

国家自然科学基金项目《基于多特征融合的视频足球比赛中的团队行为识别方法研究》(项目编号: 61462008)资助
广西自然科学基金项目《团体行为识别的关键技术研究》(项目编号: 2013GXNSFAA019336)资助
广西科技大学博士基金项目《视频图像处理在信息隐藏方面的应用研究》(项目编号: 院科博12Z14)资助

JIYU TEZHENG RONGHE DE
ZUQIU BISAI SHIPIN ZHONG DUO YUNDONGYUAN DE
XINGWEI SHIBIE FANGFA YANJIU



基于特征融合的 足球比赛视频中多运动员的 行为识别方法研究

王智文 ● 著



西南交通大学出版社

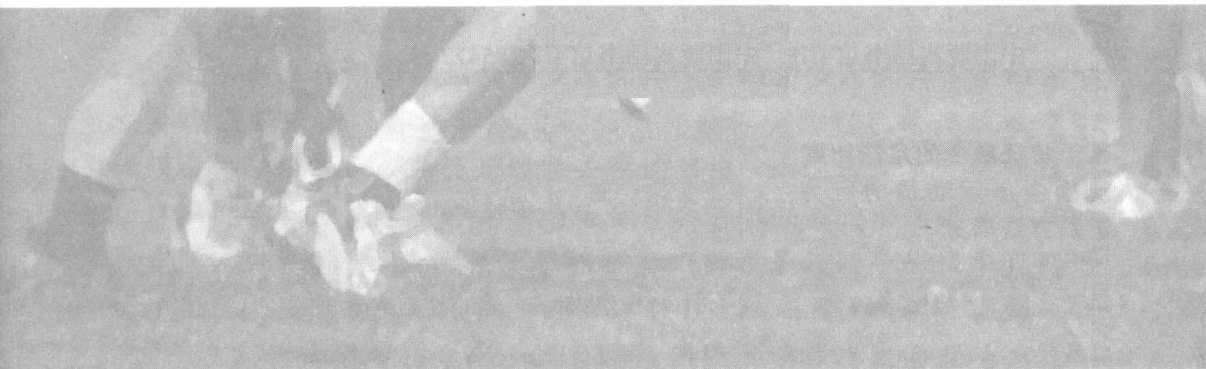
国家自然科学基金项目《基于多特征融合的视频足球比赛中的团队行为识别方法研究》(项目编号: 61462008)资助
广西自然科学基金项目《团体行为识别的关键技术研究》(项目编号: 2013GXNSFAA019336)资助
广西科技大学博士基金项目《视频图像处理在信息隐藏方面的应用研究》(项目编号: 院科博12Z14)资助

JIYU TEZHENG RONGHE DE
ZUQIU BISAI SHIPIN ZHONG DUO YUNDONGYUAN DE
XINGWEI SHIBIE FANGFA YANJIU



基于特征融合的 足球比赛视频中多运动员的 行为识别方法研究

王智文◎著



西南交通大学出版社
· 成 都 ·

内容提要

全书共分6章。其中,第1章绪论介绍了多运动员行为识别研究的背景及意义。第2章首先介绍了6种具有代表意义的典型的去噪算法,并比较了它们的性能;其次介绍了常用的足球比赛视频图像增强及本书提出的基于正交小波分析和伪彩色处理的足球比赛视频图像增强算法,并对它们的性能进行了对比分析。第3章介绍了足球比赛视频中多运动员行为的特征提取及融合的具体实现方法。第4章介绍了足球比赛视频中多运动员的行为识别方法。第5章介绍了迁移学习的定义、类型、实现途径及其在多运动员行为识别过程中的应用。第6章对研究工作进行了总结和展望。

本书着眼于足球比赛视频中多运动员行为识别方法的研究,涉及的研究内容可以给图像处理、特征提取、行为表示及识别等研究领域的相关研究人员提供技术参考,也可作为高校研究生、本(专)科生行为识别教学方面的教材,还可作为有关技术人员的参考书籍。

