摘要

目标跟踪一直以来都是计算机视觉研究的热点问题，在国防军事、智能交通和生活安防等众多领域得到了广泛的应用。近年来，视觉注意机制（显著性检测）的研究取得了实质性的进展和突破。将显著性检测引入到目标跟踪中，可以使目标处理过程更接近于人类认知机制，并提高算法的准确性。因此，将显著性检测应用到目标跟踪中具有重要的意义和应用价值。本文在均值漂移跟踪算法的框架下，通过模拟人类视觉特性，提出了基于（视觉）显著性检测的目标跟踪算法，研究工作主要包括以下两个方面。

1）提出一种利用目标先验信息的时空显著性检测（模型）算法。时间显著性检测方面，引入时域运动特征，利用光流法检测运动矢量，并对其矢量场进行滤波处理。利用运动熵值对运动矢量的显著性进行分析，建立能够表征场景中运动情况的时间显著图。空间显著性检测方面，本文模拟人类视觉特性的“记忆性”，利用初始目标区域的颜色先验信息改进自底向上的（认知视觉注意）显著性检测模型，使得生成的空间显著图能够更好地适应目标跟踪任务。最后，时空信息融合阶段，根据人眼对不同场景的主观感知，提出自适应加权的融合方法，通过设置自适应权重得到时空显著图。

2）针对传统均值漂移跟踪算法采用单一颜色特征导致跟踪鲁棒性不强的问题，提出一种基于视觉显著性特征的目标跟踪算法。首先，利用本文提出的显著性检测模型计算时空显著图，将颜色特征和视觉显著性特征融合，用以描述目标模型，并根据相似系数的大小自适应地调节转移向量融合权值，避免采用单一颜色特征跟踪时易发生的漂移现象。其次，提出核函数带宽的自适应调整策略，以适应目标的尺度变化，并且引入卡尔曼滤波器预测目标的位置，实现了复杂背景下的准确目标跟踪。实验结果表明，本文依据目标先验信息构建的显著性检测算法，检测效果较好，将其应用于均值漂移跟踪框架中，能够有效弥补仅采用颜色特征的不足，解决了目标跟踪过程中候选目标信息描述单一的问题，提高跟踪的鲁棒性和准确性。

关键词：目标跟踪，显著性检测，特征融合，均值漂移

1. 绪论

1.1课题研究背景及意义

计算机视觉是融合了图像处理、计算机图形学、模式识别、人工智能、心理学和数学等领域的一门交叉性很强的学科。计算机视觉研究的目的是利用计算机技术来处理序列图像，以完成对图像中目标的检测、识别、跟踪等任务。近些年来，计算机视觉的研究重点已经逐渐从对静态图像的形态学操作过渡到对动态图像的关联分析，己经成为广大国内外学者研究的热点问题。目标跟踪作为计算机视觉领域的核心研究方向，己经有了几十年的发展历史。目标跟踪就是在序列图片中将目标与背景持续分离的过程［5］，一般利用相关特征为目标建立模板，采用一定的跟踪策略，在图片当中搜索到与目标模板最相似的位置作为跟踪结果。其研究成果在人机交互［１］、视频监控［２］、智能交通［３］、机器人视觉定位导航［４］等多个方面都有着广泛的应用。因而研究目标跟踪技术有着重大的实用价值和广阔的发展前景。

目标跟踪的应用主要有以下几大方面：

1）智能安防监控系统：智能视频监控技术是计算机视觉与人工智能的一个重要研究领域，该技术的目标是用计算机视觉的方法,在不需要人为干预的情况下,通过对摄像机拍录的视频序列进行自动分析来对被监控场景中的变化进行定位、识别和跟踪,并在此基础上分析和判断有关目标的行为,从而既能完成日常管理,又能对异常情况及时做出反应。

