

# 中山大学 大学生创新训练计划项目申请书

项目名称：面向机器人视觉系统的深度估计硬件  
处理器设计

相关学科：微电子科学与工程

申 请 人：卓子莹，姚思颖，周彤昕，彭逸杰，  
肖懿珊

所在院系：集成电路学院

指导教师：毛文东

预期成果类别：

☐ 社科类社会调查报告及学术论文

☐ 自然科学类学术论文

填表时间 二〇二四 年 1 月 16 日

项目名称											
申请人或申请团队	姓名 (负责人请加*号)	学号		性别	身份证号		年级	专业	联系电话	手机	E-mail
	卓子莹*	21312216		女	440514200303120063		21 级	微 电 子 科 学 与 工 程	13822823136	13822823136	Zhuoziying1@163.com
	姚思颖	21312486		女	230605200306111824		21 级	微 电 子 科 学 与 工 程	13089070393	13089070393	1402067113@qq.com
	周彤昕	21312832		男	371702200302251019		21 级	微 电 子 科 学 与 工 程	13396210825	13396210825	Zhoutx8@mail2.sysu.edu.cn
	彭逸杰	21312420		男	440902200211010015		21 级	微 电 子 科 学 与 工 程	13592997801	13592997801	2401545262@qq.com
	肖懿珊	22362088		女	440305200404150424		22 级	微 电 子 科 学 与 工 程	13641488485	13641488485	810752052@qq.com
指导教师	姓名	性别	年龄	职称	职务		所属院系	联系电话	手机	E-mail	
	毛文东										
研究起止时间		2024 年 1 月至 2024 年 12 月									
一、申请理由											
1. 项目概述： 为应对机器人硬件资源受限及高实时性要求的挑战，本项目将面向机器人视觉系统应用，基于可编程逻辑阵列（FPGA），设计一适用于搭载机器学习单目深度估计算法的硬件处理器，通过软硬件多维度协同优化，在满足轻量级、低功耗等需求下，优化其运行性能。											
2. 团队概述： (1) 团队成员均进行过系统的模拟电路、数字电路、FPGA 实验课程、数据结构与算法等专业课程修习，具有大学生创新训练项目实践经历，具备扎实的项目推进能力。 (2) 团队 5 人对于图像分析、FPGA 硬件处理器搭建、python 算法设计、深度学习模型构建等具备浓厚兴趣。同时，项目成员积极参与数学建模比赛、电子设计比赛等，对于模型数据收集处理、深度学习模型构建等具备一定经验，实现过基于深度学习的道路坑洼识别模型构建。											
3. 指导老师概述：											

- (1) **基本情况：**本团队的指导老师由毛文东老师担任。主持中山大学百人计划启动经费一项；参与江苏省重点研发、科技部重点研发、国家自然科学基金项目三项；参与和华为、快手公司等大型企业与研究所合作项目多项。主要研究方向为人工智能系统与集成电路设计、深度学习与硬件加速、计算机视觉等，在 IEEE 主流期刊和会议上共发表论文近 20 篇，申请专利 10 余项。
- (2) **项目支持：**统筹规划项目方案，设计合理的研究路线，为研究提供设备与技术支持。

二、项目创新特色概述(50 字以内)

1. 目前搭载机器学习单目深度估计算法模型的硬件处理器研究尚不够深入。
2. 本项目计划设计搭载面向机器人视觉系统的深度估计硬件处理器，利用软硬件协同设计思路实现处理器多维度优化。

三、计划项目实施思路（3000 字以内）

1. 研究意义与目的，同类研究工作国内外研究现状与存在的问题等

(1) 研究背景：

机器人工程作为新型工业化中最主要的趋势，其产业需求不断扩大。目前的机器人仍然在外界感知能力方面存在短板，诸如视觉感知等，这极大地影响了机器人的推广和使用，因此提高机器人的外界视觉感知能力显得尤为重要。而深度估计技术是实现三维场景识别的重要组成部分，相较于利用双幅图间的细微偏差实现三维场景构建的双目深度估计算法，通过机器学习构建的单目深度估计算法由于其环境适应性、部署灵活性、识别准确性更受机器视觉领域重视。由于机器人单目深度估计算法本地部署，对专用硬件处理器的轻量级、低耗能等工作特性提出了更高的要求，搭建一个可用于单目深度估计的本地硬件处理器将对机器人产业的发展起到广泛的推动作用。