图书在版编目(CIP)数据

基于特征融合的足球比赛视频中多运动员的行为识别方法研究 / 王智文著. —成都:西南交通大学出版社, 2015.4

ISBN 978-7-5643-3589-2

I. ①基… II. ①王… III. ①视频设备—应用—足球运动—运动员—行为分析—方法研究 IV. ①G843

中国版本图书馆CIP数据核字(2014)第285477号

基于特征融合的足球比赛
视频中多运动员的行为
识别方法研究

王智文 著

责任编辑 张宝华
装帧设计 何东琳设计工作室

印张 10 字数 169千

成品尺寸 170 mm×230 mm

版本 2015年4月第1版

印次 2015年4月第1次

印刷 四川煤田地质制图印刷厂

书号: ISBN 978-7-5643-3589-2

出版发行 西南交通大学出版社

网址 <http://www.xnjdcbs.com>

地址 四川省成都市金牛区交大路146号

邮政编码 610031

发行部电话 028-87600564 028-87600533

定价: 36.00元

图书如有印装质量问题 本社负责退换

版权所有 盗版必究 举报电话: 028-87600562

前 言

足球比赛视频中的运动员行为是一种有计划、高协同性的多运动员团队行为。团队行为的理解和识别是计算机视觉研究领域的重要研究问题之一,有许多方面的应用,如视频监控、对象视频摘要、人机交互、体育视频分析、运动员辅助训练、比赛辅助判罚和视频检索浏览等。因此,开展足球比赛视频中的多运动员行为识别研究具有极大的经济价值和社会价值。

足球比赛视频中的行为识别涉及特征提取、目标跟踪与检测、行为表示、分类器的构建及行为识别等具体研究内容,需要使用运动目标的检测、分割、跟踪识别、语义表示与推理等技术,涉及模式识别、图像处理、图形学、计算机视觉、机器学习和人工智能等学科,是一个具有挑战性的、跨学科的研究方向。

本书在全面综述足球比赛视频中的多运动员行为识别研究现状的基础上,提出用模糊推理系统来融合多特征进行足球比赛视频中的行为识别,并针对当前多运动员行为识别的研究过程中存在的不足展开研究。主要研究成果归纳如下:

(1) 本书在视频图像预处理过程中提出了基于多元统计模型的分形小波自适应图像去噪算法和基于正交小波分析和伪彩色处理的足球比赛视频图像增强算法:在视频图像去噪过程中,通过结合多元统计模型与分形小波去噪方法,能够更准确地估计各种相关信息,选择高品质的图像空间。在适度的噪声方差下根据拼贴距离在最好的子树域中找到近优父子树,从而预测出无噪声的图像分形小波编码,达到优化去噪的目的。基于多元统计模型的分形小波自适应图像去噪算法在去除噪声的同时,能有效地保持图像的边缘及纹理特征,很好地保留图像的精细结构,取得良好的去噪效果。由于采用了预测小波分形编码,优化了算法结构,算法的处理速度比较快。基于正交小波分析和伪彩色处理的足球比赛视频图像增强算法既可克服采用正交小波分析足球比赛视频图像增强算法处理后的图像偏亮及对比度较差等缺陷,又可克服伪彩色处理的足球比赛视频图像增强算法不能够充分处理图像中某些细节信息的缺陷。

(2) 提出用自动生成 RBF 网络来融合多特征: 为了克服光照、遮挡、尺度变化等影响, 满足实时性识别的要求, 提出用自动生成 RBF 网络融合足球比赛视频中的多运动员行为识别过程中提取的球员的服装颜色矩特征、球员和裁判的轮廓特征、球场线的坐标参数特征及运动目标的运动轨迹特征。定义一个动态特征模型, 首先提取有关足球比赛视频中的多运动员行为的主要特征, 当这些特征不足以完成识别与理解时, 系统逐步提取候选的细节特征。同时通过采用 3D 局部方向直方图特征, 能有效解决遮挡和姿态变化的多样性, 使足球比赛视频中的多运动员行为识别与理解具有更强的鲁棒性。

(3) 提出了足球比赛视频中的团体行为模式的时空驱动力模型和模糊推理系统: 利用由时空限制的轨迹、变化的人数和行为之间的时空变化的集合组成的行为特征为足球比赛视频中的团体行为进行建模。将足球比赛视频中的团体行为作为一个区域密集分布的时空驱动力的动态过程, 用简单的离散轨迹点集来取代运动的发生。通过将 $F(t_n, x, y)$ 的 Lie 群非线性流形空间转化为 $f(t_n, x, y)$ 的 Lie 代数的线性空间来大大减少模型的计算量。模型充分利用了从运动轨迹中获得的位置和速度等低层次特征, 使得模型的学习简单, 模型的特征融合能力比较强, 与其他模型相比性能优越。模型在建模复杂行为模式方面具有通用性和灵活性。

