

第一章

绪论

1.1. 研究背景

人们从外界环境获取信息，一般是通过视觉、触觉、听觉、嗅觉等感觉器官来进行的。研究已经表明，人类 75%~80% 的信息是由视觉来获取的。利用视觉信息具有实时性和直观性，比其他方式获得信息更容易被接受。近些年，随着计算机技术与电子技术的飞速发展，计算机硬件的性价比不断提高，已能满足视频图像处理的实时性要求，加之视频图像处理柔性大、适应能力强，应用前景十分广阔。人们逐渐尝试将计算机视觉用于公共安全监控、工业现场监控、居民小区监控、交通状态监控等各种监控场景中，能够实现犯罪预防、交通管制、意外防范和检测、老幼病残监护等功能，具有极广阔的市场应用前景。

近年来，视频目标的检测、分类、跟踪及运动行为理解等相关问题一直是计算机视觉研究的热点问题之一。由于这些研究热点具有非常重要的科学意义和广泛的应用前景，因此得到了国内外研究人员的广泛关注和浓

厚兴趣。

国际上每年都会召开很多与视频目标检测、分类、跟踪及运动行为理解相关的国际会议，其中一流的国际会议包括：International Conference on Computer Vision and Pattern Recognition (CVPR); International Conference on Computer Vision (ICCV); Asian Conference on Computer Vision (ACCV); European Conference on Computer Vision (ECCV); British Machine Vision Conference (BMVC); International Conference on Image Processing (ICIP); International Conference on Pattern Recognition (ICPR); IEEE International Workshop on Visual Surveillance (IWVS)等。这些会议收录了很多计算机视觉领域的最新的研究成果，其中的文章代表着该领域的当前的走向和热点。

在很多重要的国际期刊中也收录了大量视频目标检测、分类、跟踪及运动行为理解方面的论文，这些期刊包括：IEEE Transactions on Pattern Analysis and Machine Intelligence; IEEE Transactions on Image Processing; IEEE Transactions on Medical Imaging; IEEE Transactions on Circuits and Systems for Video Technology; IEEE Transactions on Vehicular Technology; International Journal of Computer Vision; Computer Vision and Image Understanding; Image and Vision Computing; Pattern Recognition; Pattern Recognition Letters; Machine Vision and Application; Real-time Imaging 等。

国外的一些大学和研究所，如卡内基梅隆大学、斯坦福大学、麻省理

工大学、牛津大学、剑桥大学、多伦多大学、瑞典皇家工学院、大阪大学、MERL (Mitsubishi Electric Research Laboratories)实验室等都设有专门的计算机视觉研究组或者研究实验室，对视频目标检测、分类、跟踪及运动行为理解等方面进行了专项研究，并且将这些研究成果应用到实际中。

本专著以视频监控和交通工程为主要应用背景，将本人近几年独立和合作完成的研究成果进行了介绍，主要从阴影检测、目标检测、目标分类、目标跟踪、和行为理解等这几个主要方面进行了详细介绍，并在不同章节中梳理了近几年来出现的各类相关的优秀算法，希望能够为从事计算机视觉领域的相关专家和工程师提供一些参考和借鉴。

1.2. 研究意义

随着计算机技术与电子技术的飞速发展，计算机硬件的性价比不断提高，已能满足视频图像处理的实时性要求，加之视频图像处理柔性大、适应能力强，使得视频图像处理技术几乎可用于计算机视觉的所有相关应用。

目前，计算机视觉研究领域中很多热点问题都是艰难的挑战，主要原因在于：

(1) 复杂环境下的图像信息往往易受到复杂背景、噪声、光照变化等干扰因素的影响，这使得运动目标检测在实际应用中相当困难。

(2) 视频目标可能会发生旋转、缩放、位移等各种复杂的变化，单一特征往往难以保证准确识别目标。

(3) 在光照变化、前景背景相似、目标遮挡等情况下，目标的识别更加困难。

总的说来，当前对计算机视觉相关问题的研究尚处于对于特定问题设计特定方法的阶段，对复杂多变环境的自适应能力很差。环境噪声、照度的变化、目标边缘或区域特征不明显、背景复杂、遮挡、前景背景相似等各种因素的存在都会影响系统的性能。

本专著在借鉴已有研究成果的基础上，重点从阴影检测、目标检测、目标分类、目标跟踪、行为理解等这几个方面进行分析与综合研究，并在此基础上提出新的设计方法，这为视频目标的后续处理奠定基础。目前，这几方面的研究内容还有许多关键和突出的问题没有解决，主要有以下几方面：