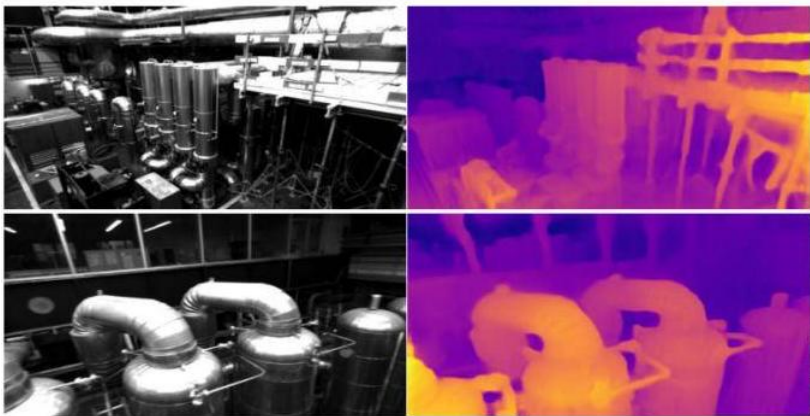


图 深度估计原输入图像与结果图

(2) 同类研究工作国内外研究现状：

单目深度估计算法的发展主要经历了传统算法阶段、机器学习算法阶段，并在近 10 年过渡到深度学习算法阶段，其中深度学习算法分为有监督单目深度估计算法和无监督/自监督单目深度估计算法。Eigen 等人(2014)首次将深度神经网络用于单目深度估计任务。Liu 等人(2015)将深度卷积神经网络与连续条件随机场结合，提出深度卷积神经场，用以从单幅图像中估计深度。Laina 等人(2016)提出了一种基于残差学习的全卷积网络架构用于单目深度估。Wang 等人(2016)提出使用并行卷积神经网络将深度与多种信息结合的框架。Fu（2018）等人提出了新的深度有序回归 DORN 网络。Chen 等人(2021)提出了一种适用于单目深度估计的拉普拉斯金字塔网络结构。Zhang（2022）等人提出了基于条件对抗网络的深度估计模型。

近年来，在优化硬件处理器性能方面，已有部分研究对存储器存储方式，在时序优化和并行化流水线化等方面进行设计。

存储器优化方面黄伟等人通过检测和标记 NAND FLASH 芯片中的坏块并替换，保证数据完整性；采用多平面流水线式的编程方法，提高了数据传输效率和准确率。减少数据传输时间，以此降低功耗[1]；通过加权优

先级算法，以减少内存访问的总执行时间优先级标准重新组织访问顺序通过顺序优化所带来的预期性能提升能够大大补偿预处理工作的时间成本，来改进了内存访问性能，并实现了理想的性能与资源比率。[2]曾翔宇等人通过 OpenCL 存储器模型的 FPGA 存储器层次结构，来有效地处理内存管理，无需针对特定应用进行重新设计。在时序优化和并行化流水线化方向[3]Monteiro 等人通过使用 OpenMP 标准的软件并行化技术来提高算法的加速。以及一种基于流水线结构的多项式除法的并行方法。

通过大量的资料查询，团队发现，截止目前，国内外主要将精力集中在深度估计的算法开发上，而在设计专用于实现深度估计算法的轻量级、低功耗硬件处理器方面的研究不够深入。

**(3) 研究意义与目的:**

单目识别算法，采用了复杂的神经网络，这就需要算力强大的硬件处理器。ASIC、GPU、FPGA 为三个可用于增加 CNN 数据吞吐量的硬件设计方向。通过对比，GPU 具有高存储能力、大数据吞吐量和并行处理图像的能力，但需要与 CPU 配合工作，产生高功耗。ASIC 硬件处理器虽然体积小、功耗低、速度快，但开发周期长、技术成本高，且不利于机器人功能更新。故上述两种硬件处理器在规模、功耗、成本、运算速度等方面并不适合直接在机器人平台部署。相比于其他硬件加速器，现场可编程门阵列 FPGA 具有更高的灵活性和可定制性，高性能，功耗低的特点。同时 FPGA 硬件结构提供了高度定制化的数据流和并行处理能力，在实时性方面具有更好的表现。

在图像处理的过程中，主要存在两方向的限制：在软件方向，相关算法由大量计算密集型指令组成，需要大量的资源和计算能力，导致算法执行时间过长；在硬件方向，需要大量的存储单元来存储中间和最终结果，耗能巨大。而通过软硬件协同优化，可以实现系统整体性能的提高，能够更好地应对复杂的任务和大规模的数据处理。

因此面向机器人深度估计算法的硬件处理器应当满足轻量级、低耗能等需求。团队通过查询文献发现，虽然深度估计算法发展迅速，但专用的协同硬件处理器设计及多维度优化方面的研究较为匮乏。

