



同济大学大学生创新训练项目计划申请书

项目编号	X2024492		
项目名称	三维图像传感器信号增强网络的轻量化方法		
项目负责人	林继申	联系电话	15143305542
所在学院	软件学院		
学号	2250758	专业班级	软件工程
指导教师	曾进		
E-mail	2250758@tongji.edu.cn		
申请日期	2023 年 12 月 13 日		
项目期限	一年期		

同济大学 创新创业学院

填写说明

1. 本申请书所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要。
2. 申请人可以是个人，也可为创新团队，首页只填负责人。“项目编号”一栏不填。
3. 本申请书为大 16 开本（A4），左侧装订成册。可网上下载、自行复印或加页，但格式、内容、大小均须与原件一致。
4. 负责人所在学院认真审核，经初评和答辩，签署意见后，将申请书（一式两份）报送同济大学项目管理办公室。

一、基本情况

项目名称	三维图像传感器信号增强网络的轻量化方法						
所属学科	学科一级门：工学 学科二级类：计算机类						
项目来源	<input type="radio"/> A、学生自主选题，来源于自己对课题的长期积累与兴趣 <input checked="" type="radio"/> B、导师推荐课题 <input type="radio"/> C、SITP 直升课题						
项目申请级别	国家级						
申请金额	10000.00 元	项目期限			一年期		
负责人	林继申	性别	男	民族	汉族	出生年月	2004 年 3 月
学号	2250758	联系电话	宅：15143305542 手机：15143305542				
指导教师	曾进	联系电话	宅：13265807412 手机：13265807412				
项目简介	<p>研究目的：</p> <p>三维成像与传感技术是获取物理世界几何信息的重要手段，为实现高精度的重建、识别、定位奠定了基础。在众多三维传感技术中，基于飞行时间法（Time-of-Flight, ToF）的三维图像传感器得益于其成本优势和优越的实时三维感测能力，是目前主流的三维成像方式。</p> <p>然而在已有研究中缺少平衡功耗和精度的三维成像硬件加速方案，无法满足传感器小型化需求。物联网、智能无人系统对传感器的小型化有着迫切需求，然而目前大量最新的三维图像处理算法，尤其是基于深度神经网络的算法，虽然能够达到较高的精度，但在设计过程中忽略算法复杂度，难以提供实时输出；抑或不满足硬件实现时的存储限制，导致无法部署在边缘设备运算芯片。因此，迫切需要为边缘设备设计低功耗高性能的三维图像处理算法，在算力有限的条件下保持算法输出精度，为“感算一体”三维图像传感器提供解决方案。</p>						



	<p>本研究旨在提升由 ToF 传感器采集的深度图和三维点云的成像质量，并致力于优化模型的推理效率，以满足轻量化的需求，使其适用于机器人、智能手机、自动驾驶等各种移动端应用。研究重点包括减少模型的计算复杂性，开发轻量级的算法和模型结构，并保证模型在不同应用场景下的适应性和泛化能力，这项工作的成功将大大推动智能设备在三维空间感知和理解方面的发展。</p> <p>研究内容：</p> <p>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。</p> <p>国内外研究现状：</p> <p>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。</p> <p>创新点：</p> <p>内核范围的动态调整：对内核范围进行动态调整，允许网络高效捕捉从长距离到短距离的依赖关系。实施类似策略预计可以使网络更有效地处理不同细节的尺度，改善输出质量，提高模型性能。</p> <p>改善模型大小和复杂度：减少层数和参数数量，同时通过智能设计（如使用更少但更有效的迭代）维持性能，预计可以实现适合移动或嵌入式应用的轻量级模型。</p>
负责人曾经参与科研的情况	项目负责人曾参加同济大学创新创业训练计划项目 2023 年导师发布课题（定向 2022 级本科生）。
指导教师承担科研课题情况	主持国家自然科学基金青年科学基金、上海市启明星项目（扬帆专项）
指导教师对本项目的支持情况	指导教师现任同济大学软件学院助理教授，香港科技大学博士，毕业后曾在商汤科技任高级研究员。主要研究方向为 3D 视觉，图信号处理，共发表 20 篇国际顶级期刊及会议，包括 IEEE TIP, IEEE TSP, CVPR, ECCV 等。



