhang Xudong, Shen Yuliang, Hu Liangmei, Chen Jingjing (480)

Image superreconstruction for Micro-CT based on compressed sensing Wang Liyan, Wei Zhihui, Luo Shouhua, Gu Ning (487)

Dual quaternion of space resection with single-image Ji Ting, Sheng Qinghong, Wang Huinan, Liu Weiwei (494)

Efficient adaptive motion estimation algorithm based on motion intensity Guo Xiaomin, Yao Rui, Liu Zhiyue, Wang Youren (504)

Image Analysis and Recognition

Divergence thresholding method in kernel space Wu Chengmao (512)

Spectral clustering based on neighboring adaptive local scale
..... Kong Wanzeng, Sun Changsihe, Zhang Jianhai, Hu Sanqing, Yang Can (523)

Adaptive spectral clustering algorithm based on Nyström method with multi-level structure in LUV color space
..... Liu Yarong, Wang Xili (530)

Image enhancement based media-adventitia border detection in intravascular ultrasound images
..... Qiu Xuan, Huang Jing, Yang Feng, Xing Dong, Tu Shengxian (537)

Coherent point drift registration combined with image feature and its application
..... Zhang Jiulou, Li Chunli, Feng Qianjin, Chen Wufan, Yang Wei (546)

Image Understanding and Computer Vision

Dynamic multi-colony ant cooperative optimization schemes for road image segmentation
..... Lin Lili, Zhou Wenhui (553)

Behavior prediction of ball carriers in basketball match videos Wang Qian, Xia Limin, Tan Lunzheng (560)

Estimate of craniofacial geometry shape similarity based on principal warps
..... Zhu Xinyi, Geng Guohua, Wen Chao (568)

Computer Graphics

Realization of GPU parallel spatial interpolation method
..... Zhao Yanwei, Cheng Zhenlin, Dong Hui, Fang Jinyun (575)

Virtual Reality and Augmented Reality

GPU-based real-time terrain rendering algorithm using batched LOD
..... Zhang Bingqiang, Zhang Limin, Zhang Jianting (582)

Remote Sensing Image Processing

Method for ship recognition using optical remote sensing data
..... Du Chun, Sun Jixiang, Li Zhiyong, Teng Shuhua (589)

SAR images de-speckling algorithm via an adaptive over-complete learning dictionary
..... Yang Meng, Zhang Gong (596)

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2012)04-0560-08

论文引用格式: 王千, 夏利民, 谭论正. 篮球比赛视频中持球队员行为预测[J]. 中国图象图形学报, 2012, 17(4): 560-567

篮球比赛视频中持球队员行为预测

王千, 夏利民, 谭论正

中南大学信息科学与工程学院, 长沙 410075

摘要: 体育视频分析是近年来计算机视觉领域备受关注并且十分具有挑战性的研究方向。提出一种新颖的篮球比赛中持球队员行为预测方法。针对篮球比赛视频中背景复杂, 球员移动速度快, 头部图像分辨率低的特点, 提出运用协方差描述子对持球队员的多种有效头部视觉特征融合, 构成黎曼流形, 再将其映射到切空间中, 用训练好的多类 LogitBoost 进行持球队员的头部姿态识别, 从而确定持球队员的视野范围。根据持球队员视野范围内双方球员分布, 提出基于人工势能场的球员信息量, 对持球队员投篮、传球、运球 3 种行为进行预测。最后对实况篮球赛进行测试, 验证了这种方法的有效性。

关键词: 协方差描述子; 切空间; 多类 LogitBoost; 人工势能场

Behavior prediction of ball carriers in basketball match videos

Wang Qian, Xia Limin, Tan Lunzheng

School of Information Science and Engineering, Central South University, ChangSha 410075, China

Abstract: Sports video analysis has received much attention in recent years and is a challenging research direction in the field of computer vision. A novel prediction method for the behavior of ball carriers in basketball matches is proposed in this paper. Aiming at the cluttered background, fast motion of the sportsmen, and the low resolution of the head images in basketball match videos, we propose the adoption of a covariance descriptor to fuse multiple visual features of the head region, which can be represented as Riemannian Manifolds. Then we map the covariance descriptor to the tangent space and complete the head pose classification through the trained multiclass Logitboosts directly in this space for determining the range of vision of the ball carrier. According to the distribution of all the sportsmen within the range of vision of the ball carrier, we predict the behavior of him—shooting, passing, and dribbling, through sportsmen information based on artificial potential field (APF). Finally, the tests on the basketball match videos verify the effectiveness of our method.

Key words: covariance descriptor; tangent space; multiclass LogitBoost; artificial potential field

0 引言

国内外体育专家经过长期实践与研究表明, 在运动训练中引入数字视频技术, 可以极大提高训练效率。如何利用计算机图像处理、视频分析等技术手段, 分析运动员训练和比赛的视频图像, 得到运动

员和教练员感兴趣的各类信息, 实现科学指导训练, 并已成为当前热点^[14]。目前计算机技术在体育视频中的应用主要体现在运动员的动作识别与分析。Ramasso 等人^[5]利用 TBM (transferable belief model) 对田径比赛中的背越式跳高、撑杆跳高、三级跳远等动作进行了识别; Roh 等人^[6]对低分辨率网球比赛视频中球员的各种姿态进行了识别; Min 等人^[7]则

