

**论文答辩情况和学位授予决议书**

**答 辩 人**  周宗伟

**学 号**  201618014628053

**论文题目** 基于深度卷积网络的多目标跟踪

方法研究

**指导教师** 胡卫明 研究员

兴军亮 研究员

**学位类别** 工学博士

**学科(专业)** 模式识别与智能系统

**研究方向** 多目标跟踪

**培养单位** 中国科学院自动化研究所

**中国科学院大学制**

**填 表 说 明**

1. 本表内容须真实、完整、准确。
2. “学位类别”名称填写：哲学博士、教育学博士、理学博士、工学博士、农学博士、医学博士、管理学博士，哲学硕士、经济学硕士、法学硕士、教育学硕士、文学硕士、理学硕士、工学硕士、农学硕士、医学硕士、管理学硕士，工程硕士、工商管理硕士等。
3. “学科(专业)”名称：学术型学位填写“二级学科”全称，专业学位填写“培养领域”全称。
4. 本表“学位授予决议书”中：学术型学位填写“同意（或不同意）授予博士学位”、“同意（或不同意）授予硕士学位”。

**学位论文答辩记录**

|  |  |  |  |
| --- | --- | --- | --- |
| 学位论文题目 | 基于深度卷积网络的多目标跟踪方法研究 | | |
| 答辩日期 | 2021年03月 04日 | 答辩地点 | 中国科学院自动化研究所  智能化大厦三层第五会议室 |
| 用简明语言记录答辩过程，答辩委员会提出问题和研究生回答情况：  (赵耀教授)   1. 论文研究内容是多目标跟踪，那么像火车站这种密集场景中是否会遇到遮挡和交互问题？如何解决？   答：目标交互和遮挡是多目标跟踪问题的难点。处理这类问题一般有两个思路。其一是提取目标可见部分的表观，由于邻近帧中相同目标的可见部分变化缓慢，因此这部分特征有较好的鉴别性，一定程度克服轨迹编号交换；其二是利用时序信息，当目标处于遮挡情形时，该轨迹停止跟踪，在交互过程完成之后，利用表观关联找回跟丢目标。   1. 用可见部分处理遮挡的思路是对的，但怎么判断是否遮挡呢？   答：这种情况一般利用检测框的密度来判断，若轨迹的预测框与多个检测框都具有较大的重叠，那么认为轨迹在此刻可能处于遮挡，然后采用遮挡模块进行处理。也有一些方法通过判断目标的深度信息来区别被遮挡目标。  （陆耀教授）   1. 高阶图匹配方法中，是将图送入图网络中处理吗？这样做效果是否会更好？   答：在我们采用的高阶图匹配模型中，仅利用了深度网络提取可见部分的表观特征，而运动信息和结构信息依然是手工特征，并没有利用图网络处理。理论而言，采用图网络能够取得更好的特征融合，会获得更好的效果。   1. 图5.2需要添加适当的标注。   答：已添加引用。  （张文生研究员）   1. 在干扰感知鉴别学习模型中，如何选择困难和简单样本呢？这种阈值是否具有普适性？   答：在该模型中，我们利用锚框与标签之间的交并比筛选正负样本，而其阈值始终超参，往往需要针对于不同的任务进行人工调节，比如在多目标跟踪数据库中经过多次实验分析，我们选择上下阈值分别于0.6和0.3。后续工作可以在普适性上展开进一步的深入探讨。  （王以政研究员）   1. 该任务中的多目标是如何定义的？提出的算法有在实际场景中实验过吗？   答：在多目标跟踪任务中，目标一般指独立的个体，比如行人和车辆等，此时如果需要引入额外的约束，比如对小孩的跟踪，那么就需要对分类器添加额外的约束，但特征提取和数据关联阶段的方法是通用的。本文提出的算法所使用的评测数据库本身就采自于真实场景，但对于非常复杂的应用场景，算法暂时没有进行过尝试。  （王文成研究员）   1. 目标跟踪中的目标框如何获得？对于尺度差异较大目标，如何大人和小孩需要分别处理吗？目标运动速度不同该怎么处理？能够处理表观非常相似的目标吗？   答：多目标跟踪任务中的目标框一般是由检测器检测获得，针对于不同尺寸的目标其检测得到的目标框大小也不同。对于不同尺寸的目标框，最简单的方式是利用目标框的交并比进行区分和关联，复杂情况下，是对目标进行缩放使其尺度相同利用表观抽取网络提取表观特征，该特征用来区分目标。而对于运动速度不同的目标，一般会扩大待搜索区域进行表观关联，但此时如果存在多个表观非常相似的目标就无法处理。  （李兵研究员）   1. 在高阶图匹配模型中，目标的三种特征是融合到一个网络中端到端处理的吗？   答：高阶图匹配模型中目标的特征包括表观特征、运动一致性和空间位置关系，这里只有目标可见部分的表观特征是利用深度网络提取的，运动一致性和空间位置关系则基于规则设计，而最终的优化问题始终离散的组合优化，因此不存在端到端处理的做法。但如果利用图卷积网络刻画时空图，那么整个系统是能够实现端到端处理的。   1. 四个工作是否能统一在一起形成体系？   答：四个工作中，第一个工作室针对于数据关联，其余的三个工作针对于目标检测和特征抽取，这三个工作是逐步推进的。理论上第一个工作提出的数据关联方法可以应用到后三个工作中，但是从跟踪速度上考虑，第一个工作提出的数据关联方法速度较慢，难以应用到实际场景，因此将该方法结合不同的目标检测和特征抽取没有实际的应用意义。但如果采用图卷积网络重新处理第一个工作中的图匹配模型，应该可以和后三个工作进行统一，得到较好的跟踪性能。  答辩委员会记录人（签字）： | | | |