项目组 主要成员	姓名	学号	专业班级	所在学院	项目分工
	林继申	2250758	软件工程	软件学院	项目统筹和任务分工 文献调研 数据处理 模型搭建
	刘垚	2253215	软件工程	软件学院	文献调研 数据处理 论文撰写
	刘淑仪	2251730	软件工程	软件学院	文献调研 数据处理 经费管理
	梁斯凯	2253540	计算机科学与技术	电子与信息 工程学院	文献调研 数据处理 算法设计
	杨宇琨	2252843	软件工程	软件学院	文献调研 数据处理 答辩演讲

二、立项依据

(1) 研究目的

三维成像与传感技术是获取物理世界几何信息的重要手段，为实现高精度的重建、识别、定位奠定了基础。在众多三维传感技术中，基于飞行时间法（Time-of-Flight, ToF）的三维图像传感器得益于其成本优势和优越的实时三维感测能力，是目前主流的三维成像方式。

然而在已有研究中缺少平衡功耗和精度的三维成像硬件加速方案，无法满足传感器小型化需求。物联网、智能无人系统对传感器的小型化有着迫切需求，然而目前大量最新的三维图像处理算法，尤其是基于深度神经网络的算法，虽然能够达到较高的精度，但在设计过程中忽略算法复杂度，难以提供实时输出；抑或不满足硬件实现时的存储限制，导致无法部署在边缘设备运算芯片^{[12][13]}。因此，迫切需要为边缘设备

设计低功耗高性能的三维图像处理算法，在算力有限的条件下保持算法输出精度，为“感算一体”三维图像传感器提供解决方案。

本研究旨在提升由 ToF 传感器采集的深度图和三维点云的成像质量，并致力于优化模型的推理效率，以满足轻量化的需求，使其适用于机器人、智能手机、自动驾驶等各种移动端应用。研究重点包括减少模型的计算复杂性，开发轻量级的算法和模型结构，并保证模型在不同应用场景下的适应性和泛化能力，这项工作的成功将大大推动智能设备在三维空间感知和理解方面的发展。

(2) 研究内容

围绕“三维图像传感器信号增强网络的轻量化方法”这一课题，本项目的研究内容主要分为以下四个方面：

1. 三维图像传感器数据分析

这一阶段聚焦于深入分析 ToF 传感器采集的深度图和三维点云数据，识别并理解其成像过程中的关键问题，如噪声等问题。

2. 建立有效的神经网络进行画质提升

在此部分，重点是设计并实现一个深度神经网络模型，该模型旨在针对三维图像数据进行有效的去噪、清晰度提升和细节恢复。

3. 对网络进行轻量化设计

本环节的核心是对已有的深度学习模型进行优化，降低模型的计算复杂度和内存需求，使其更适合于移动设备。

4. 进行实际传感器数据的画质及计算复杂度评估

本部分将通过使用真实的 ToF 传感器数据来评估优化后网络的画质提升效果和计算效率，确保其在实际应用中的实用性和高效性。

(3) 国、内外研究现状和发展动态

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

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]。

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。

2. 基于深度学习的原始数据处理

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。

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，还存在以下局限性：

(1) 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；

(2) 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；

(3) 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，在已有基础中），以及准确度和功耗的平衡。

3. 三维图像传感器信号增强网络的轻量化

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，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。

参考文献：

- [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.
- [12] Zhuang Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, and Trevor Darrell. Rethinking the value of network pruning. In *International Conference on Learning Representations*. 2019.
- [13] Aojun Zhou, Anbang Yao, Yiwen Guo, Lin Xu, and Yurong Chen. Incremental network quantization: Towards lossless CNNs with low-precision weights. In *International Conference on Learning Representations*. 2017.