2）机器人技术：机器人系统中对外界数字图像信息的处理和分析是机器人进行后续行为分析的基础，机器人跟踪系统主要包括对行为人的跟踪、特定路径的跟踪和固定物体目标的跟踪三个方面。1994 年美国斯坦福大学研制的 Flakey 移动机器人成功完成了对办公区域内行为人的识别与跟踪，随着现在视觉认知模型的发展和计算机技术的提高，越来越多的学术科研人员将人类视觉注意机制应用于机器人视觉系统，使得其能够快速、实时和智能的处理复杂多变的庞大外界信息，并从中找出最相关的部分并做出相应的反应。

3）军事领域：近年来人们将视觉跟踪技术应用于各种安全防卫系统、导弹的精确定位、作战侦查和武器的精确定位控制等方面。在军事作战和环境考察过程中，存在一些极度危险和人类无法亲自到达的地方，此时就需要智能机器人来完成探索任务。在军事领域中的另一热点研究为武器的精确制导，武器中装备能够精确定位目标的系统组件，通过视频传感技术采集窗口信息，并应用视频目标跟踪技术准确定位攻击目标，在这个过程中跟踪的准确性和稳定性就显得尤为重要。

近年来，视觉显著性检测作为计算机视觉领域的热门方向受到广泛关注。所谓显著性，是指一个物体从其周围环境突显出来的特性。人类视觉系统具有高效的数据筛选能力，能够迅速将注意力集中到显著目标上。视觉显著性检测试图模拟视觉系统的这一能力对图像信息进行筛选处理。与一般的图像或视频处理方法不同，显著性检测模型是基于已经被证实的人类视觉对于外界处理的机制建立起来的图像和视频的处理框架。它更加符合人类视觉对于图像和视频的感知，而传统的计算机视觉处理方法大多由任务驱动，由所要完成的任务来决定处理的过程。在图像和视频的处理与分析的应用中，视觉显著性的计算往往作为一个预处理的过程，是图像与视频分割、目标检测以及目标跟踪等任务的重要基础，它为解决这些问题提供了一条新的思路。不仅如此，显著性检测在目标识别、图像视频的自适应压缩、视频检索等领域也有重要应用。高效而准确地对图像和视频的显著区域提取对后续的处理和应用都有着积极的影响。将视觉显著性检测模型引入目标跟踪系统中，通过模拟人类特有的视觉选择性注意机制，可以为目标的搜索提供先验知识，势必会极大提高跟踪的鲁棒性，是一个值得尝试的方向。

1.2国内外研究现状

1.2.1 目标跟踪技术

目标跟踪作为计算机视觉领域最具挑战性的研究课题之一，受到国内外许多研究者的广泛关注。近年来许多学术会议将其作为研究讨论热点，其中包括机器视觉与模式识别领域的著名国际会议ICCV(International Conference on Computer Vision)和CVPR(International Conference on Computer Vision and Pattern Recognition)。在国外许多大学和研究机构一直在进行目标跟踪的研究，如：卡内基·梅隆大学的 Digital Mapping 实验室，康奈尔大学的机器人与计算机视觉实验室等。在国内也有许多团队在做目标跟踪方面的工作，如西安交通大学人工智能与机器人研究所，香港中文大学，上海交通大学图像处理与模式识别研究所等。中国科学院自动化研究所下属的模式识别国家重点实验室，已经在智能交通监控、人脸识别与跟踪、异常行为检测方面取得了巨大研究成果。

视觉跟踪算法众多且各自具有优缺点，广大学者针对不同的跟踪算法进行优化改进，并取得了卓越的效果。视觉跟踪方法可以分为生成式跟踪方法(Generative tracking methods)和判别式跟踪方法(Discriminative tracking methods)两大类。