提出了足球比赛视频中的多运动员行为识别的模糊推理系统, 将行为模型抽象为事件模型, 建立传球、射门、控球、带球、丢球、进球、角球、任意球、越位、球出界、红黄牌等事件的推理规则, 系统应用这些推理规则进行足球比赛视频中的多运动员行为识别。

(4) 设计了尺度自适应局部时空特征 Harris 检测操作数来解决复杂背景中的光照、多尺度和遮挡问题: 依据视觉理解的整体性和层次性原理, 将空间金字塔模型推广并应用到局部时空特征中, 设计了尺度自适应选择局部时空特征 Harris 检测操作数。该操作数方法简单, 计算速度快, 能解决复杂背景中的光照变化和多尺度问题, 并能在一定程度上解决遮挡问题。

(5) 首次将迁移学习算法引入足球比赛视频中的多运动员行为识别与理解中解决多视角及遮挡问题: 借鉴迁移学习在图像分类、手势识别等领域研究的成功经验, 设计了基于迁移学习的局部时空码本原型构建

算法,该算法使得不同视角的码本之间能够共享特征,以更紧凑的方式来表示足球比赛视频中的多运动员行为,能在一定程度上解决多视角问题,提高足球比赛视频中的多运动员行为识别与理解方法的鲁棒性。

总之,本书从特征提取、团队行为表示、团队行为建模和分类器的构建技术等核心问题入手来研究足球比赛视频中的多运动员的行为识别。首先提出了用基于多元统计模型的分形小波自适应图像去噪算法和基于正交小波分析和伪彩色处理的足球比赛视频图像增强算法来改善视频图像的视觉效果,然后提出了用自动生成 RBF 网络来融合提取的多特征,提出了足球比赛视频中的团体行为模式的时空驱动力模型,并用它来建模足球比赛视频中的团体行为,更具通用性和灵活性。设计了尺度自适应局部时空特征 Harris 检测操作数来解决复杂背景中的光照、多尺度和遮挡问题,并能在一定程度上解决遮挡问题。提出了基于先验知识和人工神经网络的树结构混合分类器,提高了识别的准确率及识别速度,利用神经网络的独立性以及自适应性解决了单一分类器难以不断学习和适应环境、光照、运动员人数变化的多运动员行为识别问题。最后针对多视角及遮挡问题首次将迁移学习算法引入足球比赛视频中的多运动员行为识别与理解中解决部分遮挡问题和多视角问题。本书提出的方法提高了足球比赛视频中的多运动员的行为识别的识别性能,有利于促进足球比赛视频中的多运动员的行为识别技术不断向前发展和进一步的实用化。

各项实验结果表明,本书提出的解决方法是有效的,基本达到了预期的研究目标,为其他相关的应用研究提供了借鉴。

本书由王智文著。在著作过程中得到了广西科技大学计算机学院、西南交通大学出版社的大力支持,在此表示衷心的感谢!另外,还参考了大量的国内外著作、学术论文和有关文献,在此谨向这些文献的作者表示深深的谢意!

由于作者水平有限,书中难免存在不当和疏漏之处,热诚欢迎各位专家及读者批评指正。

作 者

2014 年 9 月

目 录

1	绪 论	1
1.1	研究背景与意义	1
1.2	足球比赛视频中多运动员行为识别的研究现状及存在的问题	3
1.3	研究内容及创新点	13
1.4	本书的组织结构	15
2	足球比赛视频图像预处理技术	17
2.1	足球比赛视频图像去噪	17
2.2	足球比赛视频图像增强	34
2.3	小 结	49
3	足球比赛视频中多运动员行为的特征提取及融合	50
3.1	运动目标跟踪与检测	50
3.2	足球比赛视频中多运动员的行为特征提取	61
3.3	足球比赛视频中多运动员行为的特征融合	71
3.4	小 结	75
4	足球比赛视频中多运动员的行为识别	76
4.1	足球比赛视频中的多运动员行为表示	76
4.2	足球比赛视频中团体行为模式的时空驱动力模型	84
4.3	基于先验知识和人工神经网络的树结构混合分类器	92
4.4	足球比赛视频中多运动员的行为识别	103
4.5	小 结	116

5 基于迁移学习的多运动员的行为识别	118
5.1 迁移学习	118
5.2 迁移学习在多运动员行为识别过程中的应用	120
5.3 实验及结果分析	124
5.4 小 结	130
6 总结与展望	131
6.1 研究工作总结	131
6.2 未来研究工作展望	133
参考文献	136

绪 论

足球比赛视频中的行为识别是人工智能的前沿研究方向，这方面的研究具有重大的经济意义和社会价值。本章介绍了足球比赛视频中运动员行为识别的研究背景与意义、研究现状及存在的问题、本书的研究内容及创新点。