收稿日期: 2011-04-28; 修回日期: 2011-09-22

基金项目: 国家自然科学基金项目 (50808025); 国家教育部博士点基金项目 (20090162110057)

第一作者简介: 王千 (1986—), 男, 中南大学信息科学与工程学院信息与通信工程专业硕士研究生, 主要研究方向为图像处理和计算机视觉。E-mail: utopiamail@qq.com

通过跟踪芭蕾舞舞者某些身体部位关键点,获得其舞蹈轨迹,实现对舞蹈动作的自动分析;Tong 等人^[8]实现了对游泳比赛视频中4种泳姿的识别;Zhu 等人^[9]综合考虑视频中球场上的多种信息,达到了自动分析足球比赛战术打法的目的。

篮球比赛是一项广为开展且具代表性的团体运动,比赛的输赢取决于双方队员间的密切配合,而队员间配合的好坏又取决于持球队员的决策。因此,对篮球比赛中持球队员的投篮、传球、运球3种行为进行分析和预测具有重要的指导意义。篮球比赛中持球队员行为预测属于复杂行为分析,位于运动人体行为分析的高层。它不仅要求识别持球队员的行为,还要考虑持球队员与场上其他队员、对方篮筐的各种关系等。当前国内的研究主要集中在人体的行为识别方面,进行行为预测的研究较少。Duque 等人^[10]提出一种基于N-ARY TREES分类器的异常行为预测方法,后来又提出一种实时性更强、具有自适应能力的动态有向图(DOG)^[11],对异常行为进行预测。但在现实中,人体的行为预测比行为识别往往更有意义,比如在银行、机场、政府大楼等对安全比较敏感的场所,智能监控系统若能提前预测可疑人员的犯罪行为,及时向安保人员发出报警,则可以有效预防犯罪的发生,同时也减少了大量人力、物力和财力的投入;在体育比赛中,若能提前得知持球队员的行为决策,便能起到预测比赛态势发展的作用。

本文提出一种新颖的对篮球比赛视频中持球队员行为进行预测的技术。针对篮球比赛视频中背景复杂,球员移动速度快,头部图像分辨率低的特点,提出运用协方差描述子对持球队员的多种有效头部视觉特征融合,构成黎曼流形,再将其映射到切空间中,用训练好的多类LogitBoost进行持球队员的头部姿态识别,确定持球队员的视野范围。根据持球队员视野范围内的双方球员分布,利用基于人工势能的球员信息量,完成对其投篮、传球、运球3种行为的预测。

1 持球队员的视野判定

在进行持球队员的头态识别之前,以人工标记和粒子滤波器^[12]跟踪相结合的方法,确定了图像中每名球员的位置。为了对球场上持球队员的视野范围进行判定,需要进一步确定持球队员的头态即将头部的旋转角度判定为如图1所示的8个方向之

一。现有的人体头部姿态识别算法大都针对人机交互等画面质量较好、分辨率较高、人的面部头部特征可以有效被提取的情况,而篮球比赛视频中存在较多噪声,运动模糊,头部图像在画面中所占比例小,分辨率低,能被提取到的特征有限。

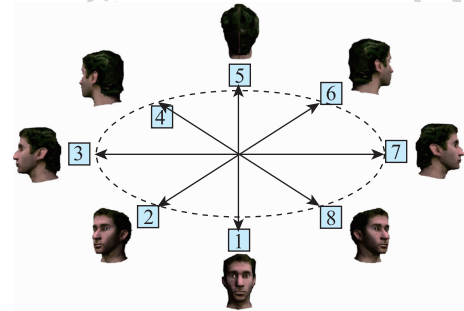


图1 8种头部姿态

Fig. 1 8 Head poses

Tuzel 等人^[13]提出一种对外界环境变化具有较强鲁棒性的人体检测算法,他把对输入的图像进行人体检测看成是一个前景(人体)和背景的二分类问题,并在INRIA和DaimlerChrysler两个数据库上实验,取得了较好的检测效果。本文在其基础上,将人体检测二分类问题扩展为多分类问题,同时完成持球队员的头部检测与头态识别,原因是两者的适用条件相同,且都是检测分类问题。

1.1 基于黎曼流形的头部特征提取

对每一个训练样本,首先抽取图像 N 个面积相等的正方形区域,这 N 个小区域允许有部分重叠。对所提取的小区域内每一个像素点,计算一个12维特征向量

$\phi = [x \ y \ R \ G \ B \ I_x \ I_y \ O \ Gab \ \{0, \pi/3, \pi/6, 4\pi/3\}]$
 式中, x, y 为像素点的坐标, R, G, B 为3颜色通道值, I_x, I_y 为一阶导数, O 为梯度方向, $Gab \ \{0, \pi/3, \pi/6, 4\pi/3\}$ 表示分别用 2×4 ,正弦频率为16,方向为 $0, \pi/3, \pi/6, 4\pi/3$ 的Gabor滤波器滤波的响应值。提取颜色特征是因为不同角度下的球员头部的肤色和非肤色所占的比例不同(如正面和背面),而Gabor滤波是为了提取头部的纹理特征。然后按照式(1)计算每个小区域所对应的协方差描述子

$$X_k^r = \frac{1}{S_k^r - 1} \sum_{i=1}^{S_k^r} (z_i - \zeta_k^r) (z_i - \zeta_k^r)^T \quad (1)$$

式中, ζ_k^r 表示第 k 个样本第 r 个小区内所有像素点特征向量的均值向量, S_k^r 为第 k 个样本第 r 个小区内像素点个数。 z_i 为该区域内第 i 个像素点的