**2. 研究内容及工作方案**

**(1) 研究内容:**

**• 硬件需求分析**

对机器人视觉系统的深度估计任务做硬件需求分析，对硬件处理器的性能、存储需求等指标进行评估，后需根据评估完成设计；

**• 网络结构选择**

选择合适的深度估计网络结构，作为主要的卷积层网络结构。本项目选择 ResNet18 或 ResNet50 网络结构，不仅在深度估计任务中表现良好，而且相对适合硬件加速器的设计。

**• 硬件加速器设计**

使用 Xilinx Vivado 进行编码设计，基于选定的网络结构进行硬件加速器的设计，包括对卷积层、池化层、激活函数等关键网络层的硬件实现。

采用一定的通信网络来连接不同的计算节点，将多个计算节点或处理器通过高速通信网络连接起来，设计并实现节点之间的通信协议，编写并行计算程序，设计并实现硬件阵列的结构，将输入数据划分为多个子数据集，以构成一个大规模的并行计算系统。

**• 硬件-软件协同设计**

进行硬件与软件的协同设计，确保设计的硬件处理器能够与深度估计网络的软件模型进行有效集成和协同工作，以实现深度估计任务的高效执行。

**• 性能评估和优化**

使用 EDA 工具等，测量器件能耗、布局布线、工作频率、处理速度等数据，将硬件处理器与 GPU 平台进

行比较，对性能优劣进行分析和评估。

根据测试结果和比较分析，对硬件处理器进行迭代优化，以提高其性能。

使用 Xilinx Vivado 等工具生成硬件描述语言的比特流文件，并进行上板测试，以验证硬件处理器的性能和可靠性。

(2) 工作方案：



3. 拟解决主要问题
- (1) 面向基于机器学习方法构建的单目视觉伺服深度估计算法，设计一种搭载有深度估计算法模型的硬件处理器。
  - (2) 结合机器人视觉的应用场景优化硬件处理器性能，满足机器人视觉系统的轻量级、实时性等应用需求。
  - (3) 对硬件处理器进行多维度系统优化，并结合性能测试结果和与 GPU 的比较分析进行迭代优化，以软硬件协同设计的思路实现更高效准确的深度估计与更优的性能。
  - (4) 针对项目实际应用，视情况进行上板等应用实践，验证硬件处理器的性能及其应用效果，为后续的进一步研究和开发提供有价值的实践经验。

#### 4. 研究支持条件（使用仪器或设备）

- （1）软件支持：Matlab、Cadence、Vivado 等设计软件；
- （2）研究器材：搭载单目摄像头的 FPGA 开发板；
- （3）配置有仿真平台软件的 windows 系统计算机；
- （4）实验平台：集成电路学院实验室，用于模型训练的云服务平台

#### 5. 预期研究成果（研究论文、设计、专利、产品、鉴定、推广应用等）

- （1）**设计成果：**实现基于 FPGA 的面向机器人视觉识别系统的深度估计硬件处理器设计，进行实际烧录与测试。
- （2）**研究报告：**面向机器人视觉识别系统的深度估计硬件处理器设计；
- （3）**论文：**发表一篇关于面向机器人视觉识别系统的深度估计硬件处理器设计的论文；

6. 项目创新之处（原始创新：重大科学发现、技术发明；集成创新：融合多种相关技术，形成新产品、新产业；引进消化吸收再创新：在引进国内外先进技术的基础上，学习、分析、借鉴，形成具有自主知识产权的新技术）

（1）在选题思路，针对现有领域中基于可编程逻辑阵列的机器视觉研究尚不深入这一问题，围绕我国自主研发的人工智能芯片和开发框架，发展软硬件协同和系统级优化技术，本项目从硬件角度切入人工智能这一主题，设计适用于搭载机器学习单目深度估计算法的硬件处理器。

（2）在研究方法上，融合采用并行计算的可编程逻辑阵列进行多维度迭代优化，实现硬件处理器的整体优化。

（3）在项目成果上，进行软硬件协同设计，结合机器人视觉系统的硬件架构与单目深度估计的软件算法，设计一种面向机器人视觉的硬件处理器；

（4）在应用场景上，为应对机器人硬件资源受限及高实时性要求的挑战，面向机器人视觉系统，聚焦轻量级、实时性等性能应用需求，对硬件处理器进行性能优化。

四、计划项目管理（800 字以内）

1、计划项目人员分工

（1）调查准备阶段：

资料的查阅、项目方案的研究	项目的选题与方向：全体成员 项目资料的收集：全体成员 与指导老师讨论并对项目选题方向进行确定：全体成员
项目资料完善	对项目相关的文献进行整理分类：全体成员
进行项目实施方案的制定	项目具体实施方案的制定：全体成员 与指导老师讨论并对项目实施方案进行调整：全体成员
完善项目实施计划细则	项目具体应用技术的确定：全体成员 项目具体流程与研究时间安排的确定：全体成员