(1) 准确分割目标与背景

虽然分割前景目标和背景是视频目标检测过程中最基础的部分，但却是非常重要的部分，对它的研究仍然非常活跃。长期以来，前景与背景分割技术的研究和应用备受研究人员的关注。减背景法是当前前景目标与背景分割的主要技术之一，其核心问题是如何建立一个自适应的背景模型。

一个性能良好的背景模型要能够有效描述各种条件下背景在空域和时域所发生的变化，而这是非常困难的。当前的背景建模方法或者过于复杂或者适应能力差，归咎起因，主要是缺少一个有效的信息描述和处理的方法，而这也是本文研究重点之所在。

(2) 准确分割目标与阴影

由于视频目标的运动阴影与背景有着明显的差别，因而运动阴影常被当成运动的前景目标被一起检测出来。而前景目标的阴影会影响目标的真实形状，产生大量的虚假运动区域。当前，阴影的检测仍有一定的困难：首先，阴影和运动目标一样，都显著区别于背景；其次，大多数情况下阴影和投射它们的运动目标是相邻的，而且具有相同的运动规律。在目标的检测中，由于阴影与目标相连而被当作是目标的一部分，影响对目标形状的判断；最后，两个甚至多个目标还可能因为阴影而互相粘连在一起，严重影响对目标的判断和识别。

(3) 更准确的目标分类

运动目标检测出来之后，接下来就是要把感兴趣的目标从其它目标中区分出来。然而，对于同种类别的目标，它们的外在表现形式往往是不同的。比如轿车、客车和货车，虽然它们属于同一类别，但它们的外在形状不一样，计算机很有可能将它们区分成三类。而属于不同类别的目标对象

也会因为图像数量的不断增加而具有相近的内容，比如说，在很多场合下姿态表现相似的人和猩猩这个两个类别。因此，一个鲁棒的目标分类算法要能够适应同种类别图像以及不同类别的图像变化。

尤其在复杂的交通环境下，视频目标受距离、尺度、姿态、外貌、视角、状态和噪声等因素的影响较大，现有的目标分类算法有必要在鲁棒性和准确性方面进一步提高。

(4) 更快更准确的目标跟踪

在跟踪过程中，由于被跟踪目标姿态发生改变、运动目标所处环境的光照变化、目标被部分遮挡所引起的运动目标不规则变形、目标之间或者目标和背景之间的遮挡导致目标暂时消失等原因，目标的准确跟踪定位是一个非常困难的问题。在通常情况下，使用单一特征信息(如颜色)对目标跟踪，很难达到良好的跟踪效果。因此，基于多线索融合的视频目标跟踪引起了人们越来越多的关注。

视频目标跟踪算法处理的对象是由一帧帧视频图像组成的图像序列，每帧图像中都包含着巨大的信息量。在实际应用中，通常需要对视频图像进行实时的采集和分析，这就需要算法能在很短的时间内提取出目标的大小、位置、形状、轮廓等信息，这就是视频目标跟踪算法的实时性要求。

(5) 目标的行为理解及描述

对一个视频监控系统的操作人员来说，他的真正目的往往不是对视频目标的准确快速跟踪，而是对目标的运动行为的理解和描述。也就是说，人们通常不关心目标在每一帧中的位置，而更关心目标是在做什么(如走、跑、蹲、跳、爬等)，以及目标运动对场景造成的影响(如放置物体、进入危险区域等)，并希望计算机能够用人类的语言把目标的行为表达出来(如目标进入警戒区、目标拿了东西走、车辆逆行等)。

目标的行为理解和描述本质上是一个人工智能的问题。相对于目标检测和跟踪方面技术的显著发展，当前对目标的行为理解和描述的研究还不多，仍然处于萌芽阶段。

基于上述分析，本专著围绕视频处理技术，对复杂背景环境下的阴影检测、目标检测、目标分类、目标跟踪、行为理解等相关问题进行了研究。本专著的研究内容涉及到计算机视觉技术的主要方面，其研究成果是视频处理技术的法杖，提出的新思想和新方法具有指导性意义和应用价值。在本专著中，只研究固定单摄像机、复杂场景下的相关问题，不涉及移动摄像机和多摄像机的情况。

1.3. 本专著的主要研究内容

根据 1.2 节所提出的阴影检测、目标检测、目标分类、目标跟踪、行为理解的几个关键问题，本专著分章进行了相关问题的研究。

第二章对运动阴影检测问题进行了研究。提出了一种基于 TSK 模糊系统的运动阴影检测方法。该方法通过 TSK 模糊系统融合了扩展的 LBP 纹理特征和颜色特征，并针对这两种特征分别使用了前景像素和背景像素的特征差异性度量。为了训练该模糊系统，还结合了粒子群优化算法和递归最小二乘估计算法。通过在不同的视频序列上进行测试，所建议方法比其他方法具有更高的阴影检测性能。