1.1 研究背景与意义

行为识别的目标是现实生活中建立起来的人类共同行为。精确的行为识别具有挑战性，因为人类行为具有复杂性、高度多样化等特征。足球比赛视频中的运动员行为是一种有计划、高协同性的多运动员（智能体）的团队行为。团队行为的理解和识别是计算机视觉研究领域的重要研究方向之一，它有许多方面的应用，如视频监控、对象视频摘要、人机交互、体育视频分析、运动员辅助训练、比赛辅助判罚和视频检索浏览等，可以获得极大的经济价值和社会价值。虽然单一智能体的行为识别问题已得到广泛研究，但是多智能体的行为识别的研究却相对少得多，况且目前还不能让计算机对足球比赛视频中的团队行为的理解、描述和识别达到具有生物那样高效、灵活的功能，对这一工作缺乏系统和深入研究，因此，足球比赛视频中的行为识别与理解成为了当前计算机视觉研究领域中的前沿方向。足球比赛视频中的行为识别易受光照、噪声、视角、多尺度以及遮挡等因素的影响，且要在复杂场景下实时鲁棒地识别团队行为，因此，足球

比赛视频中的团队行为的理解和识别仍然是一个具有挑战性的研究课题。原因如下:

(1) 很难找到球队行为的有效描述, 因为球队行为是由多人协同完成的行为, 球队中个体的单独行为和球队整体行为具有一定的关联性。一般不能通过观察单一智能体的单独行为来特征化球队的总行为。

(2) 当用局部特征来描绘单个人的行为时, 需要对很多对象进行跟踪。此外, 球队行为的环境、背景往往非常混乱, 存在自遮挡和互遮挡现象。

(3) 球队行为中的人数、相互遮挡和自我遮挡、不规则的摄像机参数等干扰因素也给球队行为的理解和识别带来很多困难。

因此, 开展此项研究, 具有重大的理论意义和广泛的应用前景及潜在的经济价值。

足球比赛视频中的多运动员的行为识别方法的研究旨在对图像或者视频进行分析, 从而获得足球比赛视频中的行为姿态和运动参数, 并进一步进行语义分析及行为识别与理解。足球比赛视频中的行为识别覆盖了特征提取、目标跟踪与检测、行为表示、分类器的构建及行为识别等具体研究内容, 需要使用运动目标的检测、分割、跟踪识别、语义表示与推理等技术, 涉及模式识别、图像处理、图形学、计算机视觉、机器学习和人工智能等学科, 是一个具有挑战性、跨学科的研究方向。

总之, 足球比赛视频中的多运动员的行为识别是使计算机更加人性化(从“looking at people”向“understanding people”转变)的关键所在。但是, 目前足球比赛中的行为识别和理解的研究还没有找到切实可行的办法, 因为足球比赛场景中多运动目标、遮挡、多视角、多尺度及行为变化多样性的问题是不可避免的。虽然有的研究人员提出通过 3D 来解决遮挡问题, 但是这种方案需要多摄像机的协同工作, 成本高, 而且无法达到实时效果, 不能满足足球比赛转播的实时性识别与理解的要求。因此有必要研究一种能在真实复杂背景中工作的实时的、鲁棒的和低成本的足球比赛视频中的行为识别和理解方法, 即考虑场景中存在各种各样的变化和干扰因素, 如 ① 多视角; ② 遮挡与无遮挡的; ③ 多运动目标; ④ 光照变化; ⑤ 噪声干扰; ⑥ 有球和无球场景, 如图 1.1 所示。

