

基于机器视觉的目标跟踪算法研究

张熙鹏 陈恒

西京学院 陕西西安 710123

DOI:10.12238/ems.v7i7.14295

[摘要] 随着计算机视觉和人工智能技术的快速发展,基于机器视觉的目标跟踪算法得到了广泛关注。本文系统梳理了目标跟踪的基本理论、传统与深度学习方法,重点分析了关键技术如特征提取、目标定位及模型更新策略。同时,探讨了算法在跨场景适应性、实时性能、多目标跟踪和遮挡处理等方面的挑战。研究表明,深度学习与多模态融合技术将推动目标跟踪向更智能、更高效的方向发展,具有广泛的应用前景。

[关键词] 机器视觉;目标跟踪;跟踪算法

1. 引言

目标跟踪作为计算机视觉的重要研究方向,旨在从视频序列中实时定位并跟踪特定目标。随着机器视觉技术的进步,目标跟踪算法在智能监控、无人驾驶、机器人导航等领域发挥着关键作用。如何在保证高精度的同时实现轻量化和实时性,以及应对遮挡、多目标等挑战,仍是当前研究的热点问题。本文将系统分析现有目标跟踪技术,探索其关键技术与未来发展趋势。

2. 目标跟踪技术基础

2.1 目标跟踪的定义与基本流程

目标跟踪是机器视觉领域的一个核心任务,指的是在视频序列中持续定位和跟踪特定目标的位置和状态。其基本目标是在每一帧图像中准确检测并确定目标的位置,保证跟踪的连续性和稳定性。一般来说,目标跟踪的基本流程包括目标初始化、特征提取、目标检测与匹配、状态估计和模型更新几个关键步骤。首先,在视频的初始帧中对目标进行标注或检测,确定其位置。然后通过提取目标的视觉特征,结合当前帧图像信息进行目标识别和定位,利用匹配算法确定目标的新位置^[1]。接着,根据跟踪模型对目标状态进行估计,判断目标的运动轨迹。最后,动态更新目标模型以适应目标的外观变化和环境变化。这个循环过程持续进行,直至跟踪任务结束。

2.2 目标跟踪的性能指标

评估目标跟踪算法的性能,通常依赖于多个指标。准确率(Accuracy)是最基本的指标,反映跟踪算法在定位目标中心或边界框时的精度,常用重叠率(Intersection over Union, IoU)来量化。鲁棒性(Robustness)表示算法在复

杂场景下抗干扰和恢复能力,评估其在遮挡、快速运动、光照变化等情况下是否能保持跟踪不中断。实时性(Real-time Performance)是目标跟踪应用的重要要求,尤其在无人驾驶、智能监控等领域,算法需要在限定时间内完成计算,保证视频流的连续处理。跟踪的稳定性和连续性也是重要指标,避免出现目标漂移和丢失现象。

2.3 目标跟踪面临的主要挑战

目标跟踪任务在实际应用中面临诸多挑战,主要包括目标遮挡、光照变化、背景干扰、目标形变和快速运动等。遮挡是最常见的问题,部分或完全遮挡目标会导致跟踪失败。光照变化影响图像的颜色和纹理特征,给目标检测带来困难。复杂背景可能导致误匹配,特别是在目标与背景相似时,容易出现跟踪漂移。目标本身的形变和姿态变化增加了特征提取的难度,传统方法往往难以适应这些变化。快速运动则会导致目标在连续帧间位置变化剧烈,增加跟踪的不确定性。面对这些挑战,现代目标跟踪算法不断引入深度学习、多模态融合和自适应模型更新等技术,以提升算法的鲁棒性和准确性。

3. 基于机器视觉的目标跟踪算法分类

3.1 传统目标跟踪算法

3.1.1 基于滤波的方法

基于滤波的目标跟踪方法是传统目标跟踪领域的重要分支,主要通过对目标状态进行动态估计实现连续跟踪。卡尔曼滤波是一种线性高斯系统的最佳估计算法,适用于目标运动规律较为简单且噪声服从高斯分布的场景。它通过预测目标的下一状态并结合当前观测数据,实现对目标位置和速度的估计,具有计算效率高、实现简单的优点。但卡尔曼滤波对模型线性和噪声分布的要求较高,难以处理复杂的非线性

和非高斯问题。为此, 粒子滤波(又称序列蒙特卡洛方法)被引入跟踪任务。粒子滤波通过生成大量粒子样本对目标状态的后验分布进行近似, 适用于非线性、非高斯系统, 能够更好地应对遮挡和快速运动等复杂情况。