1）生成式跟踪方法通过学习目标的知识建立目标模型，并将跟踪看作寻找与目标模板最相似的候选区域的过程，其中最具代表性的就是均值漂移(Mean Shift)算法、粒子滤波(Particle Filter)算法和卡尔曼滤波(Kalman Filter)算法。Mean Shift算法是基于密度梯度上升的一种非参数方法，它通过迭代来寻找跟踪目标的位置，从而实现对目标的有效跟踪，它是核函数统计迭代算法应用到目标跟踪的代表。Particle Filter算法是将贝叶斯滤波应用到目标跟踪上的代表，它将目标跟踪过程看作是目标状态的概率密度函数在时间上的传播过程。它的核心思想来源于蒙特卡洛方法(Monte Carlo methods)，将目标状态的后验概率密度通过一组带有权重的离散粒子集来表示，并利用从该后验概率中随机抽取的离散粒子来表示目标状态的分布情况，从而准确定位目标位置，实现对目标的有效跟踪。Kalman Filter算法在目标跟踪中运用滤波原理，通过建立目标的运动模型预测估计目标位置。该算法为一种最优滤波方法，但是其要求系统噪声统计模型为高斯分布。 近些年涌现了许多基于以上算法的改进算法，使其得到不断的完善。例如GR Bradski 在Mean Shift算法的基础上提出的连续自适应Mean Shift算法，即CamShift(Continuously Adaptive Mean Shift)算法，使得跟踪窗口可以自适应地根据目标尺度的变化而变化。该方法提出之后迅速成为目标跟踪领域的研究热点，为许多研究人员开辟了新思路。

2）判别式跟踪方法也称为基于检测的跟踪(Tracking by Detection)，将跟踪问题看作目标与背景的分类问题，是一种在线识别学习算法，通过寻找训练分类器的决策边界，有效分离目标与背景样本。典型方法有Adaboost算法、K近邻(K-Nearest Neighbor，KNN）分类算法、支持向量机(Support Vector Machine，SVM)等。Adaboost算法是判别式跟踪中广泛应用的算法，它将目标跟踪过程看作简单的目标和背景的二分类问题，在对目标进行跟踪时，Adaboost算法通过对一组训练集合中的每个样本赋予不同的分布概率，从而得到不同的训练集合，并且对每一个训练集合进行训练从而产生若干个不同的弱分类器，最后将这些弱分类器通过不同的权值结合起来，就组成了最终的强分类器。K近邻分类算法的决策分类只与其最近邻的几个样本有关，如果一个样本的最近邻样本中的大多数样本是属于某一个类别的，那么判别该样本也是属于这个类别的，其中选择的最近邻样本是已经正确分类的。SVM是一种监督式学习的方法，它将低维空间向量映射到一个更高维的空间中，然后在这个高维空间中建立一个最大间隔超平面，该最大间隔超平面使得两个不同类别的数据间隔最大，并通过这个最大间隔超平面来对目标与背景进行有效的分类。判别式跟踪方法的有效性主要依赖于特征的选择和分类器的有效更新，跟踪器的识别能力与特征空间的选择有直接的关系，所选特征区分背景与目标的能力越强，跟踪分类效果越好，反之亦然。

1.2.2 显著性检测技术

人类视觉系统具有从大量复杂的视觉信息中迅速识别目标的特点，视觉显著性检测能够模拟视觉系统的这一能力对图像信息进行筛选处理，在目标检测、图像分割、视频压缩等方面具有重要的意义，如今已经成为计算机视觉领域的一个新的热点问题。现有的显著性检测模型大体可以分为三类。第一类是基于底层视觉特征的视觉注意模型，目前应用最为广泛的是 Itti 等[20]提出的认知视觉注意模型。该模型提取场景中的特征显著图并采用线性合并的方式整合为最终显著图，并以赢者取全和返回抑制相结合的方式引导视觉注意焦点的选择和转移。第二类是基于纯数学计算的视觉注意模型。比如 Hou 等[22]提出的基于频域的谱残差法(SR)，对图像进行二维傅里叶变换后得到频域的相位谱和幅度谱，认为频谱域上的统计奇异对应图像的异常区域，因此该区域的物体显著性高。四元相位谱傅里叶变换法(PQFT)[23]是基于频域的注意模型，利用图像的相位谱保留各个频率成分的位置信息，并在颜色、亮度等静态初级视觉特征基础上，增加了运动信息通道。第三类是将前两种相结合的方法，最具代表性的是 Schölkopf 等[24]提出的基于图论的算法(GBVS)，底层特征采用场景中颜色、亮度、方向特征，在显著图的生成部分提出引入马尔科夫链的方法来计算显著性。

