三维图像传感器信号增强网络的**轻量化**方法

研究目的：针对**ToF传感器**采集的深度图或三维点云，进行成像质量提升，并同时优化模型推理效率，满足模型轻量化需求，从而适配机器人、手机、自动驾驶等移动端应用。

研究内容：包括三维图像传感器数据分析，建立有效的神经网络进行画质提升，对网络进行轻量化设计，进行实际传感器数据的画质及计算复杂度评估等。

国内外研究现状：国内外已有大量关于深度图及三维点云画质提升的研究，**但已有基于深度学习的方法往往忽略网络性能优化，**和传感器小型化背道而驰，亟需提出三维图像增强网络的轻量化方法。

创新点：面向三维图像传感器，基于其**成像原理**进行三维图像增强，创新地提出基于**动态稀疏卷积核（这是我瞎写的，后续方案再议）**的图像增强网络轻量化方法。

Related works

ToF原始数据需要通过深度重建的过程，获得深度图或三维点云的输出，其中包括深度值解算、自动曝光控制、噪声去除、空洞修复等功能模块。

噪声在深度重建过程中被非线性放大，增大处理难度，所以目前越来越多的研究工作从数据的源头开始，针对原始数据进行处理，有助于更好地恢复真实深度。

面向ToF原始数据的复原算法研究目前还处于较为初级的阶段，从研究思路上可分为基于模型和基于深度学习两类方法，具体如下。

1. **基于模型的原始数据处理 （简要了解，此为早期方法）**

这类方法基于经典统计学模型，通过图像结构先验知识设计优化算法。首先，libfreenect2是基于Kinect v2开发的原始ToF信号处理软件，整体处理框架被广泛使用，其中复原模块基于双边滤波的去噪模块，可实现实时处理[1]。（*也是目前项目里在使用的框架*）

Lenzen等人基于实验，对比了去噪模块在处理流程中位置和去噪方法的不同对结果锐利度、平滑度的影响，其中在原始自相关数据上使用双边滤波并加以中值滤波获得了最优的效果[2]。