(4) 创新点与项目特色

1. 内核范围的动态调整

对内核范围进行动态调整，允许网络高效捕捉从长距离到短距离的依赖关系。实施类似策略预计可以使网络更有效地处理不同细节的尺度，改善输出质量，提高模型性能。

2. 改善模型大小和复杂度

减少层数和参数数量，同时通过智能设计（如使用更少但更有效的迭代）维持性能，预计可以实现适合移动或嵌入式应用的轻量级模型。

(5) 技术路线、拟解决的问题及预期成果

1. 技术路线

(1) 三维图像传感器数据分析

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。对 ToF 传感器采集的数据进行详细分析。分析包括深度图和三维点云数据的特性，以及在成像过程中的潜在问题，如噪声等。

(2) 建立有效的神经网络进行画质提升

设计和实现深度学习模型，专注于提升三维图像的质量。利用深度神经网络处理原始 ToF 信号，实现去噪等功能。

(3) 对网络进行轻量化设计

在确保画质提升效果的基础上，对深度学习模型进行轻量化处理，降低其计算复杂度和内存需求，减少模型的参数数量和计算负担；此外，通过量化压缩或混合精度处理对模型进行压缩，使模型适合于资源受限的移动设备，如智能手机、机器人和自动驾驶系统。

(4) 进行实际传感器数据的画质测试及计算复杂度评估

对轻量化后的模型进行实际测试，使用 ToF 传感器采集的数据评估模型的画质提升效果和计算效率。评估包括对比轻量化前后的画质差异、运行速度、功耗等指标，以确保模型在实际应用中的有效性和实用性。

2. 拟解决的问题

本项目旨在解决当前三维图像传感器信号处理网络在轻量化方面的不足。目前，许多基于深度学习的方法虽然在画质提升方面效果显著，但它们往往具有较高的计算复杂度和功耗，这在资源受限的移动设备上尤为突出。这些问题包括但不限于：

(1) 高功耗：现有的深度学习模型通常需要大量计算资源，导致在移动设备上运行时耗电量。

(2) 计算复杂度高：复杂的模型结构导致处理速度慢，不利于实时应用场景，如自动驾驶或机器人导航。

(3) 模型尺寸大：大型模型难以部署在存储和内存受限的移动设备上。

(4) 优化效率不足：现有模型对于 ToF 传感器的特定需求优化不足，可能无法充分发挥传感器的潜能。

3. 预期成果

项目的核心目标是实现三维图像传感器信号增强网络的轻量化，以提升模型在移动设备上的应用潜力。预期成果包括：

(1) 降低功耗：通过轻量化设计，大幅减少模型在移动设备上的能耗。

(2) 提升处理速度：简化网络结构，减少计算量，实现快速处理，以适应实时应用需求。

(3) 减小模型参数量：通过剪枝、参数量化等技术减小模型尺寸，使其适合于存储和内存受限的设备。

(4) 优化模型效率：定制化模型设计，针对 ToF 传感器的特点进行优化，提高其在特定应用场景下的效能。

(5) 实用性验证：通过实际传感器数据对模型进行全面评估，尝试在移动端部署模型，确保其在真实应用环境中的有效性和实时性。

(6) 项目研究进度安排

第一季度项目研究安排

文献调研和基础学习

(1) 深入研究相关文献，包括三维图像传感器技术、深度学习在图像处理中的应用，特别是 ToF 传感器数据的处理。

(2) 学习前人在该领域的研究基础，包括已有的方法和技术难点。

第二季度项目研究安排

1. 三维图像传感器数据分析

- (1) 收集和分析 ToF 传感器采集的数据，包括深度图和三维点云。
- (2) 识别和分析成像过程中的潜在问题，如噪声等。
- (3) 基于分析结果设计初步的数据处理策略。

2. 初步神经网络模型设计

- (1) 开始设计用于画质提升的初步深度神经网络。
- (2) 实施初步的去噪和图像质量提升实验。

第三季度项目研究安排

1. 建立有效的神经网络进行画质提升

- (1) 完善深度神经网络模型，优化其在画质提升方面的性能。
- (2) 进行深入的实验，以验证模型在不同场景下的效果。
- (3) 开始探索网络轻量化的可能性。