特征向量。 \mathbf{X}_k 为一个对称的矩阵,主对角线上的元素为每个特征的方差,非对角线元素为每类特征间的相关系数。Tuzel 还提出一种利用积分图像快速计算协方差描述子的方法^[14],任何矩形图块所对应的协方差描述子可以在时间复杂度为 $O(d^2)$ 内计算出来, d 为每个像素点特征向量的维数。

由于协方差描述子不能构成向量空间,无法用经典机器学习的方法来构造分类器。但在数学上,它具有对称、半正定的结构属性。Fletcher 等人^[15]指出, $n \times n$ 大小的对称半正定矩阵的集合构成了 $n(n+1)/2$ 维线性空间的一个凸锥。又因为在计算各类图像特征时它们的方差不可能等于零,只需要考虑协方差矩阵非奇异(对称且正定)的情形。而对称正定矩阵的集合对应于对称半正定矩阵所构成凸锥的内部,该内部是一个微分流形,因此可以运用微分几何学的方法进行后续处理。Pennec 等人^[16]提出微分流形上的一种仿射不变度量。它是一个连续黎曼度量,其主要思想是:经过黎曼流形 M 上的任意一点 X ,均可做一个切空间 S_X ,并构造切空间 S_X 和黎曼流形 M 微分同胚。切空间是某一点所有切向量组成的线性空间。对切空间 S_X 中的向量 V ,可以通过指数映射 $\exp_X(V)$ 将 V 映射为黎曼流形 M 上从点 X 出发的等长同向的测地线。该指数映射 $\exp_X(V)$ 定义如下

$$Y = \exp_X(V) = X^{\frac{1}{2}} \exp(X^{-\frac{1}{2}} V X^{-\frac{1}{2}}) X^{\frac{1}{2}} \quad (2)$$

而指数映射的逆映射,即对数映射 $\log_X(Y)$,则将黎曼流形 M 上的点 X 到点 Y 的测地线映射为切空间 S_X 中等长同向的向量

$$V = \log_X(Y) = X^{\frac{1}{2}} \log(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}) X^{\frac{1}{2}} \quad (3)$$

上述指数映射和对数映射,构成了切空间 S_X 和黎曼流形 M 之间的同胚映射。通过该同胚映射,可以将黎曼流形 M 上的点转换到切空间 S_X 上进行研究。因此,将该仿射不变度量的思想用于图像特征的协方差描述子所构成的微分流形上,使其成为黎曼流形,然后转换到切空间,在线性空间上运用机器学习方法进行球员的头部检测和头态角度识别。

但是经过黎曼流形 M 上任意一点 X ,均可做一个切空间 S_X ,因此有无数个切空间存在,必须有一个统一标准的切空间作为任意点 X 的映射空间。对于欧氏空间中的一簇点 $(y_1, y_2, \dots, y_{n-1}, y_n)$,通常以这簇点的均值作为它们的最佳代表进行各种操作。对于黎曼流形 M 上的一簇点 $(X_1, X_2, \dots, X_{n-1},$

$X_n)$,类似可以推出,在这一簇点的均值处做一个切空间,可以得到流形的最优逼近。文献[16]中提出一种微分流形的 Karcher 均值

$$\mu = \operatorname{argmin}_{Y \in M} \sum_{i=1}^N d^2(X_i, Y) \quad (4)$$

公式要求所得均值点和黎曼流形 M 上的各点 X_i 之间的距离平方之和最小。其中,距离平方为

$$d^2(X, Y) = \operatorname{tr}(\log^2(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}})) \quad (5)$$

对于协方差描述子所构成的凸锥流形, Karcher 均值是唯一的,可以采用梯度下降过程 $\mu^{t+1} = \exp_{\mu^t} \left[\frac{1}{N} \sum_{i=1}^N \log_{\mu^t}(X_i) \right]$ 求解均值 μ ,当 $d(\mu^t, \mu^{t+1}) < \varepsilon$ 时,循环终止。在所求的 μ 点处做切空间 S_μ ,就可以得到黎曼流形 M 的最合理逼近。最后将投射到切空间中的向量按式(6)正交化^[16]

$$\operatorname{vec}_\mu(V) = \operatorname{vec}_i(\mu^{-\frac{1}{2}} V \mu^{-\frac{1}{2}}) \quad (6)$$

式中,令 $\mu^{-\frac{1}{2}} V \mu^{-\frac{1}{2}} = W$,则

$$\operatorname{vec}_i(W) =$$

$$[w_{1,1}, \sqrt{2}w_{1,2}, \dots, w_{2,2}, \sqrt{2}w_{2,3}, \dots, w_{n,n}] \quad (7)$$

1.2 检测分类器的训练

经过上述计算,可以将所有样本第 r 个区域的协方差描述子 X_k^r 变换为均值点处切空间中正交化的特征向量,构成特征集合 $A_r = \{x_k^r \mid 1 \leq k \leq M\}$, $1 \leq r < N$,其中 x_k^r 表示第 k 个样本第 r 个小区域经变换后所提取的特征向量。然后以 A_r 为训练集,为每个小区域训练一个多类 LogitBoost 检测分类器,最终的检测分类结果由所有小区域的 LogitBoost 投票决定,扩展后的多类 LogitBoost 训练算法步骤如下:

输入:不同外界条件下的各种不同角度的头部图片训练集与非头部图片集 $(\mathbf{X}_1^k, y_1), \dots, (\mathbf{X}_n^k, y_n)$,其中 \mathbf{X}_i^k 为第 i 个样本第 k 个小块对应的融合多种头部特征的协方差描述子, $y_i \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$,分别代表图 1 中 8 种不同的头部姿态以及非头部图像类别的标签。

1) 由梯度下降法 $\mu^{t+1} = \exp_{\mu^t} \left[\frac{1}{N} \log_{\mu^t}(X_i^k) \right]$ 计算所有输入的头部样本第 k 个小区域上协方差描述子的切空间 Karcher 均值 μ^k ,并将其通过对数映射 $\log_{\mu^k}(X_i^k)$ 映射到均值的切空间上,并用正交化的形式表示,即 $x_i^k = \operatorname{vec}_{\mu^k}(\log_{\mu^k}(X_i^k))$ 。

2) 初始化权重 $w_{ij}^k = 1/N, i = 1, \dots, n, F_j(x_i^k) =$

$0, Pr_j[x_i^k] = 1/9$ 。

3) 循环迭代得到弱分类器假设

for $l = 1, 2, \dots, L$ do
for $j = 1, 2, \dots, 9$ do

(1) 计算回归目标值和权重

$$w_{ij}^k = Pr_j[x_i^k](1 - Pr_j[x_i^k])$$

$$z_{ij}^k = \frac{y_{ij}^{k*} - Pr_j[x_i^k]}{Pr_j[x_i^k](1 - Pr_j[x_i^k])}$$

式中, $y_{ij}^{k*} = (j = y_i)$, $Pr_j[x_i^k]$ 表示训练样本 x_i^k 属于第 j 类的概率。

(2) 采用加权最小二乘法拟合弱分类器函数。

(3) 更新 $F_j(x_i^k)$

$$F_j(x_i^k) \leftarrow F_j(x_i^k) + f_{ij}(x_i^k)$$

式中,

$$f_{ij}(x_i^k) = \frac{8}{9}(g_{ij}(x_i^k) - \frac{1}{9} \sum_{d=1}^9 g_{1d}(x_i^k))$$

(4) 更新 $Pr_j[x_i^k]$

$$Pr_j[x_i^k] = \frac{e^{F_j(x_i^k)}}{\sum_{d=1}^9 e^{F_d(x_i^k)}}$$

(5) 保存

$$F_j = \{g_{ij}, \mu^k\}$$

End for

End for

4) 保存 $\{F_1, \dots, F_9\}$ 。

其中, 参数 L 为 LogitBoost 的迭代次数, $L = \max_l \{ \text{accuracy}_l(j) \geq \tau_{\text{accuracy}} \vee (\text{accuracy}_l(j) - \text{accuracy}_{l-1}(j)) \leq \tau_{\text{rate}} \vee \text{accuracy}_{\text{background}} \geq \tau_{\text{background}}, \forall j \in \{1, \dots, 8\} \}$ 。 $\text{accuracy}_l(j)$ 代表第 j 类分类器在第 l 次迭代后的正确率, $\text{accuracy}_l(j) - \text{accuracy}_{l-1}(j)$ 为学习率, 代表连续两次迭代后分类器正确率的提高程度。

1.3 检测与分类

由于我们假定已经完成了场上持球队员的检测, 并用矩形框将其标记出来, 因此对其进行头部检测和头态识别, 只需用检测窗口在身体矩形框上 $1/4$ 区域内扫描, 流程如图 2 所示。

扫描窗口在移动过程中, 分别计算窗口内每个小区域的协方差描述子, 再由 Karcher 均值点 μ^k 用对数映射式(3)投影到切空间, 并对其正交化。用所得的特征向量作为已训练好的多类 LogitBoost 的输入, 最后由所有小区域的识别结果按式(8)来统计判定头态的类别。

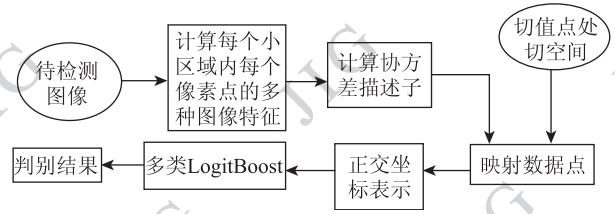


图 2 头态识别流程图

Fig. 2 The flow chart of head pose recognition

$$\text{Label} = \underset{k}{\text{argmax}} \{c_k\}, k = 1, \dots, J \quad (8)$$

式中, $c_k = \sum_{i=1}^N (LB_i = k)$ 表示判定结果为第 k 类的小区域总数, LB_i 为第 i 个小区域的 LogitBoost 分类器给出的判定类别。

头态识别算法的复杂性主要在于头部特征的计算, 而计算协方差描述子的复杂度为 $O(d^2)$, 将协方差描述子映射到切空间中所作的对数变换时间复杂度为 $O(d^3)$ ^[13], 其中 d 为提取特征的维数。

2 持球队员的行为预测

通过球员检测分类和持球队员的头态识别, 可以获得双方球员在球场上的分布和持球队员的视野信息。在此基础上, 本小节提出基于人工势能场的持球队员行为预测算法。

持球队员在进行行为选择的过程中根据自己和队友的信息量以及距离, 选择不同的行为。作为持球队员, 他可以选择投篮、传球和运球 3 种行为。而这 3 种行为中, 投篮的优先级最高, 其次是传球、运球, 即如果持球队员投篮成功的可能性大, 则选择投篮; 否则寻找位置更好的队友传球, 如果找到, 则选择传球, 否则选择运球。