Frank等人对深度重建过程进行建模，得到深度图噪声模型为偏移正态分布，并且噪声方差反比与信号幅值的平方，可以依据信号幅值推测噪声分布[3]；

基于此噪声模型，Reynolds利用信号幅值作为置信度，作为均值滤波的系数，从而避免低置信度的像素对滤波结果造成影响[4]。

总结：该类算法以先验知识为基础，可解释性强、可灵活调节，且**功耗低实时性强**；但使用的方法对ToF传感器缺乏针对性，**精度有限**。

1. **基于深度学习的原始数据处理 (此为重点，是最新方法)**

近年来，深度学习的引入大大提升了三维成像质量，取得了令人鼓舞的成果。

Su等人利用端到端的深度神经网络整体替换模块化3D ISP，对ToF原始信号进行去噪和多径干扰纠正，输出经过复原的深度图[5]。

Guo等人建立了ToF的仿真数据集，并利用该数据集进行分阶段的深度神经网络训练，进行去噪、多径干扰纠正和运动模糊去除[6]。

Dong 等人通过考虑不同尺度下的场景全局几何信息，预测不同尺度下的深度残差并进行融合，提升去噪精度[7]。

另外，同时利用原始数据和深度图、信号强度图作为输入进行多次迭代[8]和表面法向量联合预测[9]，进行精细处理。

Agresti等人关注由于网络在仿真数据上训练，真实场景泛化能力不足，所以通过生成对抗网络在真实数据上进行自监督提升泛化性[10]。

总结：这类方法在特定场景下精度提升显著，算法设计尤其是深度神经网络忽略复杂度，导致功耗高，传感器小型化困难。

**局限性分析**

虽然对原始数据进行处理可以从源头上抑制错误的发生，但目前的研究依然较为初步，还存在以下问题。

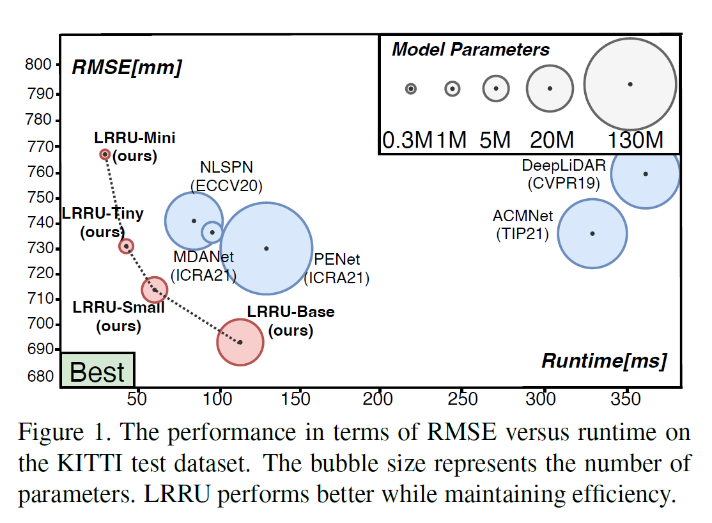
• 准确度：缺乏针对成像过程的退化过程分析，导致算法准确度不足；

• 泛化能力：基于深度学习的方法虽然在精度上有所提升，但在数据集之外的场景中泛化能力不足，缺乏表征能力和可解释性的平衡；

• 功耗：算法设计尤其是深度神经网络忽略复杂度，导致功耗高，传感器小型化困难。

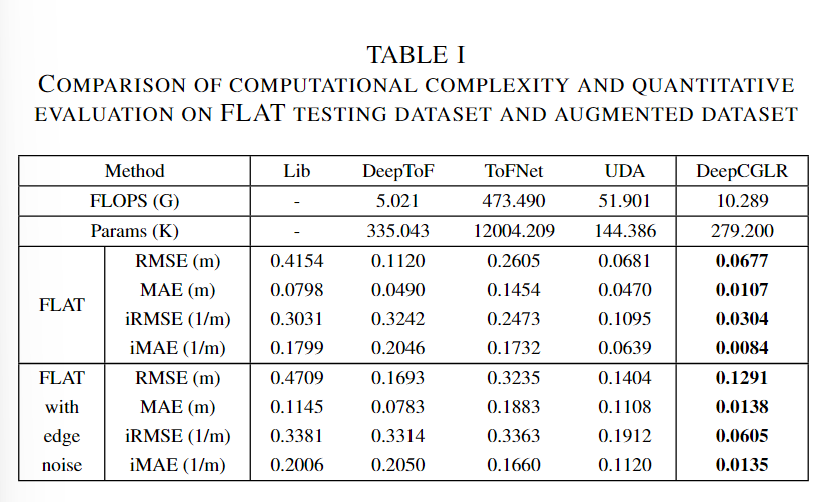
因此需要做到表征能力和可解释性的平衡（已解决，在4.1中），以及准确度和功耗的平衡。

1. **三维图像传感器信号增强网络的轻量化 (此为重点，直接相关的较少)**

最新的方法开始关注到三维图像增强网络的轻量化，在保证网络准确度的同时降低算法功耗。例如在深度补全任务中（相关任务，但并非我们的任务），LRRU[11] 通过spatially-variant kernels对其进行迭代更新，迭代更新过程是内容自适应且高度灵活的，核范围（kernel scope）动态调整以捕获长到短范围的依赖关系（long-to-short range dependencies），有效地用较少的可学习参数和推理时间将其优化为准确的深度地图。

然而，针对ToF传感器原始数据处理的网络对网络加速/轻量化还鲜有研究，因此亟需提出三维图像增强网络的轻量化方法，推动ToF传感器在实际应用中的使用。

1. 已有基础（此为本项目开发基础，为前两届同学大创成果，细节内容后续与大家同步）

（1）基于ToF成像原理，提出confidence-aware **graph** Laplacican regularization（CGLR）作为信号先验，与深度神经网络的结合得到DeepCGLR网络。一方面利用DNN增强信号先验的准确性，同时通过信号先验限制DNN的解空间，提升模型泛化能力。结果如下表所示。

（2）网络轻量化也在其他领域引起了广泛关注。例如传统GNN在处理大规模图数据时面临计算资源的巨大挑战，因此需要对GNN进行轻量化。虽然现有的图彩票 (Graph lottery ticket, GLT) 方法可以通过同时剪枝网络参数和图结构来缓解这一问题，但它们通常依赖于手动的剪枝率调整和不具弹性的剪枝范式，不适用于深层图神经网络。