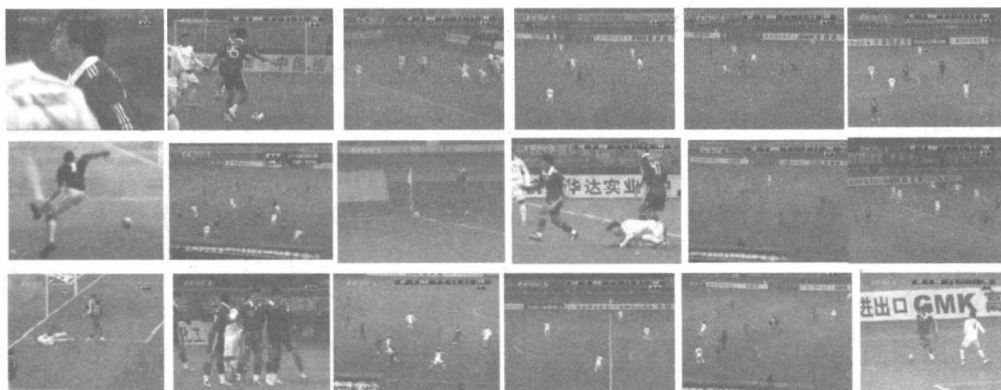


图 1.1 足球比赛各种各样场景变化和干扰因素图

从模式识别的角度看,要构建一个包含如此多的类内和类间变化的、鲁棒的分类器,不管是特征的提取和选择,还是分类器的构造,都是一个难题。由此可见,足球比赛视频中的行为识别和理解极具挑战性,同时又为众多行为识别的实际应用提供核心技术。开展足球比赛视频中的行为识别和理解的研究,具有重大的学术价值,其相关的研究成果可以被人脸识别、步态识别等生物特征识别方法研究及数字媒体内容理解的理论与方法研究所借鉴;同时该方面的研究具有广泛的应用前景,可以创造巨大的社会价值和经济效益。

1.2 足球比赛视频中多运动员行为识别的研究现状及存在的问题

国内外有大量的研究机构和研究者从事人类行为识别方面的研究。其中法国 INRIA 实验室的 Bill Triggs 研究组针对单面摄像机的人体运动估计做了大量研究^[1-5],对人的运动进行检测和运动分析,并对人体进行鲁棒描述,通过机器学习方法回归运动数据与形状参数^[6-8]。另外,还有基于局部特征的目标识别研究^[9-14],基于局部时空兴趣点的人体行为识别研究等^[15-17]。瑞士 EPFL 实验室的 Pascal Fua 研究组利用立体视觉视频序列进行人体建模^[18, 19],用视频图像中的轮廓来估计运动参数;Urtasun 等在人体运动跟踪方面首次将高斯过程隐变量模型和高斯过程动态模型应用到人体跟踪^[20]。美国 Brown 大学计算机系的 Michael J Black 研究组基于模型的

人体运动分析研究,通过单目或多目视频图像进行人体区域特征提取,将得到的特征与模型的投影进行相似性比较,然后基于退火粒子滤波或者图模型的方法进行跟踪^[21-24]。加拿大多伦多大学计算机系的 C. Sminchisescu 研究组开展了人体运动分析方面的研究^[25-32]。意大利国家研究委员会的 T. D'Orazio 等从事特征提取、强化学习、自主机器人、行为协调等领域的研究^[33, 34]。日本大阪大学的 T. Shimawaki, T. Sakiyama 和丰桥大学智能系统实验室的 J. Miura 等研究严重遮挡情况下球的路由估计^[35]。国外的 SoccerMan 是一个基于视频的足球比赛的三维重建系统^[36]。它首先提取每帧的摄像机运动参数,求出图像坐标系和空间坐标系之间的转换关系,构成三维重建系统;然后提取出球场背景纹理、球的运动轨迹和球员头部的运动轨迹、球员的纹理等;最后通过三维重建系统重建上述元素,从而达到重建比赛的目的。

国内体育视频的语义分析起源于 20 世纪 90 年代末,其中中科院自动化所在步态识别视觉监控方面取得了较好的研究成果,北京大学视觉与听觉信息处理国家重点实验室、清华大学、上海交通大学、西安交通大学、南京大学、哈尔滨工业大学、南京理工大学等也在该领域做了许多研究工作。代表机构和人员包括谭铁牛教授负责的 NLPR 实验室、高文教授负责的 JDL 实验室和清华大学徐光佑教授负责的研究组、西安交大利物浦大学徐明等。