3.1.2 基于模板匹配的方法

基于模板匹配的目标跟踪方法通过在当前帧中搜索与目标模板最相似的区域实现目标定位。该方法核心思想是先在第一帧中选定目标模板, 然后在后续帧中利用相关系数、互相关匹配或像素差异等度量指标寻找与模板匹配度最高的位置。模板匹配方法简单直观, 计算成本较低, 适用于目标外观相对稳定且背景变化不大的场景。然而, 该方法对目标外观变化和遮挡较为敏感, 容易出现跟踪漂移和失效问题。为提升匹配的鲁棒性, 后续研究中引入了多模板更新策略、自适应模板调整以及多特征融合等技术, 以增强算法在复杂环境中的适应能力。

3.2 基于深度学习的目标跟踪算法

3.2.1 监督学习方法

基于深度学习的目标跟踪算法中, 监督学习方法是最常见的一类。该方法依赖于大量带有精确标注的训练数据, 通过构建深度神经网络学习目标特征表示和运动模式。典型的监督学习跟踪算法会采用卷积神经网络(CNN)提取目标的高维特征, 再通过分类器或回归器实现目标定位。监督学习的优势在于能够自动学习复杂的目标特征, 极大提升了跟踪的准确性和鲁棒性。例如, Siamese网络通过构建模板和搜索区域的相似性度量, 实现高效的匹配跟踪, 成为近年来的研究热点。

3.2.2 无监督/半监督学习方法

为解决监督学习对标注数据依赖较大的问题, 无监督和半监督学习方法逐渐成为研究新方向。无监督学习方法通过设计自监督任务或利用目标的内在结构特征, 减少对人工标注的依赖, 能够从大量未标注数据中学习有效的特征表示。半监督学习则结合少量标注数据与大量未标注数据, 通过标签传播、伪标签生成等技术提升模型的泛化能力^[2]。这类方法在处理多样化场景和目标时展现出较强的适应性和灵活性, 特别适用于数据匮乏或变化多端的应用环境。

3.2.3 端到端跟踪网络

端到端跟踪网络是将目标跟踪任务设计成一个整体的深度学习模型, 直接输入视频帧序列, 输出目标位置或边界框。

该类方法避免了传统分步骤设计中的信息丢失和误差传递问题, 实现了特征提取、目标检测和状态估计的联合优化。典型代表包括基于循环神经网络(RNN)、长短期记忆网络(LSTM)以及Transformer架构的跟踪模型, 能够捕捉目标的时序动态特征。端到端方法不仅提升了跟踪的准确性, 还增强了模型对目标外观变化和复杂运动的适应能力。

3.2.4 多目标跟踪技术

多目标跟踪(MOT)旨在同时跟踪视频中多个目标, 面临目标交互、遮挡和身份切换等复杂挑战。基于深度学习的多目标跟踪技术通过结合检测与关联方法, 实现对多个目标的准确识别与连续跟踪。常见方法包括“检测-关联”框架, 即先用深度检测网络定位所有目标, 再通过关联算法(如匈牙利算法、图神经网络)实现目标身份匹配。深度学习的强大特征表达能力为多目标跟踪提供了有力支撑, 广泛应用于视频监控、智能交通和机器人导航等领域。

4. 关键技术与算法改进

4.1 特征提取与表示方法

特征提取是目标跟踪算法中的基础环节, 其质量直接影响跟踪的准确性和鲁棒性。传统方法依赖手工设计的特征, 如颜色直方图、边缘、纹理等, 这些特征对环境变化的适应性较差。随着深度学习的发展, 基于卷积神经网络(CNN)的特征提取成为主流, 能够自动学习多层次、多尺度的目标表征。深度特征不仅具备更强的判别能力, 还能有效抑制背景干扰, 提高目标区分度。结合多模态特征融合(如颜色、形状和运动信息)的方法, 进一步增强了跟踪器对复杂场景的适应性。

4.2 目标检测与定位技术

目标检测与定位是实现精准跟踪的核心步骤。现代跟踪算法通常依赖目标检测模块快速识别目标区域, 定位目标位置。基于深度学习的检测器, 如Faster R-CNN、YOLO和SSD等, 通过端到端训练获得较高的检测精度和速度, 广泛应用于目标跟踪系统。目标定位不仅仅是粗略的框定目标, 还包括对目标边界的精细调整以应对目标大小和形状变化。为此, 许多算法引入回归方法精确拟合目标边界框, 提升定位的准确性。同时结合运动估计技术, 如光流法和卡尔曼滤波, 有效预测目标的动态位置, 提高跟踪的实时性和连续性。

4.3 跟踪模型的更新策略

跟踪模型的动态更新是保持跟踪性能的关键环节。由于

目标在运动过程中会出现外观变化、姿态变化甚至遮挡, 静态模型容易导致跟踪漂移或失败。有效的模型更新策略需要平衡适应新环境与防止错误累积之间的关系。当前常用的更新方法包括基于置信度的自适应更新、在线学习机制以及长短期记忆模型的应用。置信度较高时, 模型可适时引入新样本进行微调, 以捕捉目标最新特征; 反之则暂停更新, 避免引入噪声。部分算法还采用双模型结构, 主模型稳定跟踪, 辅助模型快速响应变化, 从而提升整体鲁棒性。