（2）项目开展阶段：

项目的规划与开展	项目流程及进度的统筹规划：卓子莹
深度估计算法模型的选择与量化	算法模型分析：卓子莹，姚思颖 模型训练与移植：周彤昕，彭逸杰，肖懿珊
硬件处理器设计	Verilog 编码：周彤昕，卓子莹 仿真验证与物理验证：姚思颖 硬件综合与时序分析：彭逸杰，肖懿珊 软硬件协同迭代优化：全体成员
项目测试与改进	将项目系统移植到测试设备：全体成员 对项目系统进行测试与调整：全体成员 向指导老师展示并根据老师建议对项目系统进行再调整：全体成员

（3）项目结项阶段：

项目研究报告与论文撰写	项目成果整理：全体成员 研究报告及论文撰写工作：全体成员	
2、计划项目研究时间安排（查阅资料 、选题、项目研究方案、开题报告、实验研究、数据统计、处理与分析、研制开发、结题、撰写研究论文和总结报告、参加结题答辩和成果推广等）		
<p>此项目预计持续 1 年，预计 2024 年 12 月完成结题报告。将时间分为六个阶段，并对各阶段分配相应的阶段任务，预计每 2 周召开一次小组会，汇报成员进度并明确阶段工作计划，在每阶段完成后向指导老师汇报工作，根据指导老师的意见调整工作方案。</p>		
<b>阶段一（2023 年 11 月-2023 年 12 月）</b>		
<p>通过小组讨论，初步确定项目研究方向，联系指导老师并确定选题，确定初期成员工作分工与工作流程，学习国内外相关领域研究现状并研读重要研究成果，以产生本项目的基本思路，最终形成并递交项目申请书。</p>		
<b>阶段二（2023 年 12 月-2024 年 1 月）</b>		
<p>采购相关设备，对项目所需知识技能补漏，深入研读初期获得的资料，保持指导老师的密切联系，深入对项目问题的认识，为文本阅读研究打好基础，针对机器人视觉、深度估计等研究方向进行探究式学习与梳理，并辅以国内外相关研究进行探究性思考，在小组讨论中拟定可行思路。</p>		
<b>阶段三（2024 年 1 月-2024 年 3 月）</b>		
<p>按照工作方案流程，进行算法训练数据集收集处理，并完成机器学习深度估计算法模型的筛选，量化与训练。</p>		
<b>阶段四（2024 年 3 月-2024 年 6 月）</b>		
<p>完成面向机器人视觉系统的深度估计硬件处理器的设计，利用 verilog 实现处理器模块，并确定并行方案与硬件阵列，初步实现硬件处理器搭建，并进行调试与测量。</p>		
<b>阶段五（2024 年 6 月-2024 年 9 月）</b>		
<p>针对初步完成的器件进行仿真，并结合仿真结果进行多次迭代优化，从而实现软硬件协同多维度优化。</p>		
<b>阶段六（2024 年 9 月-2023 年 11 月）</b>		
<p>完成硬件处理器搭建优化后，尝试进行上机测试，实测其在搭载深度估计算法模型下的运行性能。有余力的话，将该硬件处理器作为实验机器人的视觉感知处理器，测试其实际应用性能。</p>		
<b>阶段七（2023 年 11 月-2023 年 12 月）</b>		
<p>负责撰写论文的组员利用已经已有的研究成果和分析报告撰写论文，审批结题，准备结题答辩。</p>		



五、项目成员是否有申请延期的大学生创新创业训练计划项目（如有，请填写项目成员姓名、立项时间、立项等级、项目名称和指导教师姓名

无。

五、经费预算及使用计划

总计：15000 元人民币

序号	支出项目	经费测算依据	金额（元）
1	设备购置费用		5000
2	云服务器平台租借费用		3000
3	论文版面费		5000
4	事务费（打印费，查阅资料费，培训费）		2000

经费使用计划：  
（1）设备购置费用：主要用于携带单目摄像头的 FPGA 开发板，以及其他实验开发设备的购置费用，预计花费 5000 元。  
（2）云服务器平台租借费用：主要用于深度估计算法模型的训练与调参，预计花费 3000 元。  
（3）论文版面费用：主要用于论文发表的版面购置费用。  
（4）事务费：进行项目科研期间所花费的资料查询、打印及培训相关费用。

申请团队签名：

指导  
教师  
意见

签名：  
年 月 日

院 系  
意见

签名（公章）：  
年 月 日

学 校  
审 核  
意见

签名（公章）：  
年 月 日

备注