足球比赛视频中的行为识别相关研究起源于 20 世纪末,图 1.2 总结了

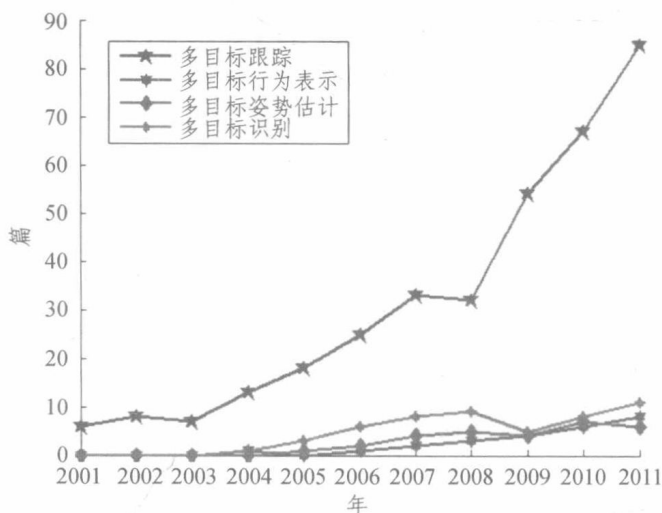


图 1.2 多目标行为识别相关研究论文统计图

从 2001 年起到 2011 年 12 月止的一些重要期刊和会议上发表的有关多运动员行为跟踪与检测、行为表示、姿势估计及行为识别方面的论文数目。图 1.2 从侧面反映了多运动员行为识别相关研究发展过程。从图 1.2 可以看出,除了多运动员行为跟踪与检测方面的研究相对多些外,多运动员行为识别的其他方面研究还存在一定的难度,值得进行更深入、更广泛的研究。

足球比赛视频中的多运动员行为识别涉及特征提取、目标跟踪与检测、行为表示、分类器的构建及行为识别等具体研究内容。

1.2.1 目标跟踪与检测技术的研究

足球比赛视频中的多运动员行为识别过程中的目标跟踪与检测主要是对运动员及球的跟踪与检测,属于多目标跟踪与检测范畴。

1.2.1.1 多运动员跟踪与检测技术的研究

过去 20 多年,多目标跟踪与检测问题得到了广泛研究^[37],提出了许多目标检测与跟踪算法,但是对学习和识别框架中的整个球队的行为模式的跟踪与检测方面的研究很少。这些算法可以概括为四大类:① 基于特征的;② 基于模型的;③ 基于运动的;④ 基于数据关联的。表 1.1 总结了从 2000 年 1 月起到 2011 年 12 月止的一些重要期刊和会议上发表的有关行为跟踪和识别使用特征统计。

表 1.1 人体行为识别使用特征统计分析 (%)

年 份	外观形状特征		运动特征	混合特征	数据关联	时空特征
	外观	模型				
2010—2011	22.5	10.2	29.4	5.6	12.1	15.8
2009—2010	26.3	13.9	28.5	4.2	10.4	16.7
2008—2009	31.2	15.9	19.1	1.9	11.5	20.4
2006—2007	37.8	6.3	31.5	3.7	9.9	10.8
2004—2005	40	6	40	5	13	9
2003—2004	42.3	8	34.5	6.9	14.1	7
2002—2003	38	9	42	5	8	6
2001—2002	39.7	16.2	28.8	4.6	7.1	5
2000—2001	43	7	36	10	5	4

在基于特征的算法中,跟踪目标的某些特征被用来区分一帧视频中的跟踪目标和其他物体。有些算法利用了背景图像作参考,即所谓的背景帧:利用当前帧减去背景帧所得到的“差额帧”中的所有对象就是计算出的跟踪目标^[38]。为了从其他物体中鉴别跟踪目标,用跟踪目标的特征来特征化特征状态空间中的跟踪目标。跟踪目标表示中的参数化形状^[33, 34]、颜色分布^[39]、形状和颜色^[40]可以作为特征。文献[34]用特征与手动标记的跟踪目标来训练神经网络分类器,然后用训练好的神经网络分类器来区别跟踪目标与其他对象。文献[41]用几何模块实现图像到模型的单应矩阵估计,通过提取描述跟踪目标的位置和局部外观的兴趣点来进行局部跟踪。椭圆形区域内的颜色直方图用于跟踪球场上的运动员^[42]。这些算法更多地利用了低层图像信息,获取特征方式简单,用一种粗糙的特征来描述整个行为,对于噪声、视角变化和行为的主体变化很敏感。

基于模型的算法,包括反模型算法,使用特征、高层语义表示和领域知识来区分跟踪目标与其他对象^[43-47]。Gammeter 等用人类姿势统计模型来完善行人跟踪系统^[48]。Ali 和 Shah 在密集人群中用基于力学模型的现场结构来跟踪个体目标^[49]。文献[50]用激光范围搜索器和动态 RFID 传感器的组合来解决跟踪与识别问题,用概率模型进行实时跟踪。这些算法的缺陷是难以建立精确的行为表示模型,各种不同的比赛不能共享模型,且受限于姿势估计算法的发展。

