一是往轻量级,小型化方向发展,让 SLAM 能够在嵌入式或手机等小型设备上良好的运行,然后考虑以它为底层 功能的应用,毕竟大部分场合中,我们的真正目的暴展实现机器人、AR/VR 设备的功能,比如设运动、导航、教 学、娱乐,而 SLAM 是为上层面用推销自身的一个位姿估计。在这些应用中,我们不希望 SLAM 占据所有计算资 源,所以对 SLAM 的小型化和轻量化有非常强烈的要求。 发展趋势 另一个方面,则是利用高性能计算设备,实现精密的三维重建、场景理解等功能。在这些应用中,我们的目的是 完美地重建场景,而对于计算资源和设备的便携性则没有多大限制,由于可以利用 GPU、这个方向和深度学习亦 实际的 机器人也好,硬件设备也好,通常都不会只携带一种传感器,往往是多种传感器的融合。 学 水界的研究人员喜爱"大而且干净的问题"(Big Clean Problem),比如说仅用单个摄像头实 但产业界的朋友们则更注重于让算法更加实用,不得不面对一些复杂而 且琐碎的场景。在这种应用背景下,用视觉与惯导融合进行 SLAM 成为了一个关注热点。 背景 惯性传感器(IMU)能够测量传感器本体的角速度和加速度,被认为与相机传感器具 有明显的互补性,而且十分有潜力在融合之后得到更完善的 SLAM 系统 1.IMU 可以测得角速度和加速度 2.IMU 虽然可以测得角速度和加速度,但这些量都存在明显的漂移(Drift),使得积分两次得到的位姿数 据非常不可靠。好比说,我们将 MU 放在桌上不动,用它的读数积分得到的位姿也会漂出十万八千里。但 是,对于照时间内的协建运动。IMU 能够提供一些较好的估计。这正是相机的弱点。 当运动过快时,任帝珍仁门的 相似会出现运动模糊,或者商帧之间重极区域太少以至于无法进行特征匹配,所以纯税党 SLAM 非常害怕快速的运动。而有了 IMU,即使在相机数据无效的那段时间内,我们还能保持一个较好的位姿估计,这是纯视觉 SLAM 无法做到的 视觉 + 惯导 SLAM 惯性传感器(IMU)的特点: 第9章: SLAM的未来及展望 3.相比于IMU,相机数据基本不会有漂移。如果相机放在原地固定不动,那么(在静态场景下)视觉 SLAM的位姿估计也是固定不动的。所以,相机数据可以有效地估计并修正IMU读数中的漂移,使得在慢 谏运动后的位姿估计依然有效。 4. 当图像发生变化时,本质上我们没法知道是相机自身发生了运动,还是外界条件发生了变化,所以纯视觉 SLAM 难以处理动态的障碍物。而IMU 能够感受到自己的运动信息,从某种程度上减轻动态物体的影响。 松耦合是指,IMU 和相机分别进行自身的运动估计,然后对它们的位姿估计结果进行融合。 目前 VIO 的框架的分支: 緊耦合是指,把 IMU 的状态与相机的状态合并在一起,共同构建运动方程和观测方程,然后进行状态估计 紧耦合 (Tightly Coupled) 紧耦合理论分为基于滤波和基于优化的两个方向。 1. 语义帮助 SLAM 语义 SLAM 物体识别和分割都需要大量的训练数据。要让分类器识别各个角度的物体,需要从不同视角采集该 物体的数据。然后进行人工标定,非常辛苦。而SLAM中,由于我们可以估计相机的运动,可以自 动地计算物体在图像中的位置,节省人工标志的成本。如果有自动生成的带高质量标注的样本数 2. SLAM 帮助语义。 据、能够很大程度上加速分类器的训练过程。 作者: 龙笑泽 课程: 深蓝学院《视觉SLAM理论与实践》 书籍: 《视觉SLAM十四讲》高翔老师 参考: 证明参考: 见证明部分,已标注参考链接 制作软件: xmind 2020版本(Mac系统) 制作时间: 第一版: 2021.06-2021.09 作者是深蓋学院《視觉SLAM理论与实践》第十期(课程时间: 2021.05-2021.08)的学员,第一版思维导图制作完成时大四刚开学,自动化专业,西安某211院校。 