2.1 基于人工势能场 (APF) 的球员信息量

比赛中进攻方的球员分为无球队员和持球队员。进攻方球员的信息量可以使用人工势能场 (APF)^[17] 的概念和方法来表示和计算。球员 x 的人工势能场可以表示如下

$$U_{\text{APF}}(x) = U_{\text{goal}}(x) + U_{\text{obs}}(x) \quad (9)$$

式中, $U_{\text{APF}}(x)$ 、 $U_{\text{goal}}(x)$ 、 $U_{\text{obs}}(x)$ 分别表示人工势能场、由目标产生的吸引势能以及由障碍产生的排斥势能。 x 描述了球员的位置信息。吸引势能可表示为

$$U_{\text{goal}}(x) = \frac{1}{2}kp \left(\frac{1}{x - x_{\text{goal}}} \right)^2 \quad (10)$$

式中, $x - x_{goal}$ 表示进攻方队员距离对方篮筐的距离, 其值越小, 吸引势能越大; 其值越大时, 表示进攻方队员距离对方篮筐越远, 吸引势能越小。排斥势能可以表示为

$$U_{obs}(x) = \begin{cases} \frac{1}{2}kd\left(\frac{1}{\rho} - \frac{1}{\rho_0}\right)^2 & \rho \leq \rho_0 \\ 0 & \rho > \rho_0 \end{cases} \quad (11)$$

障碍指的是对方队员, 式中 ρ 是进攻方队员与其周围存在的对方队员之间的距离。 ρ_0 是一个距离阈值, 超过 ρ_0 , 球员就不会受到对方队员的排斥力, 即对方队员不会对该球员造成威胁。如果进攻方球员的有效范围内 ($\rho \leq \rho_0$) 有多名对方防守队员, 则在计算该名进攻方球员的人工势能场时就需要同时考虑由多名对方队员产生的排斥势能 $U_{obs}(x)$ 。

这里需要强调的是式(10)中的参数 kp 是一个正值, 表示球员的得分能力。对于不同角色的进攻方队员, 其得分能力不同, 因此不同进攻方队员的 kp 是不同的, 前锋较大, 中锋次之, 后卫的较小。而式(11)中的 kd 是一个负值, 反映了对方防守队员对球员的威胁程度。对于不同角色的对方防守队员, 其防守能力不同, 因此不同防守队员的 $|kd|$ 也是不同的, 后卫的较大, 中锋次之, 前锋较小。

由于无法从比赛视频的图像特征中直接判断球员的角色(前锋、后卫等), 这里提出一种根据场上球员抢球和投篮次数自动更新参数 kp (投篮能力)和 kd (抢球能力)的算法。首先建立一张场上 10 名球员的投篮抢球能力表, 见表 1。

表 1 投篮抢球能力表(初始状态)

Table 1 Shooting and tackling ability (initial state)

攻防能力	球员号									
	A_1	A_2	A_3	A_4	A_5	B_1	B_2	B_3	B_4	B_5
kp	α	α	α	α	α	α	α	α	α	α
kd	β	β	β	β	β	β	β	β	β	β

设 P_{ball} 表示球所处的位置。该算法伪代码如下:

```

初始化各队员的投篮抢球能力值:
T=1; for (i=1; i≤5; i++)
 $kp_{A_i} = \alpha; kd_{A_i} = \beta; kp_{B_i} = \alpha; kd_{B_i} = \beta$ 
球员投篮和抢球能力更新:
While (T≠end)
    
```

```

If ( $P_{ball} = A_i \rightarrow P_{ball} = B_i (i = 1, \dots, 5) \cap kd_{B_i} < \theta_d$ )
//球员  $B_i$  对持球球员  $A_i$  抢球成功
 $kd_{B_i} = kd_{B_i} + \gamma$  // 球员  $B_i$  的抢球能力增加
Else If ( $P_{ball} = B_i \rightarrow P_{ball} = A_i (i = 1, \dots, 5) \cap kd_{A_i} < \theta_d$ )
//球员  $A_i$  对持球球员  $B_i$  抢球成功
 $kd_{A_i} = kd_{A_i} + \gamma$  // 球员  $A_i$  的抢球能力增加
Else if ( $A_i (i = 1, \dots, 5)$  投篮  $\cap kp_{A_i} < \theta_a$ )  $kp_{A_i} = kp_{A_i} + \tau$ 
//球员  $A_i$  投篮能力增加
Else if ( $B_i (i = 1, \dots, 5)$  投篮  $\cap kp_{B_i} < \theta_a$ )  $kp_{B_i} = kp_{B_i} + \tau$ 
//球员  $B_i$  投篮能力增加
    
```

式中, τ 和 γ 为常数, θ_d 和 θ_a 为抢球能力和投篮能力的上限值。

2.2 持球队员行为预测

每个时刻, 为每名球员 x 计算信息量 $\tau_x = U_{APF}(x)$, 持球队员根据自己与队友的信息量, 选择不同的行为。可以将持球队员的行为预测抽象为如图 3 所示的路径选择问题(其中蓝白小矩形框代表双方球员在场上的分布情况, 五角星代表对方篮筐)。