注：可另附页

**答辩委员会决议（限1200字）**

|  |
| --- |
| 多目标跟踪是模式识别和计算机视觉领域的重要研究课题之一，在智能监控、自动驾驶和体育分析等领域有着广泛的应用前景，论文选题具有重要的理论意义和应用价值。  论文取得如下创新性研究成果：  1. 提出了一种高阶图匹配模型，以目标可见部分的表观特征、轨迹的运动信息和空间结构信息构建高阶张量，并通过张量幂迭代算法获得更加稳定的跟踪轨迹，提高了帧间关联中的特征表示能力。  2. 提出了一种基于干扰感知鉴别学习神经网络的多目标跟踪方法，通过逐步利用表观相似度、空间交并比进行数据关联，既能召回漏检目标，又能提升检测率。  3. 提出了一种基于长短时线索联合抽取网络的多目标在线跟踪方法，能够同时进行目标的局部检测和表观特征提取，并能利用表观特征建立目标关联从而找回跟丢目标，改善了全局检测结果。  4. 提出了一种在单个阶段中同时完成目标全局检测和表观特征提取的联合神经网络，提升了检测精度和跟踪速度。  在国际公开数据库上验证了上述方法的有效性。  论文工作表明，作者在本学科领域掌握了坚实宽广的基础理论和系统深入的专门知识，独立从事科研工作的能力强。论文写作规范、条理清晰、逻辑性强、实验结果翔实。答辩过程中，思路清晰，回答问题正确。答辩委员会经无记名投票，一致同意周宗伟同学通过博士论文答辩，并建议授予周宗伟工学博士学位。  答辩委员会主席（签字）：  2021 年 03 月 04 日 |

**答辩委员会组成**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 答辩委员会成员 | 姓 名 | 专业技术职 务 | 是否博导 | 研究领域 | 工作单位 | 签 字 |
| 主 席 | 赵 耀 | 教 授 | 是 | 信息取证、信息隐藏、计算机视觉、人工智能 | 北京交通大学 |  |
| 委 员 | 王以政 | 研究员 | 是 | 神经信号转导研究 | 中科院神经科学研究所 |  |
| 陆 耀 | 教 授 | 是 | 计算机视觉、模式识别与跟踪 | 北京理工大学 |  |
| 王文成 | 研究员 | 是 | 计算机图形学 | 中科院软件研究所 |  |
| 张文生 | 研究员 | 是 | 人工智能理论与方法、社会计算与大数据解析 | 中科院自动化研究所 |  |
| 李 兵 | 研究员 | 是 | 图像与视频处理 | 中科院自动化研究所 |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| 表决结果 | 记录 | 投票 6 人， 其中同意 6 票，不同意 0 票, 弃权 0 票  综合评价意见：优秀 6 票，良好 0 票，中 0 票，差 0 票 | | | | |
| 结论 | √ 通过 □修改论文重新答辩 □ 不通过 | | | | |
| 答辩委员会秘书（签字）： 答辩委员会主席（签字）：  2021 年 3 月 4 日 | | | | | |