基于特征与模型的算法主要有三要素:目标表示、特征提取和对象区分。建立目标表示的原则是从其他物体中区分目标并可以很容易地提取用于表示的特征。因此,目标表示可以包括外观特征、运动特征,而模型是用来解决不同问题的。初始化过程中建立的表示随着帧的变化而不断更新。算法利用了隐含的假设,即在同一帧内目标与其他物体之间有“某种不同”。算法的缺陷是特征的选取及特征对行为识别的影响难以估计,模型的适应性需要提高。

基于运动的算法依赖于提取随帧(或时间)变化的运动的一致性来分割运动对象^[45, 51]。文献[52]通过消除阴影为足球运动员检测取得良好的分割效果。虽然运动的一致性涉及一些帧,基于运动的算法通过水平目标而不是水平轨迹来区分目标和其他物体。这些方法的缺陷是很难找到遮挡目标的位置。

基于数据关联的算法的目的是解决数据关联问题,是一个寻找检测对象与已知路径之间的正确对应问题^[51, 53-56]。数据关联问题有四种基本解决

方法: ① 近邻法是一种在高度混乱的环境下能有效计算但跟踪不可靠的算法^[53]。② 文献[53, 54, 55]中的跟踪操作是解决数据关联问题的一种技术。当前跟踪操作包括由 Smith 和 Buechler 提出的轨迹分裂^[22]、轨迹合并和轨迹删除^[53, 54]。③ 联合概率数据关联, 即利用联合概率目标测量执行“排除”原则以防止两个或更多的跟踪器锁定相同目标^[40, 54]。④ 多假设跟踪^[53-55]是一种基于多种扫描方法的跟踪技术。这些算法需要较高的内存和过多的计算, 从而使问题的复杂度指数增长^[53]。

1.2.1.2 足球跟踪与检测技术的研究

足球的跟踪是比球员跟踪更难的问题, 因为图像中足球的尺寸小、方向变化突然。相对于一般的目标检测与跟踪, 有多种专门用于足球定位与检测的算法。文献[57, 58]报道了成功检测和跟踪固定摄像机记录的足球比赛视频中的球, 但这些方法运用到 BSV 时, 效果都欠佳。Gong 等^[59]设计了检测 BSV 中的球的算法, 该算法只使用球的形状和颜色属性来区分球与其他物体, 但没有提供详细的实验性能分析。

Seo 等人提出了一种跟踪 BSV 中的球的算法^[39], 利用卡尔曼滤波的模板匹配程序来跟踪球并使用反投影法来预测可能发生的遮挡。然而, 该算法需要手动输入球的出发位置且没有跟踪结果报告。

D'Orazio 等使用改进的圆形霍夫变换 (Circle Hough Transform, CHT) 与神经网络分类器来从实时视频选定的帧中检测足球^[33, 34]。当球在视频中足够大、颜色单一、变形不严重时, 这种方法可以取得良好的检测效果。此外, 改进的圆形霍夫变换与神经网络分类器无法识别一些似球的非球。由于球不够大、颜色是黑白相间的、没有明显的纹理, 以前检测球的方法并不适用。一些研究提出了基于球的路径分析的足球检测方法。文献[60]提出用基于轨迹的检测和跟踪算法来定位足球视频直播中的球。文献[61]提出了对固定摄像机拍摄的图像序列中的模糊小球进行跟踪的方法。足球跟踪的过程是对每 m 帧进行批处理, 以产生仅包含球区的像素的积累图像。对积累图像用粒子过滤器将球与噪声区分开来, 并确定球是否可见。文献[35]用球员之间、场线和球的时空关系来估计球的路线。文献[62]用固定摄像头来检测三维球的轨迹并解释球滚动、控球状态下的球飞出及球出球场的四个不同运动阶段。文献[63]提出用视域非常狭隘的 16 PAL 制式摄像机来跟踪球员和球, 提取三维球的轨迹, 并计算出球员相当于越位线的坐标和位置。

1.2.1.3 目标颜色特征强化跟踪技术的研究

色彩特征不仅可以用来提高目标跟踪能力,而且还可以被用来分开属于不同球队的球员。文献[64]用颜色分类和分割来获取斑块以对应球队球员和裁判的运动服。通过将图像像素映射到各自的颜色类,利用色彩分类集来寻找感兴趣的区域,然后用形态算子来分组像素。文献[65]用混合颜色空间来检测空间,最佳区分属于对手球队球员的像素。文献[66]提出多摄像机跟踪球员的方法。通过交叉被分割的斑块的RGB直方图来分类球员服装类别,通过半监督策略获得了5种模式的直方图,该策略需要在比赛开始前对球员观察样本进行标签。文献[67]为足球比赛视频中的球员控球状态开发了一个半自动化系统。支持向量机的颜色直方图用于团队识别。对于每类人,作者事先手动确定一种区别于其他人的颜色,为每一颜色建立几个颜色素材库用来评估落入预置的颜色素材的像素分布,通过支持向量机,这种分配形成的颜色直方图用来评估团队。文献[58]从由摄像机通过图像和球场之间的单应性标定、旋转及缩放摄像机自我标定的单筒足球比赛视频中估计球员和球的位置。