视频由具有较强相关性的图像序列组成，研究表明，视频条件下动态语义特征的显著性远高于与静态特征的显著性[34]，因此视频的显著性检测模型往往需要加入运动信息。Itti在认知视觉注意模型的框架下引入了运动特征和帧间闪烁[13]，使得改进后的模型能够检测运动信息。Guo[14]等人则将谱残差方法扩展到视频中，对视频帧分别提取运动、红绿对比、蓝橙对比色以及亮度特征，然后通过四元傅里叶变换得到各特征的相位谱，最后将各通道融合成最终的显著图。Lu[15]等人在他们的显著性模型中使用了底层特征如颜色、纹理和运动以及一些认知特征。Cheng[16]等人也在他们的静态图像的显著模型中加入了运动信息，通过分析水平方向和垂直方向的像素运动，计算显著图。Bioman[17]等人提出了视频中的时空维度上不规则性的检测方法，该方法通过视频块的2维和3维纹理与训练数据进行比较，检测得到视频中的不规则的运动。Meur[18]等提出了基于视觉注意机制的时空模型，该方法的时间显著图是通过对仿射分量的分析得到的。Kienzle[19]通过对眼动数据的学习，构建了基于时空兴趣点的检测器对输入信号分别在时间和空间域内滤波，从而检测得到显著目标。

1.3目标跟踪的挑战和难点

运动目标跟踪技术虽然经历了几十年的发展，但各种算法大多都有其局限性，需要满足一定的使用条件，到目前为止，提髙跟踪的鲁棒性依然是运动目标跟踪中的核屯、课题。影响跟踪鲁棒性的因素非常多，化如图像从Ｈ维空间投影到二维空间中发生的信息丢失、图像自身的噪声[9]等。但总的看来，图片中目标与背景的复杂性是影响跟踪鲁捧性的两个最主要因素，详细描述如下：

1）目标自身不确定性

主要包括在跟踪中出现的目标自遮挡和姿态变化。自遮挡，是指目标在跟踪过程中出现的水平旋转现象，当目标发生自遮挡后，目标表观在图片中会发生很大变化，导致与模板差异较大，给跟踪带来难度；姿态变化，是指非刚体目标出现的表观形变等变化，也会导致目标区域与其模板有一定差异，若目标外观模型和模板更新机制不能适应这种变化，就会出现＂漂移＂现象，因此跟踪算法的外观模型和模板更新机制的设计是技术关键。

2）目标所处背景的复杂性

主要包括场景光照变化、目标与背景的相似性、目标被静止的障碍物遮挡等情况。颜色特征是在目标检测与跟踪中经常使用的特征，其对部分遮挡和姿态变化抵抗性较强，但是对光照变化非常敏感，当场景发生光照变化时，会对目标表观颜色造成很大改变，导致模板的不准确；目标跟踪实际是目标与背景的持续分离问题，当目标周围出现相似背景的干扰的时候，会降低这种区分性，可能会造成跟踪结果转移到该相似背景上，出现误跟踪现象；遮挡是跟踪中最常见，也是最难处理的问题之一，如何对遮挡情况进行正确判断以及充分利用目标未被遮挡部分的信息进行跟踪是解决送类问题的关键。