面向图神经网络（GNN）的图结构及参数稀疏化，我们提出了一种自适应的图神经网络联合稀疏框架（简称AdaGLT），从而进行网络轻量化。该方法通过以下几个方面的创新来解决现有技术的局限性：1) 针对不同数据集和GNN的层自适应稀疏结构，以支持更深层的GNN；2) 将剪枝和训练过程整合，实现动态工作流程，包括剪枝和恢复；3) 自动识别不同稀疏级别的图彩票，无需广泛的剪枝参数调整。我们还理论上证明了AdaGLT能够在深层GNN场景中缓解过平滑问题，并获得改进的稀疏结构。结果展示的是可获得的最大图稀疏度或权重稀疏度。

|  |  |  |  |
| --- | --- | --- | --- |
| GCN上对邻接矩阵进行稀疏 | | | |
| 方法 | Cora | Citeseer | Pubmed |
| 随机剪枝 | 2.13% | 6.22% | 8.26% |
| UGS | 9.84% | 18.17% | 47.10% |
| RGLT | 17.52% | 20.02% | 52.13% |
| 本方案 | 28.90% | 42.66% | 65.77% |

|  |  |  |  |
| --- | --- | --- | --- |
| GCN上对权重进行稀疏 | | | |
| 方法 | Cora | Citeseer | Pubmed |
| 随机剪枝 | 11.04% | 24.65% | 10.45% |
| UGS | 32.65% | 69.21% | 91.95% |
| RGLT | 68.55% | 70.34% | 92.38% |
| 本方案 | 95.37% | 92.64% | 99.17% |

References（标红的为重点关注）

1. Lingzhu Xiang, et.al. libfreenect2: release 0.2. Zenodo, 2016.
2. Frank Lenzen, Kwang In Kim, Henrik Schäfer, Rahul Nair, Stephan Meister, Florian Becker, Christoph S. Garbe, and Christian Theobalt. Denoising strategies for Time-of-Flight data. In Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications, pp. 25-45. 2013.
3. Mario Frank, Matthias Plaue, Holger Rapp, Ullrich Köthe, Bernd Jähne, and Fred A. Hamprecht. Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. Optical Engineering 48, no. 1: 013602. 2009.
4. Malcolm Reynolds, Jozef Doboš, Leto Peel, Tim Weyrich, and Gabriel J. Brostow. Capturing Time-of-Flight data with confidence. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 945-952. 2011.
5. Shuochen Su, Felix Heide, Gordon Wetzstein, and Wolfgang Heidrich. Deep end-to-end Time-of-Flight imaging. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6383-6392. 2018.
6. Qi Guo, Iuri Frosio, Orazio Gallo, Todd Zickler, and Jan Kautz. Tackling 3D ToF artifacts through learning and the FLAT dataset. In Proceedings of the European Conference on Computer Vision (ECCV), pp. 368-383. 2018.
7. Guanting Dong, Yueyi Zhang, and Zhiwei Xiong. Spatial hierarchy aware residual pyramid network for Time-of-Flight depth denoising. In European Conference on Computer Vision, pp. 35-50. 2020.
8. Zhuolin Zheng, Yinzhang Ding, Xiaotian Tang, Yu Cai, Dongxiao Li, Ming Zhang, Hongyang Xie, and Xuanfu Li. Iterative error removal for Time-of-Flight depth imaging. In International Conference on Artificial Neural Networks, pp. 92-105. 2021.
9. Rongrong Gao, Na Fan, Changlin Li, Wentao Liu, and Qifeng Chen. Joint depth and normal estimation from real-world Time-of-Flight raw data. In 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 71-78, 2021.
10. Gianluca Agresti, Henrik Schafer, Piergiorgio Sartor, Yalcin Incesu, and Pietro Zanuttigh. Unsupervised domain adaptation of deep networks for ToF Depth refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
11. Wang, Yufei, Bo Li, Ge Zhang, Qi Liu, Tao Gao, and Yuchao Dai. "LRRU: Long-short Range Recurrent Updating Networks for Depth Completion." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9422-9432. 2023.