DF-SLAM：基于深度描述子的增强视觉SLAM系统

Abstract：

SLAM算法发展很快，但是传统SLAM系统的非几何模块受限于数据关联，变成了SLAM的发展瓶颈。为了解决中问题，许多研究转入深度学习。但是现在的很多研究受限于视觉数据集或者特定环境，也为了精度牺牲了效率。所以这些方法不够有实践性。

我们提出的DF-SLAM用了深度局部特征描述子来替换hand-made描述子。实验表明提升了效率和稳定性。也在很多场景outperforms了受欢迎的传统SLAM，包含光照变化剧烈的场景。他的多面性和可移植性契合探索新环境的需要。因为我们采用了一个浅层网络提取描述子，所以也能在GPU上达到实时。

Introduction

传统的SLAM基于多视图几何取得了巨大的成果，但是非几何模块限制了性能。跟踪线程中像素级水平的匹配和位姿优化中的轨迹估计漂移会累积。

（深度学习适合图片问题）同时，数据驱动的深度学习在图像分类和匹配方面带来了巨大的发展，表明其可以解决数据关联问题。越来越多的研究者相信像素级或者图片级关联能解决SLAM系统的瓶颈问题。

（深度学习优越性，但是有局限性）深度学习证明了其在SLAM系统中的优越性。很多替换传统SLAM中的pipeline中的一块。有的探索更高级的特征来获取语义信息。也有一些端到端的深度学习SLAM。这些在特定场景下都表明了深度学习在SLAM中的潜力。

（）但是这些方法都有局限性。很多深度学习方法依赖数据训练，不能适应未知的环境，比如语义SLAM方法我们不能保证有桌子椅子。还有的舍弃了传统SLAM的优点，牺牲了效率。

最重要的是有的深度SLAM方法将SLAM系统框架更改太多，设计了一些SLAM中的有用的特征，让研究者很难与现有的方法进行对比，更不用说移植到其他SLAM系统上。因此，基于深度学习的SLAM不足以outperform传统的SLAM。

因此我们提出了一种简单、可移植和高效的SLAM系统。主要想法就是通过深度学习来提升局部特征描述子的鲁棒性，保证帧间数据关联的精度。

在这篇文章中，我们用自己的描述子替换了传统SLAM中的描述子。主要的优势在于portability和convenience来替换传统的SLAM描述子。所以可以很好地移植到其他的几何计算机视觉任务例如SFM，相机标定等等。学习的局部特征描述子可以保证比传统描述子好。因为我们采用的浅层网络，所以可以GPU实时。

Related Work

深度学习增强SLAM

最近的端到端的SLAM效果没有传统的好，并且不适应新环境。、

所以有的研究者专注于传统SLAM系统的部分模块替换。

还有的给传统SLAM加辅助模块。比如加语义信息。开始有的将语义信息和传统的SLAM分开优化。不过现在也有一起优化的。但是计算量会比较大。

一个简单提升SLAM系统的方法就是提升一些模块，比如帧间匹配，所以有些研究专注于局部描述子的相似度计算，但是导致一些传统匹配策略的失效，比如欧式距离、cosine距离等。Superpoint训练了端到端网络提取local feature detectors and descriptors。但是其效率未被评估，因为它仅给出了合成和数据集的结果，没有继承到SLAM系统。

Local feature descriptor

传统的SIFT，SURF，ORB，也有一些深度学习的PCA-SIFT，ASD。。。有的致力于将高维信息映射到低维空间。也有的使用二进制描述符，但他们有的在特定环境增强了传统feature的性能，但缺少可移植性。许多也只是提出了新的feature，没有作进一步的测试或应用。

针对生成局部描述子，也有很多机遇CNN的结构如Siamese 和triplet网络。最终也成为局部描述子的主流框架结构。但是这些model都证明不适用与传统的最近邻搜索。

也有一些新的方法提出。

System Overview

采用传统主流的SLAM系统作为基础，来评估描述子的效率和精度。

3.1系统框架

讲述了ORB-SLAM2的框架结构和运行方法。

首先提取描述子

采用DBOW词袋做回环检测。

为了加速系统也引入词袋。

3.2 feature design

讲一下各个损失函数和训练策略

4 Experiments