图1.1为实际跟踪过程中遇到的一些挑战，分别为光照变化、目标被静止障碍物遮挡、姿态变化、目标自遮挡等，好的鲁棒性跟踪算法要能应对其中的一两种挑战。

1.4论文的主要研究内容

1.4.1论文的主要内容

本文的研究建立在视觉跟踪算法的基础之上，针对目标跟踪技术的不同环节和存在的技术难题进行了改进和优化，并在VisualStudio2013仿真平台上进行了验证和分析。本文尝试将显著性检测应用到目标跟踪当中，提出了基于视觉显著性特征的目标跟踪算法。首先，对自底向上和自顶向下的视觉显著性检测模型，以及运动显著性的引入进行了研究；在此基础上，针对目标跟踪问题构建利用目标先验信息的时空显著性检测模型；最后，利用提出的显著性检测模型对CamShift跟踪算法进行改进，以视觉显著性特征为注意引导实现目标跟踪。本文的研究内容主要包括以下三个部分。

1）构建空域静态视觉显著性模型。在对“自底向上”显著性算法分析的基础上，通过其在目标跟踪方面这一具体应用，引入基于目标的自顶向下的注意，提出将目标先验信息融入显著性检测算法当中。人眼在跟踪物体时，总能凭借对物体的记忆，通过将注意力转移到具备物体特征的区域来保持对该物体的跟踪。本文构建的空域显著性模型正是以这种现象为依据，通过在基于颜色对比度的显著性检测算法中加入目标先验信息，使具备目标特征的区域在显著图中显著度较高，更符合人类视觉的感知过程。

2）构建时空域融合的视觉显著性模型。运动特征在目标跟踪中占有主导地位，因此在空域静态显著性模型的基础上，引入时域运动特征通道，利用光流法检测运动场，利用运动熵值计算时间显著性，建立起时空域融合的视觉显著性模型。选取 i LIDS database of AVSS 2007 conference 标准库、PETS2006数据库以及实拍行人视频进行实验仿真。结果表明，与现有的多个视觉注意模型相比，本文提出的模型在行人检测中具有较高的准确率。

3）提出基于视觉显著性的行人跟踪算法。首先，采用本文提出的视觉显著性模型生成显著图，利用显著性特征对CamShift跟踪算法中的反向投影图进行融合加权，充分发挥视觉显著性特征对于跟踪过程中出现的光照变化和姿态变化的不变性，从而提高跟踪鲁棒性。最后，在CamShift跟踪框架下，以显著性特征为注意引导，实现目标跟踪。

本文对均值漂移跟踪算法中的反向投影图进行显著性加权，使得跟踪的结果不仅仅在颜色信息上与目标是相似的，而且还具有较高的显著性。实验表明，本文算法在行人跟踪中取得了良好的效果。

1.4.2论文各章节结构安排

本文主要分为五个章节，各章节的内容主要有：

第一章主要介绍本设计的研究背景和意义、研究现状、挑战和难点、本文的主要工作以及本文的结构安排。

第二章介绍信道化技术的基本理论与信号频谱检测方法。整体阐述信道化技术，介绍当前成熟的均匀信道化技术，滤波器组技术同信道化技术二者的关联引出动态信道化技术的概念。阐述了基于能量检测的频谱检测方法，分析了影响改方法性能的窗函数、信号重叠率等因素。并做出了不同窗函数、信号重叠率下检测性能的对比，并给出了相应的信号频谱的仿真验证结果。

第三章给出了一种均匀信道化技术的改进方法与一种动态信道化的实现方法。利用频率响应屏蔽技术设计窄过渡带带宽的优势，提出了一种均匀信道化技术的改进方法，该方法在相同性能下可以有效解决硬件资源。利用调制滤波器组的重构特性，本文给出了基于调制滤波器组的非均匀信道化结构，并以该结构为基础给出了动态信道化的实现方案。并对动态信道化进行了仿真验证。

第四章从工程应用的角度上研究动态信道化方法的实现。介绍新一代SCA架构的软件无线电平台的特点和整体系统结构，针对平台特点，给出了动态信道化方法的在平台上的软硬件结合的实现方案，并重点介绍了硬件FPGA的信号处理流程与上位机软件重构的程序流程。

