



网络空间安全与法治协同创新中心
COLLABORATIVE INNOVATION CENTER OF SECURITY
AND LAW FOR CYBERSPACE

反恐背景下的 信息技术革新研究

以视频序列中的行人检测与跟踪为例

刘钊·著

行人检测与跟踪是计算机视觉和模式识别领域的热点研究问题之一，在智能视频监控、智能交通、视频侦查和反恐以及人机交互等领域有着广泛的应用

时事出版社



本书由中国人民公安大学网络空间安全与
法治协同创新中心资助出版

反恐背景下的 信息技术革新研究

以视频序列中的行人检测与跟踪为例

刘钊·著

时事出版社

图书在版编目 (CIP) 数据

反恐背景下的信息技术革新研究：以视频序列中的行人
检测与跟踪为例/刘钊著. —北京：时事出版社，2017.8
ISBN 978-7-5195-0106-8

I. ①反… II. ①刘… III. ①信息技术—技术革新—
研究 IV. ①G202

中国版本图书馆 CIP 数据核字 (2017) 第 107561 号

出版发行：时事出版社

地 址：北京市海淀区万寿寺甲 2 号

邮 编：100081

发行热线：(010) 88547590 88547591

读者服务部：(010) 88547595

传 真：(010) 88547592

电子邮箱：shishichubanshe@sina.com

网 址：www.shishishe.com

印 刷：北京市昌平百善印刷厂

开本：787×1092 1/16 印张：9.75 字数：114 千字

2017 年 8 月第 1 版 2017 年 8 月第 1 次印刷

定价：60.00 元

(如有印装质量问题，请与本社发行部联系调换)

| 序 言 |

行人检测与跟踪是计算机视觉和模式识别领域的热点研究问题之一，在智能视频监控、智能交通、视频侦查和反恐以及人机交互等领域有着广泛的应用。行人受到视角、姿态、尺度变换，以及局部遮挡等影响，外观常发生很大的变化，从而给检测和跟踪带来一定的困难。本书从行人目标的稀疏特性着手，从有效描述行人目标的特征、增强外观模型的区分度和对行人形态变化进行部件建模等方面进行了深入研究。

本书的主要贡献包括：提出了用于行人检测的深度通道特征。在常见的通道特征的基础上，结合深度学习，从底层特征入手，通过神经网络在底层通道特征的指引下，有目的地学习能表示中层或更高层次的特征。特别是，对于深度通道特征的学习并非经预先设计，而是直接利用神经网络提取特征的能力，在已经有底层含义的通道特征上通过非监督学习获得。在公共数据集上的实验表明，所提出的深度通道特征可以充分描述行人的外观特性，有效

地提升行人检测的性能。

提出了多类别判别式字典学习方法并用于行人跟踪。从提高字典学习的区分度入手，在增加稀疏度的限定条件的帮助下，引入表述普遍信息的共享字典和表示类别特定信息的类别相关字典，并且在贝叶斯推理的基础上，讨论了分类决策依据和为了适应行人外观变化而进行的模型更新，实现了对行人目标的有效跟踪。在公共数据集上的实验表明，所提出的基于多类别判别式字典学习的跟踪方法可以有效地跟踪行人。

提出了基于多分量可变形部件模型的行人跟踪方法。将可以表述部件和部件之间关系的、可以随着目标形变而不断更新的可变部件模型作为行人外观表现，同时利用多个不同外观或表现的可变部件模型以适应行人外观的变化，维护多个描述目标不同外观的分量，通过图像解析估计行人位置，以最大后验概率选择最优位置作为跟踪结果。同时，为长时间跟踪目标，在模型不能很好地表现现有行人外观的情况下，利用行人检测方法，进行全局行人检测，协助跟踪算法进行错误恢复。在公共数据集上的实验表明，所提出的基于多分量可变形部件模型的行人跟踪方法可以有效地跟踪行人。

| Abstract |

Many intelligent video systems take pedestrian as primary target for analyzing. Subject to pose, scale and perspective changes, the appearance of pedestrian changes gradually. Thus, the detection and tracking algorithms expose to many difficulties. As a major element in pedestrian detection and tracking algorithms, pedestrian modeling contains features description, appearance modeling, appearance changing handling, etc. This thesis takes sparsity of the pedestrian object in video sequence in to consideration, finds the feature set which represents the appearance of the pedestrian well, increases the discriminability of the appearance model, and models the deformity of the pedestrian appearance. The main contributes of this thesis include:

We present a novel deep channel feature for pedestrian detection. Channel features have been approved effect in pedestrian description. Lacking the mid-level or high-level information, channel features don't

provide meaningful semantic information. Instead of using hand-crafted features, our method automatically learns deep features by introducing a convolutional neural network. The network is pre-trained by the unsupervised sparse filtering and learns a group of filters for each channel. We developed appearance-based pedestrian detection model using a mixture of the deep channel features with other low level channel features. Test results shows that our detectors also outperform many current state-of-the-art algorithms in both effectiveness and efficiency on public datasets.

We present a jointly learning multi-class discriminative dictionary for pedestrian tracking. Our method exploits concurrently the intra-class visual information and inter-class visual correlations to learn the shared dictionary and the class-specific dictionaries. By imposing several discrimination constraints into the objective function, the learnt dictionary is reconstructive, compressive and discriminative, thus can achieve better discriminate the object from the background. Tracking is carried out within a Bayesian inference framework where the joint decision measure is used to construct the observation model. Test results on public datasets have demonstrated that our tracker resists distracters excellently and generally outperforms existing methods.

We propose a novel online deformable part-based model (OLDPM) for pedestrian modeling and tracking. Motivated by the observation that the appearance of a pedestrian varies a lot in different perspectives or poses, we propose a multiple components model to represent distinct groups of the pedestrian appearances. Each component uses with one root and

several shared parts to represent the flexible structure and salient local patterns of one particular appearance. The top-down model extends the bottom-up model by introducing newly created OLDPMs for uncovered new appearances. The quantitative and qualitative experiments on challenge visual sequences validate our algorithm.



Contents

目 录

前 言 信息技术革新与新形态反恐	1
第一章 行人检测与行人跟踪研究现状	17
第一节 研究的目的和意义	17
第二节 国内外研究现状和发展趋势	18
一、行人检测研究现状	20
二、目标跟踪研究现状	31
第三节 行人跟踪与检测算法展望	38
第四节 研究内容	39
第五节 本书结构	39
第二章 基于深度通道特征的行人检测	41
第一节 引 言	41

第二节 深度卷积神经网络与稀疏滤波	42
一、深度卷积神经网络	43
二、稀疏滤波	47
第三节 深度通道特征	50
一、行人通道特征和深度通道特征	52
二、深度通道特征的提取	53
三、用于行人检测的特征构成	54
第四节 基于深度通道特征的行人检测	55
第五节 实验结果	56
一、实验设置	56
二、对比分析	60
三、决策树个数的影响	62
四、未识别情况分析	62
第六节 小 结	65
第三章 基于字典学习的行人跟踪方法	66
第一节 引 言	66
第二节 非冗余限定的字典学习	67
一、字典学习在分类与跟踪中的局限性	68
二、增加字典学习区分度的非冗余限定条件	70
第三节 多类别判别式字典学习跟踪算法	74
一、多类别判别式字典学习	74
二、跟踪算法与字典更新	75
第四节 实验结果	80

一、实验设置	83
二、整体性能比较	85
三、特殊预设状态性能比较	86
四、初始化对跟踪器性能的影响	87
第五节 小 结	89
第四章 基于多分量可变部件模型的行人跟踪	90
第一节 引 言	90
第二节 行人可变部件模型及其初始化	91
一、初始化可变部件	94
二、初始化 INCSVM	95
第三节 多分量可变部件模型	97
第四节 可变部件模型跟踪算法	98
一、跟踪目标函数	98
二、跟踪框架	99
第五节 自顶向下与自下而上相结合的跟踪框架	102
一、结合自顶向下策略的跟踪错误恢复	103
二、多分量可变模型的更新	104
三、难度负样本处理	106
第六节 实验结果	107
一、困难负样本对跟踪效果的提升	108
二、算法定量对比	109
三、算法对比定性分析	112
第七节 小 结	119