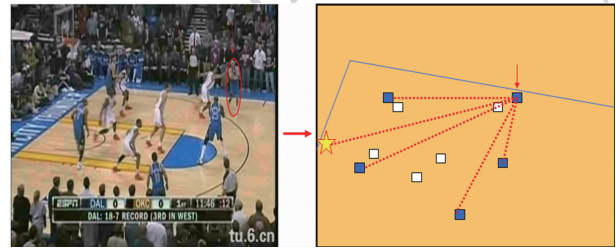


图 3 持球队员行为选择示例

Fig. 3 Example of ball carrier behavior selection

当持球队员的信息量 $\tau_c > \tau_0$ 时 (τ_0 为投篮的信息量阈值), 表示持球队员离对方篮筐较近, 并且其周围的对方球员对其造成的威胁较小, 持球队员投篮成功的可能性较大, 选择投篮; 当持球队员的信息量 $\tau_c \leq \tau_0$ 时, 表示持球队员投篮成功的可能性较小, 则选择传球或者运球。

当 $\tau_c \leq \tau_0$ 时, 持球队员是选择向队友传球还是运球, 要根据能否找到合适的接球队友而定。在寻找接球队友的过程中, 不仅要考虑该接球队友是否处在更具进球得分的有利位置(其吸引势能 $U_{goal}(x)$ 较大, 排斥势能 $U_{obs}(x)$ 较小), 而且需要考虑该接球队员与持球队员之间的距离。距离越大, 则传接球越容易失误或者越容易被对方破坏; 距离越小, 则传接球成功的可能性越大。因此, 对于持球

队员视野范围内的 k 个队友,按照下式计算持球队员 c 向队友 m 传球成功的概率

$$p_m^c = \frac{[\tau_m]^\phi [\eta_m]^\varphi}{\sum_{m \in \text{partner}_c} [\tau_m]^\phi [\eta_m]^\varphi}, m \in \text{partner}_c. \quad (12)$$

式中, partner_c 表示持球队员 c 视野范围内与其连线上无对方防守球员的 k 个队友集合。 η_m 按照下式计算

$$\eta_m = \frac{1}{d_{cm}} \quad (13)$$

式中, d_{cm} 表示持球队员 c 和队友 m 之间的距离。式(10)式(11)和式(13)中的距离 $(x - x_{\text{goal}})$, ρ , d_{cm} 的具体计算参见文献[18]。

根据式(12)可以计算出持球队员视野范围内多名队友传球的可能值,然后找出最佳接球队友 i ,即

$$p_i^c = \max \{ p_m^c \mid m \in \text{partner}_c \} \quad i \in \text{partner}_c$$

然后分情况判断:

1) 如果 partner_c 不为空且 $p_i^c \geq p_0$, 其中 p_0 是可以调节的一个概率阈值,则持球队员 c 选择向队友 i 传球。

2) 如果 $p_i^c < p_0$ 或者 partner_c 为空,则持球队员 c 找不到满足接球条件的队友,持球队员 c 将选择运球。

行为预测主要通过文中式(9)—(13)来判定。以计算场上每名球员的人工势场能 U_{APF} 为原操作进行分析,算法的复杂度主要取决于持球队员视野内队友的个数。因此算法复杂度为线性阶,可表示为 $O(n)$, 其中 n 为持球队员视野内队友的个数。

3 实验结果与分析

为了验证持球队员行为预测算法的有效性,在 CPU 主频为 2.8 GHz,内存为 3 GB 的微机,采用 C++ 和 MATLAB 进行混合编程。首先对持球队员的头态识别算法进行测试,在此基础上,进一步验证持球队员行为预测算法的有效性。

3.1 头态识别实验

为了训练 1.2 小节中介绍的多类 LogitBoost 分类器,手工从 100 段篮球比赛视频中截取 800 幅头部图像,每类各 100 幅。每类的 100 幅头部图像尺寸介于 10×10 到 30×30 之间,包括了各种不同的人物外貌特征,如秃顶、胡须、深色与浅色皮肤以及

各类光照的变化。截取后,将其统一归一化到 20×20 的尺寸(由于协方差描述子具有“尺度不变性”^[13-14],对测试图片归一化不是必要的,在这里归一化只是为了使训练集图片有一个统一的大小),作为训练多类 LogitBoost 分类器的样本,典型训练样本如图 4 所示。另外,还截取了 100 幅非头部图像作为反例训练集,主要包括局部球场背景图块和球员其他身体部位的图块。图 5 显示了在训练样本上抽取的正方形区域相对尺寸(与样本边长的比值)与识别准确率的关系。可以看出,当其边长与样本边长的比为 0.28 时,准确率最高。训练过程中,设定头部图像检测分类器的 τ_{accuracy} 为 0.99,学习率 τ_{rate} 为 0.1,非头部图像检测分类器的 τ_{accuracy} 为 0.4。我们选取了另外 400 幅头部图片作为测试集,每类各 50 幅,实验结果如表 2 所示。

由表 2 可知,误判主要出现在相邻两种头部姿态的判别上。这主要是因为篮球比赛视频中头部图像分辨率低,存在运动模糊等各种噪声干扰,以及相邻的头部姿态本身较为相似。从实验结果可以看出,融合了多种底层特征的协方差描述子能较好地适应上述情况,使用本文的方法取得了理想的实验结果,8 种头态的平均识别率达到了 96.5%。