第三章对运动目标检测问题进行了研究。提出了基于复模糊逻辑系统的运动目标检测方法。使用复 TSK 模糊逻辑系统作为背景模型的估计，并结合粒子群优化(PSO)算法和核最小均方算法(KLMS)来学习该复模糊逻辑系统；将前景像素看作背景像素的例外像素，提出了去除前景像素的方法，然后再去学习复模糊逻辑系统；最后根据估计的背景模型和前景图像的比较判断出前景像素。

第四章对运动目标检测问题进行了研究。提出应用 Hough 森林进行目标检测，并对经典 Hough 森林存在的不足进行改进。借助 Boosting 算法的思想，学习图片块样本和目标对象样本的自适应权重分布，并分别优化用于构造随机树和 Hough 森林的代价损失函数；最后根据图片块样本的权重分布，提出了改进的类标志不确定性度量方法。基于 Hough 森林的代价损失函数，还提出了随机树权重的学习方法。实验结果表明，本文所建议的方法相比于其他方法具有更高的目标检测性能。

第五章对图像分类问题进行了研究。提出了一种基于 Adaboost 算法的图像到类距离学习的分类方法。首先将图像到类的距离进行阈值化处理，并使用线性分段函数作为图像到类距离的评价函数，然后将该评价函数作为弱分类器加入到 Adaboost 算法中生成一个强分类器。为了选择最优的弱分类器，使用粒子群优化算法和误差最小化原则去优化弱分类器。最后通过实验验证，本文所建议方法在 Scene-15 和 Caltech-101 图像数据集上比其他方法具有更高的分类性能。

第六章对图像分类问题进行了研究。提出了一种基于正负模糊规则系统的分类方法。首先，根据图像特征到多个码本的距离，使用高斯型隶属度函数构造多个模糊隶属度直方图去描述图像；然后，使用正负模糊规则系统将图像描述和图像分类这两个过程融合在一个框架下学习，并使用粒子群优化算法和递推最小二乘算法来学习正负模糊规则系统的前件和后件参数，使得描述图像的参数在学习过程中不断被调整，从而使得图像描述能更好地迎合图像分类任务；最后，在 PASCAL VOC2007 图像数据集上的实验结果表明，所建议方法比原始的模糊词袋模型具有更高的图像分类性能。

第七章对目标跟踪问题进行了研究。提出了一种改进的连续自适应均值漂移算法，避免了原算法跟踪目标限于人脸的缺点，使算法可以对其他目标进行有效跟踪。在算法中，使用 HSV 三维直方图来代替原算法中的

单独的 H 直方图，并提出了 H、S、V 直方图各自的级数的灵活的分配方法。在算法中，还加入了对直方图反向投影操作产生的概率分布图像的分析 and 处理，以减少背景中相同颜色的干扰。同时，通过对目标位置的预测和调整，使用了原算法中没有利用的过往运动信息，提高了跟踪的准确程度。

第八章对目标跟踪问题进行了研究。提出了一种基于中心直方图和周边直方图对比的多线索概率分布图像融合的视频目标跟踪算法。文中采用颜色、运动历史、视觉显著性等多种概率分布图像来描述目标的观测信息，通过目标在前一帧的位置确定中心区域和周边区域，并计算不同概率分布图像基于这两个区域的直方图，即中心直方图和周边直方图。每个概率分布图像的置信度由其中心直方图和周边直方图的差异度来描述。根据不同线索的置信度，在线调节当前帧各种概率分布图像在目标位置判断中所占的权重，实现对目标的多线索融合跟踪。实验结果表明，文中算法比起常用的固定线索权值的融合算法，效果更好。

第九章对视频图像序列的多目标跟踪及其行为理解与描述问题进行了初步研究。首先建立目标的运动模型，并获取前景团块。接着分析前景团块与预测的目标的相似度，建立前景团块与目标之间的关联矩阵，并产生目标-团块匹配链表和团块-目标匹配链表。根据这两个链表，确定前景团块与目标的对应关系，实现了对视频多目标的跟踪。通过确定标记区域，限制目标异常行为的作用范围，并提供适当的行为指引。用团块和前景目标

集合定义了多种目标行为，实现了对目标行为的识别。通过识别的目标行为的先后次序，实现了用自然语言对场景的描述。

第十章对如何使用视觉分析结果改善行人跟踪效果进行了研究，提出了一种基于 MPEG-7 标准描述的行人跟踪优化算法。该算法核心思想是首先对已有的视觉分析结果进行描述，然后使用描述对行人跟踪算法进行优化，提高跟踪性能，同时优化后的跟踪结果反馈给描述信息，进一步改善跟踪效果。

最后，对本文所做的研究工作进行总结，并对今后努力的方向进行讨论和展望。