4.4 处理遮挡与目标丢失的方法

遮挡和目标丢失是目标跟踪中极具挑战性的问题, 严重影响跟踪效果。处理遮挡的方法主要包括部分遮挡检测、遮挡恢复和多模态信息融合。部分遮挡检测通过分析目标外观局部区域变化, 判断遮挡程度, 从而调整跟踪策略。遮挡恢复技术则利用运动预测、历史轨迹回溯和上下文信息, 帮助模型在目标重新出现时快速定位^[3]。近年来, 多摄像头、多传感器融合技术被引入, 通过多角度、多模态信息补充单一视觉数据不足, 显著提升遮挡情况下的跟踪鲁棒性。对于目标丢失, 现代跟踪算法多结合再检测机制, 一旦跟踪器置信度下降或目标消失, 即启动目标重新检测与匹配流程, 确保跟踪的连续性和稳定性。

5. 挑战与未来发展方向

5.1 跨场景跟踪的适应性提升

目标跟踪算法在实际应用中面临的一个重要挑战是如何在不同场景下保持稳定的性能。跨场景跟踪涉及目标在光照、背景、视角、尺度等多种条件变化下的适应能力。目前许多算法在特定环境下表现良好, 但迁移到新环境时往往效果明显下降。提升跨场景适应性, 需要研究更具泛化能力的特征提取方法和鲁棒的模型更新机制。比如, 利用域自适应技术, 通过减少源域和目标域之间的分布差异, 使得模型能在未见过的场景中依然有效。

5.2 轻量化与实时性改进

随着机器视觉应用场景的多样化, 尤其是在无人机、移动机器人和智能终端等资源受限设备上的应用, 轻量化和实时性成为目标跟踪算法必须考虑的重要因素。传统深度学习模型虽然准确率高, 但通常计算复杂度大、参数量多, 难以满足实时处理需求。为此, 研究者们致力于设计结构紧凑、计算高效的轻量级网络, 例如使用深度可分离卷积、模型剪枝和量化技术降低模型复杂度。同时, 借助硬件加速器如

GPU、FPGA 和专用 AI 芯片, 提升计算效率。实时跟踪不仅保证系统反应及时, 也极大增强了应用的实用性。

5.3 多模态融合与多目标跟踪

多模态融合和多目标跟踪是目标跟踪领域未来的重要发展方向。单一视觉信息常因遮挡、光照等问题而导致跟踪失败, 引入多模态数据(如红外、雷达、深度信息)能够有效补充视觉信息的不足, 提升跟踪的鲁棒性和准确性。同时, 多目标跟踪在智能监控、交通管理等场景中具有广泛应用。多目标跟踪不仅需要检测多个目标, 还需解决目标间交互、身份切换和遮挡等复杂问题。结合深度学习的多模态特征融合技术及图神经网络等关联算法, 为多目标的精确定位和连续跟踪提供了有力支持。

5.4 自监督学习与增强学习的应用前景

自监督学习和增强学习作为近年来机器学习领域的前沿技术, 在目标跟踪中的应用展现出巨大潜力。自监督学习通过设计无标签或少标签的学习任务, 利用大量未标注数据自动挖掘目标的内在特征, 解决了传统监督学习对大量标注数据依赖的问题。该方法能提升模型的泛化能力, 适应多变的实际环境。增强学习则通过与环境交互, 学习最优的跟踪策略和动作决策, 尤其在处理复杂动态场景和长时间跟踪任务中表现出优势。结合这两种技术, 可以构建更智能、自适应的目标跟踪系统, 实现从感知到决策的闭环优化。

6. 结论

本文全面综述了基于机器视觉的目标跟踪算法, 涵盖传统滤波、模板匹配及深度学习方法, 重点探讨了特征提取、目标检测与模型更新等关键技术。面对跨场景适应性、实时性能和多目标跟踪的挑战, 研究表明多模态融合和新兴的自监督、增强学习技术具有重要潜力。未来, 目标跟踪算法将朝着智能化、轻量化和高鲁棒性方向发展, 推动其在更多实际应用场景中实现突破。持续的算法创新和硬件支持, 将是提升目标跟踪性能的关键保障。

[参考文献]

- [1] 甘志英. 基于机器视觉的目标跟踪算法研究[J]. 工业技术与职业教育, 2024, 22(4): 10-13.
- [2] 叶聪. 基于机器视觉的目标检测与跟踪算法研究[J]. 通讯世界, 2025, 32(1): 178-180.
- [3] 周小勇. 基于机器视觉的运动目标跟踪算法优化研究[J]. 佳木斯大学学报(自然科学版), 2024, 42(2): 176-180.