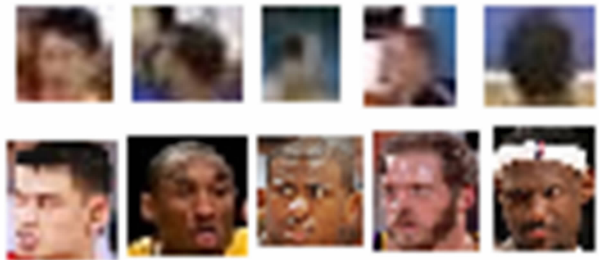


图 4 头态分类器的典型训练样本
Fig. 4 Typical training examples of the head pose classifier

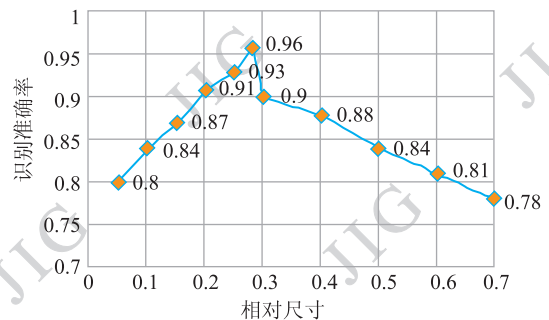


图 5 头态识别准确率与小区域相对尺寸关系曲线
Fig. 5 The relation curve between head pose recognition accuracy and relative patch dimension of extracted square region

表 2 头态识别实验结果

Table 2 Result of head pose recognition experiment

度数	0	45	90	135	180	225	270	315
0	0.98	0.01	0	0	0	0	0	0.01
45	0.01	0.96	0.02	0	0	0	0	0.01
90	0	0.02	0.95	0.01	0	0.01	0	0.01
135	0	0.01	0.02	0.95	0	0.01	0	0.01
180	0	0	0	0	0.99	0.01	0	0
225	0	0	0	0.01	0	0.97	0.02	0
270	0	0.01	0	0	0	0.01	0.96	0.02
315	0.02	0	0	0	0	0.01	0.01	0.96

3.2 持球队员行为预测实验

我们选取了 10 场帧率为 25 帧/s 的篮球比赛视频,视频中参与比赛的球队整体水平大致接近。对于篮球视频中的镜头切换,我们以人工的方式将视频切割成一个个镜头,分别对每个镜头进行分析。其中每个镜头对应于一类球员动作(投篮,传球,运球),具体实验数据见表 3。实验中,我们将各类镜头片段总数的 2/3 作为训练视频,用于估计行为预测算法中的各项参数,总数的 1/3 作为测试视频。

表 3 行为预测实验数据描述

Table 3 Data description of behavior prediction experiment

场次	投篮镜头数	传球镜头数	运球镜头数
1	41	49	28
2	45	52	33
3	38	46	39
4	47	50	36
5	37	52	31
6	48	45	40
7	44	43	41
8	53	58	30
9	51	62	29
10	46	58	37

对于投篮信息量 τ_0 ,当取值较大时,可能将投篮错误地预测为传球;当取值较小时,可能将传球错误地预测为投篮,从而导致投篮预测准确率下降。同理,对于运球信息量 p_0 ,若取值较大,可能将传球错误地预测为运球;若取值较小,可能将运球错误地预测为传球,导致预测准确率下降。通过大量实验我们发现,分别以训练视频中投篮球员 APF 信息量的均值与选择运球球员 APF 信息量的均值作为参

数 τ_0 和 p_0 的值,动作预测准确率最高。对于距离阈值 ρ_0 ,统计每次成功抢球的抢球队员与持球队员的距离,选取其中的最大值赋给距离阈值 ρ_0 。投篮能力 k_p 与抢球能力 k_d 的初始值均设为 1。为了确定式(12)中的参数 ϕ 和 φ ,令 $ratio = \phi/\varphi$,通过比较不同比值下 $ratio$ 与预测正确率的关系来确定参数 ϕ 和 φ 。实验结果如图 6 所示,预测正确率为

$$R_p = \frac{\sum_{i=1}^{10} (\tau_{shooting}^i + \tau_{passing}^i + \tau_{dribbling}^i)}{10}$$

式中, $\tau_{shooting}^i, \tau_{passing}^i, \tau_{dribbling}^i$ 分别代表第 i 场比赛投篮、传球、运球预测的正确率。实验中,分别设置 ϕ 和 φ 为 2.6 和 1。抢球能力上限值 θ_d 和投篮能力上限值 θ_a 通过实验分别设定为 5 和 25。10 次实验的统计结果见表 4。

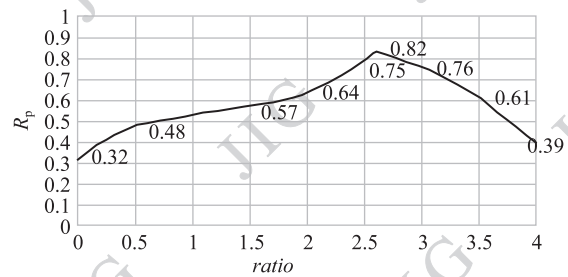


图 6 行为预测正确率与 ratio 关系曲线

Fig. 6 Relation curve between behavior prediction accuracy and the defined ratio

由表 4 可知,投篮与运球的平均预测正确率较高,均在 80% 以上,而传球的正确率较低。这主要是因为我们在计算传球预测正确率时,不仅要考虑算法能预测为传球,还要求其能准确预测接球的队友。分析实验结果,出现误判的原因主要有:1) 篮球比赛中,持球队员的行为选择由某些特定的非偶然因素决定外,还受某些偶然因素影响(如体力因素,是否有伤),同时具有随机性;2) 受到某些战术打法的制约和队友间的配合默契度等影响。由于上述原因的存在,在一定程度上造成了误判的出现。