2. 网络轻量化初步实验

- (1) 初步尝试对神经网络进行轻量化处理。
- (2) 评估轻量化对画质提升效果的影响。

第四季度项目研究安排

1. 对网络进行轻量化设计

- (1) 完成神经网络模型的轻量化设计。
- (2) 优化模型以降低计算复杂度和功耗，同时保持画质提升效果。

2. 进行实际传感器数据的画质及计算复杂度评估

- (1) 使用实际 ToF 传感器数据进行综合测试，评估轻量化后模型的性能。
- (2) 对比轻量化前后的画质、处理速度和功耗，以确保模型的有效性和实用性。

(3) 准备项目结题报告与结题答辩，整理研究成果与实验数据，进行对未来工作的展望。

(7) 已有基础

与本项目有关的研究积累和已取得的成绩

1. 基于 ToF 成像原理，提出 confidence-aware graph Laplacian 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，结果相比现有方法在 FLAT 数据集上 MAE 指标提升 77.2%。

2. 面向图神经网络 (GNN) 的图结构及参数稀疏化，提出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；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；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。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，结果在 Cora、Citeseer、PubMed 等数据集上均超过现有方法。

3. 本小组在项目进展以来进行了如下实验，并取得了阶段性成果：

(1) 实验设置

实验使用了标准化的测试集，并对两种模型进行了交叉验证以确保结果的可靠性。性能指标包括均方根误差 (RMSE)、平均绝对误差 (MAE) 以及相对误差 (rel)，这些指标能够全面评估模型在深度估计任务上的表现。

(2) EMA 模型性能

EMA 模型是一个参数量巨大的架构，总参数达到 5,611,351，占用内存 22.45 MB，在模型的一次前向传播中执行的浮点运算总数为 9,652,130,816 次。尽管 EMA 模型在信号增强任务中展现出更好的性能，但它对于计算资源消耗过大，无法在算力受限的条件下部署。

(3) UDA 模型性能

与 EMA 模型相比，UDA 模型是一个更轻量的架构，其参数总数为 144,386，占用内存 0.58 MB。然而 UDA 模型的表现并不理想，其相对误差指标也未达到理想的水平。

在三维图像传感器信号增强任务中，EMA 模型因其高参数量而对计算资源要求极高，而 UDA 模型轻量化同时却失去了准确性。我们希望能通过一种方法使得模型在轻量化的同时尽量减少甚至消除对模型效果的影响。

已具备的条件，尚缺少的条件及解决方法

1. 已具备的条件

(1) 与本项目有关的研究积累和已取得的成绩

①基于 ToF 成像原理，提出 confidence-aware graph Laplacian regularization (CGLR) 作为信号先验，与深度神经网络的结合得到 DeepCGLR 网络。

②面向图神经网络（GNN）的图结构及参数稀疏化，提出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。

(2) 指导老师在相关领域有着较高的研究水平，能够给予我们专业指导。

(3) 随着计算机科学技术的发展，有关三维成像与传感技术的相关文献与参考资料丰富，便于我们查阅相关资料以获取专业指导。

(4) 团队拥有符合本项目研究的技术基础。

2. 尚缺少的条件及解决方法

(1) 本项目成员均为本科二年级学生，相关的专业基础知识和科研经历较少，应大量查阅相关的文献资料，努力提高相关专业领域的知识水平，提升科研素养。

(2) 缺少 kinect v2 或 Azure Kinect 传感器用于 ToF 数据采集，可以使用经费购置相关设备。

(3) 缺少足够的图形计算资源（GPU），可以考虑申请电子与信息工程学院的机房机器使用。

三、经费预算

开支科目	预算经费（元）	主要用途
预算经费总额	10000.00	项目研究经费
专用设备费 (购置、维修、租赁)	4000.00	购置 Kinect 或 realsense 传感器、租用服务器等
材料费	500.00	购买专业书籍和资料等

开支科目	预算经费（元）	主要用途
测试化验加工费	0.00	无
差旅费	500.00	差旅费等
会议费	0.00	无
国际合作与交流费	0.00	无
出版/文献/信息传递/ 知识产权事务费	5000.00	专利费、会议注册费、学术论文发表费用等 （若预算充足）
劳务费 （含专家咨询费）	0.00	无

四、项目组成员签名

林继申 刘垚 刘淑仪 梁斯凯 杨宇琨

五、指导教师意见

导师（签章）：
年 月 日

六、院系大学生创新创业训练计划专家组意见

教学负责人（签章）：
年 月 日

七、学校大学生创新创业训练计划专家组意见

		负责人（签章）：
年	月	日