第五章 结论与展望	120
第一节 全书工作总结	120
第二节 未来工作	121
参考文献	123
后 记	140

前言 信息技术革新与 新形态反恐

恐怖主义伴随着人类历史的发展而不断进化蜕变，并跟随世界政治格局、经济形势和科技发展而呈现出不同态势。中国社会科学院世界经济与政治研究所研究员邵峰在总结新形势下反恐形势时指出，世界反恐形势愈加严峻，恐怖分子从伊拉克、阿富汗、叙利亚回流到欧美国家的现象增加，如何应对这些回流的极端人员可能发动的恐怖袭击是让人头疼的问题；恐怖分子混入难民中进入欧洲，因此形成的现实威胁和潜在的心理威慑将长期存在，对欧洲的政治、社会、安全局势构成严重挑战。

中国的反恐形势同样不容乐观。自 2007 年以来，中国境内恐怖主义活动进入活跃期。2008 年西藏拉萨地区“3·14”打砸抢烧暴力事件导致 18 人死亡，382 人受伤；2009 年新疆乌鲁木齐“7·15”打砸抢烧暴力事件导致 197 人死亡，1700 多人受伤；2015 年“9·18”新疆拜城煤矿暴恐事件造成 11 人死亡，18 人受伤。上述种种，

说明中国面临的恐怖主义威胁趋势严峻。

新形态恐怖主义呈现出多种新的特点，其中最显著的特点一是“独狼”式恐怖袭击增多，二是恐怖主义与信息技术革新相结合。

“独狼”式恐怖袭击，相对那些有组织的恐怖袭击活动而言更加难以防范。近年来，国际社会的主要注意力都集中在宗教型恐怖主义和民族分离型恐怖主义方面。这些恐怖主义逐步开始呈现从多人共同作案向分散式少数人作案或单人作案的转变。除了传统的极端宗教恐怖主义、极右翼恐怖主义、极端民族主义恐怖活动以外，当前出现的单一问题恐怖主义，即各种由于对某个具体问题不满，某些势力或个人对社会展开的恐怖袭击广泛存在于国际社会，在西方国家尤为突出。比如，欧洲的新法西斯主义出现上升势头，对难民、移民的个别攻击时有发生；对同性恋问题等一些社会现象抱有极端不满情绪的激进分子也有可能制造恐怖袭击，2016年6月12日，美国佛罗里达州奥兰多市“Pulse”夜总会针对同性恋群体的枪击案造成50人死亡。图1对2008年至2017年国内主流媒体公开报道的除西藏“3·14”事件和新疆“7·5”事件以外的共45件暴恐案件中的涉案恐怖分子人数进行了分析，绝大多数为5人以下的小规模恐怖袭击。在5人以下的小规模袭击中，有一半以上是由1至2人实施的。

恐怖主义与信息技术革新的结合彻底改变了恐怖活动的行为方式。随着互联网、社交平台甚至暗网在恐怖活动中的作用日趋上升，恐怖组织进行行动联络、袭击策划、人员招募和思想传播的主要途径已经发生了彻底改变，社交平台成为国际社会同恐怖组织斗争的新战场；信息的产生、传播、更新速度达到了前所未