2019.08: 开始学习机器学习 2019.11: 深度学习 学习历程: 2020.01: 计算机视觉领域 2021.01: 三维点云处理 思维导图制作相关信息 最开始了解AI是在专业课上,然后就根据专业课字老师的意见开始学习机器学习,到后来进入深度学习,正達2020年计算机视觉在深度学习,尤其是卷积神经网络的加持下发展飞快,就进入了这一个领域,自先是图像分类领域,然后进入了目标检测领域,2020年初,当时火热的有有satron系列,YOLO系列,等等,后来也了解过一些自然语言处理的相关知识,不过主要学习还是在目标检测领域。 作者简介: 到2020年底,明显感觉计算机视觉发展有些缓慢了,不是说效果不好,就是感觉没有飞跃式的进步了,正好我那时候大三,专业课比较多,学习压力也有些大,就像尔听一听公开课,不过,在这个阶段,由于知识面扩充了,而且自己思考的也足够多了,就有了一些自己对于计算机视觉的 事实上、机器学习和深度学习费认为就是干了一件很简单但是很复杂的事情——拟合函数。机器 学习的两大任务:回过和分类,主要用来预测。深度学习是机器学习的一部分,只不过、传统机 器学习基于数据的话者很大一部分是在人工燃持征构建,也叫特征工程,而欢度学习可以使用其 复杂的网络结构保护的理时待。 无论是传统的多层感动神经网络还是CNA、甚至是RNA都可 以很好的进行对不同数据格式的特征提取,保的测是时间序列或者文本序列的特征提取,甚至是 CNN可以用来进行图像的特征提取,RNN则是时间序列或者文本序列的特征提取,甚至是 Transformer结构,原本用在自然语言处理的结构,在2021年初中在图像领域大放异彩。不过在 报看来,回到综合函数这个角度,当模型足够复杂,就会产生量变到质变的过程,因此 学习感悟: (仅供参考) ner结构用在图像,CNN用在文本也就可以理解。 在经历了深度学习的计算机视觉的学习过后,我也在思考计算机视觉未来的发展方向,事实上,视觉属于计算机感知的一部分,只不过这里的传感器胜相机,数据是图像罢了。但是下一个引领社会的发展风口一定是在3000党。可以看到的是,2020年和2021年大热的自动驾驶,其内部结构很复杂。但是能发现一切控制的基础。自动等地分发展,在对土佔重要的战是计算机观觉域的进步。甚至AR、VR的发展,有了深度学习的加持。很大程度上会成为下一个发展风口;再有2021年八月大热的元宇宙的概念,这些都高不开307概划的发展。只不过我认为的发展方向是先AR,再VR。(在这里,我不是说别的组成不重要,社会的发展离不开方方面面,请读者坚信自己的选择,我坚信每个人都会有美好的未来!) 因此,在2021年,我就进入3D视觉的学习中,学习了三维点云处理和这里的视觉SLAM,目前也在不断努力学习,希望未来可以为这个领域的发展做出自己的一点价值。 第一版作为全书的第一次系统总结,所以比较详细,重点部分没有拿颜色区分,不过引用 前文公式或者内容都已拿箭头标出,部分公式我使用typora软件使用latex语法书写,没有 加粗,希望读者自行判断。 作业以及代码实践部分没有添加,后续有可能以PDF或者markdown的形式给出。 由于大部分内容来自于高翔老师的《视觉SLAM十四讲》以及深蓝学院高翔老师的《视觉

slam理论与实践》课程,所以为了保护版权,在这里仅分享xr 源xmind文件会发给高翔老师。

由于能力有限,有些错误在所难免,可以和我联系或者在GitHub的issue部分提出,我进行 更改再重新导出上传到GitHub上。总结不易,如果这份资料对您有帮助,希望您在GitHub

备注:

视觉SLAM理论与实践