4 结论

体育视频分析是当前计算机视觉领域研究的热点问题,而目前的研究大都集中于对视频中的运动员进行检测和跟踪,提取他们的动作轨迹来进行分析,属于底层视频分析技术。提出一种新颖的对篮

表4 行为预测实验结果

Table 4 Result of behavior recognition experiment

场次	预测动作		
	投篮	传球	运球
1	89	78	83
2	82	72	79
3	85	74	81
4	82	73	77
5	87	77	82
6	91	81	85
7	83	73	80
8	84	72	81
9	85	75	82
10	90	80	84
平均正确率	85.8	75.5	81.4

球比赛中持球队员行为进行预测的技术。针对篮球比赛视频中背景复杂,球员移动速度快,头部图像分辨率低的特点,提出运用协方差描述子对持球队员的多种有效头部图像特征融合,构成黎曼流形,再将其映射到切空间中,用训练好的多类 LogitBoost 进行持球队员的头部姿态识别,确定持球队员的视野范围。再根据持球队员视野范围内的双方球员分布,利用基于人工势能场的球员信息量,完成对其投篮、传球、运球3种行为的预测,并通过两个实验验证了这项技术的有效性。

篮球比赛中持球队员的行为选择受到诸多因素的影响,是一个较为复杂的决策过程,如何通过挖掘视频图像中的更多有效信息,提高行为预测的准确率是值得进一步研究的问题。

参考文献 (References)

- [1] Lashkia G, Ochimachi N, Nishida E, et al. A team play analysis support system for soccer games [C]// Proceedings of the 16th International Conference on Vision Interface. Halifax, Canada: IAPR, 2003:536-541.
- [2] Lu W L, Okuma K, Little J J. Tracking and recognizing actions of multiple hockey players using the boosted particle filter [J]. Image and Vision Computing, 2009, 27(1):189-205.
- [3] Li Fahong, Woodham R J. Video analysis of hockey play in selected game situations [J]. Image and Vision Computing, 2009, 27(1): 45-58.
- [4] Perse M, Kristan M, Pers J, et al. Analysis of multi-agent activity using petri nets [J]. Pattern Recognition, 2010, 43(4):1491-1501.
- [5] Ramasso E, Panagiotakis C, Rombaut M, et al. Human shape-motion analysis in athletics videos for coarse to fine action-activity recognition using transferable belief model [J]. Electronic Letters on Computer Vision and Image Analysis, 2009, 7(4):32-50.
- [6] Roh M C, Christmas B, Kittler J, et al. Gesture spotting for low-resolution sports video annotation [J]. Pattern Recognition, 2008, 41(3):1124-1137.
- [7] Min J, Kasturi R, Camps O. Extraction and temporal segmentation of multiple motion trajectories in human motion [J]. Image and Vision Computing, 2008, 26(12):1621-163.
- [8] Tong X F, Duan L Y, Xu C S. Local motion analysis and its application in video based swimming style recognition [C]// Proceedings of the 18th International Conference on Pattern Recognition. Hong Kong, China: Hong Kong Baptist University, 2006:32-50.
- [9] Zhu G Y, Xu C S, Huang Q M, et al. Event tactic analysis based on broadcast sports video [J]. IEEE Transactions on Multimedia, 2009, 11(1):49-67.
- [10] Duque D, Santos H, Cortez P. N-ary trees classifier [C]// Proceedings of the 3rd International Conference on Informatics in Control, Automation and Robotics. Setubal, Portugal: INSTICC, 2006: 457-460.
- [11] Duque D, Santos H, Cortez P. Prediction of abnormal behaviors for Intelligent video surveillance systems [C] // Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining. Honolulu, Hawaii, USA: IEEE, 2007:362-367.
- [12] Xia L M, Zhang L C. Object tracking based on adaptive particle filter [J]. Journal of Image and Graphics, 2009, 14(1):112-117. [夏利民, 张良春. 基于自适应粒子滤波器的物体跟踪 [J]. 中国图象图形学报, 2009, 14(1):112-117.]
- [13] Tuzel O, Porikli F, Meer, P. Pedestrian detection via classification on Riemannian manifolds [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(10):1713-1727.
- [14] Tuzel O, Porikli F, Meer E. Region Covariance: a fast detection and classification [C]// Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Graz University of Technology, 2006:589-600.
- [15] Fletcher P T, Joshi S. Riemannian geometry for the statistical analysis of diffusion tensor data [J]. Signal Processing, 2007, 87(2):250-262.
- [16] Pennec X, Fillard P, Ayache N. A Riemannian framework for tensor computing [J]. International Journal of Computer Vision, 2006, 66(1):41-66.
- [17] Khatib O. Real-time obstacle avoidance for manipulators and mobile robots [J]. The International Journal of Robotics Research, 1986, 5(1): 90-98.
- [18] Farin D, Krabbe S, Effelsberg W, et al. Robust camera calibration for sport videos using court models [C]// Proceedings of the International Society for Optical Engineering. San Jose, California, USA: SPIE 2004:80-91.