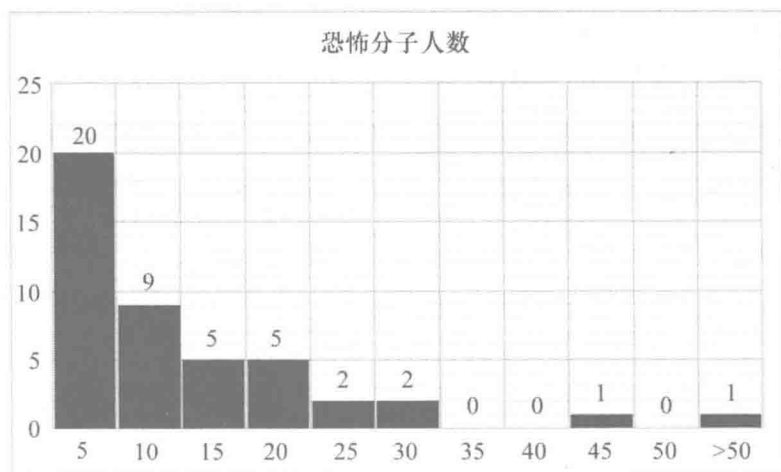


图1 涉案恐怖分子人数

有的迅速，对恐怖主义宣传、涉恐信息等方面的管控也承受着前所未有的压力。随着3D打印、无人机技术和人工智能的发展，恐怖袭击的手段也从枪击、冷武器袭击、自杀式爆炸等等发展到多方面多层级的袭击。例如，ISIS广泛采用无人机进行恐怖袭击，极大地降低了恐怖活动的成本，减少了恐怖分子自身的风险。

新常态下中国面临的恐怖活动的主要趋势包括：一是在新疆地区恐怖主义袭击事件仍有发生；二是境外的恐怖袭击对中国的海外利益和人员安全构成严重威胁，2016年以来连续发生了多起伤害中国公民和机构的恐怖袭击；三是“一带一路”沿线国家的反恐形势恶化对中国的周边环境安全和推进“一带一路”建设构成现实的重大威胁。在反恐形势恶化的压力下，中共中央总书记习近平多次围绕反恐维稳主题发表了一系列重要讲话，特别是在对做好新疆稳定

工作的重要讲话中强调“对暴力恐怖活动，必须保持严打高压态势，先发制敌，露头就打，打早、打小、打苗头”。从中可见，习近平同志对于反恐斗争中的事前预警和打击、事中及时响应提出了明确的要求。以预防代替处置将会在未来的反恐工作中占据更为重要的地位。

一、新形态反恐的技术融合

抛开对恐怖主义根源和国际恐怖主义演变的研究，单纯从新形态恐怖事件发生前的预警、发生时的快速响应和发生后的善后与总结等环节所需要的新技术的角度，将恐怖事件视为独立或连续的突发性公共安全事件，则新形态的反恐必然需要新的技术提供支持。新技术可以是信息技术、生物技术、机械制造、光学遥感技术、系统集成技术等一种或多种技术的融合。

（一）信息技术革新的反恐应用场景

信息技术革新的六大技术能够提供巨大的反恐潜力，包括系统集成技术、生物识别、非致命性武器、数据挖掘、纳米技术和定向能武器。尽管我们在不同领域都或多或少地取得了一些进展，但技术的发展永远是把“双刃剑”。如今互联网加密技术及移动互联网的成熟使恐怖主义活动更加难以预测。在智能技术的反恐应用上，成果还不足够显著，不过至少人工智能领域的突破让我们看到了一丝曙光。

每一次认知科学的进步都会带来信息技术的革新。从人类发现神经细胞的结构和功用开始,随着神经科学和脑科学的不断发展,以神经网络为代表的一系列机器学习算法得到了重视和飞速发展。没有花费太久的时间,深度学习已经攻克了计算机视觉领域的许多难题,包括人脸识别、车辆跟踪、目标分类等等。人脸识别已经逐步开始在交通枢纽投入使用,^① 视频跟踪和检测系统在目标跟踪与识别、自动防卫区域划定等领域的研究逐渐深入等 (Ge et al. n. d.)。



图2 2016年12月19日俄罗斯驻土耳其大使 Андрей

Геннадьевич Карлов 被刺杀现场视频截图,恐怖分子 Mevlüt Mert

Altıntaş 在背景中焦虑地无目的移动,其动作与一般安保人员差别较大。

^① “北京站试运行‘刷脸’进站使用人脸识别技术” 新华网, http://travel.news.cn/2016-02/05/c_128705565.htm。