第五章系统的实测与分析。给出了所实现的系统所应用的测试工具、测试环境、测试方法与实际的测试结果。并对测试的结果的误差进行了有效的分析。

最后对本文的整体研究思路进行了总结，并在自己研究设计的基础上对信道化的研究前景进行了展望，以及为下一步的具体工作提出了几点建议，同时对研究生生涯中给予我重大帮助与支持的以及同学和朋友表达我衷心的致谢。

第二章 目标跟踪算法与视觉显著性检测模型

2.1引言

2.2目标跟踪算法

2.3视觉显著性理论及其算法介绍

2.1.1视觉显著性概述

视觉感知是人类了解世界最重要的方式之一。当我们挣开双眼，各种海量信息便蜂拥而至。进入人类视野的信息如此复杂多样，但神奇的是，人类总能迅速把注意力集中到自己感兴趣的目标上。这一切主要得益于人类视觉系统在处理外界信息时表现出来的卓越的数据筛选能力。由于人类视觉系统的资源是有限的，当面对涌入的视觉信息时，它会有选择地筛选出它所认为的重要信息，进而对这些信息进行处理，而其它信息则被视觉系统过滤掉。这种具有选择性、主动性和记忆性的心理现象被称为视觉注意机制。当前一种广泛被学界接受的说法是：人类的视觉注意力是由显著性机制驱动的。显著性是由多种视觉敏感特征引起的一种局部反差，一般表现为“信息中最为特殊的部分”或“最为突出和值得关注的部分”。

计算机视觉领域的显著性检测指的是利用图像处理的方法模拟人类视觉注意系统的机能对视场中信息的重要程度进行计算。通过采用相应的显著性检测算法对输入图像进行处理得到视觉显著图。显著图一般是一幅单通道的灰度图，在显著图中，灰度值越高的地方越显著，灰度值较高的区域被认为是显著性区域。视觉显著性在人类的目标跟踪机制中起到十分重要的作用，许多专家学者将视觉显著性与目标跟踪相结合，提出了目标跟踪与视觉显著机制的假设学说[4]，即目标跟踪是通过判别目标背景区域中的显著性区域，在任务或数据驱动下完成基于目标特征调节的认知机制。目标跟踪的视觉显著性假设理解主要包含三方面的内容：(1)在跟踪过程中，视觉显著区域的贡献值大于非显著区域对目标跟踪结果的准确性贡献值；(2)跟踪结果的有效性很大程度上依赖于显著性检测过程中视觉显著特征的选择过程；(3)人类的视觉注意机制和目标跟踪机制都可由低级的生物神经机制来完成实现。该机制是视觉显著跟踪的基础性假设，并得到广大学者的论证。

“自顶向下”显著性检测算法主要跟特定任务相关，如在对特定目标进行检测的过程中，可将该目标的特征模板当作先验信息，融入显著性检测算法当中，提高生成显著图的质量。

近十多年来，随着视觉显著性研究的活跃，国内外研究人员提出了各种各样的视觉显著性检测模型，一是特征驱动的“自底向上”模型，二是任务驱动的“自顶向下”模型。其中，“自底向上”的视觉显著性检测模型是由场景底层数据驱动的，通过提取底层视觉特征来计算场景的视觉显著性。在自底向上的方式中，吸引注意的区域是那些和周围特征相比足够有判别力的区域。“自顶向下”的视觉显著性计算模型是指考虑了自顶向下因素的显著性计算模型。这些模型大多是基于具体任务，如视觉搜索、玩游戏、汽车驾驶等。由于自顶向下因素的多样性和复杂性，因此建模自顶向下的注意（计算显著性）是一项较为困难的工作。目前针对该模型的研究工作相对较少，下面对“自底向上”和“自顶向下”两种情况中一些有代表性的模型进行介绍。有关这部分的详细内容见三篇视觉显著性研究的综述性文献。