1.2.1.4 遮挡问题解决技术的研究

在足球比赛过程中,由于背景的复杂变化,运动目标经常会出现部分或全部被遮挡的情况。为了解决遮挡条件下多目标跟踪问题,文献[68]通过图形表示来实施跟踪。通过用形态算子和前、后向图形表示来分割斑块,从而处理遮挡问题。文献[69]提出用协同多摄像机跟踪来解决遮挡和错误信息传播等问题。在球场上放置四个摄像机用于同步来解决一些遮挡事件。每个目标在每个视角中由一个专用的基于粒子过滤的局部跟踪器跟踪,不同视角的跟踪器通过置信传播与其他跟踪器互动。这样,在一个视角运行的局部跟踪器利用了从其他视角传递的附加信息,从而解决了运动员之间的互遮挡和自遮挡问题。

1.2.2 足球比赛视频中的运动员行为表示技术的研究

足球比赛视频中的运动员行为表示是足球比赛视频中的运动员行为识别和理解的关键问题之一,行为表示的好坏直接影响行为识别的识别率。研究者对行为表示进行了广泛的研究,其中比较典型的研究有:Bobick和

Davis 利用减背景法推导出时间模板的表示^[70]。该行为表示方法简单,但容易受噪声影响。Laptev 为视频数据的简洁表示提出了时空(ST)兴趣点,并探讨了利用时空兴趣点来描述人的行为的优势^[17]。

ST 特征不需要分割或跟踪个体行为的实施。利用这个性质,时空特征在行为识别方面取得了相当大的成功^[15, 71-74]。Niebles, Wang 和 Fei-Fei 提取时空兴趣点,将视频序列表示为时空词的集合,并用概率潜在语义分析(pLSA)模型来识别人的行为^[72]。pLSA 不能很好地描述行为的产生式模型,因为没有依据给一个新观察到的现象赋予一个概率值。另外,pLSA 模型中的参数随训练样本数线性增加,这表明该模型很容易过度拟合。Schuldt 和 Laptev 也使用 ST 特征,但他们更喜欢用码书和词袋来表示^[71, 77]。然而,上述提到的现有行为识别方法大多数研究单个人的行为识别。Efros 等人把足球比赛视频作为他们的实验数据,但他们也只识别了单个人的行为^[75]。Kong 等把光流作为行为特征来识别足球比赛视频中的团队行为。他们的研究局限于足球比赛视频且只能处理 3 类团队行为^[76],且易受噪声干扰。文献[77]采用时间序列的因果关系来描述成对行为,但难以推广到多目标的行为识别中来。文献[37]用一个完整的四维对象-实时交互张量来描述团队行为模式,通过学习并优化张量来减少内核,使它凝聚到一个可区别的时空互动矩阵中。在视觉变化的情况下,作为团队行为模式的简洁描述的时间互动矩阵被证实是稳定的。更重要的是,给定一个黎曼度量,所有时空互动矩阵集形成一个黎曼流形,可用它来建立概率框架以特征化团队行为模式每个类,但实现起来比较困难。文献[78]提出目标之间结构化的相互作用模型,使用基于目标的原语和低阶时空关系集成的概率框架从含噪声的感知数据中识别出高度结构化的多人行为。基于模型的目标识别和概率计划识别的表示由 4 个主要假设构成:① 在智能体之间进行团队活动时,单个智能体目标是指定的时空关系的天然的原子表示单位。② 在高度结构化的多智能体的行为识别过程中,行为的时间结构的高层次描述使用较少的低阶时空关系集和逻辑限制就足以表达智能体之间的关系。③ 贝叶斯网络为不确定的视觉感知特征的多种来源提供了一种适当的融合机制。④ 可以用自动生成的贝叶斯网络来融合不确定时态信息和计算对象轨迹数据集。文献[79]引入一个能明确对动态团队成员进行编码和证明计划识别形式的适用性的新的多智能体的计划表示法。从多智能体的计划表示法中提取的局部时间依赖性能够显著地修剪潜在有效的团队计划的假设集。