

# 马睿斌 - 计算机视觉研究员 / 算法工程师

男/1993

目标城市：北京

现就职：Meta计算机视觉研究员（IC6 Research Scientist），美国西雅图

📧 iyu0126

✉️ ruibinma93@gmail.com

📞 (+1)2067376491

## ● 专业技能/知识

- 计算机视觉：三维重建，图像处理，多视角几何，Structure from Motion (SfM)，SLAM，鲁棒性估计，非线性优化，前沿机器学习模型的开发/应用/快速测试/部署/迭代，AR/VR，地图融合，地理对齐，大规模建图。
- 编程语言/工具：Python, C++, Pytorch
- 大模型基础知识如Transformer，<https://ruibinma-llm-cv.streamlit.app>是用LangChain编写的本简历的AI版本

## ● 教育背景

### 美国北卡罗来纳大学教堂山分校 - 计算机科学 - 博士

2016/8-2021/1

- 博士课题：计算机视觉，内窥镜视频的三维SLAM重建，深度学习在SLAM中的应用
- 导师：Stephen M. Pizer & Jan-Michael Frahm

2012/8-2016/7

### 清华大学 - 电子信息科学类 - 本科 工学学士 (平均成绩: 90.8/100.0)

- 2013&2015郑格如奖学金；2015挑战杯竞赛三等奖；2016汤言英奖学金

## ● 工作

### Meta (Facebook) IC6计算机视觉研究员（Research Scientist）| 在职，2021.4至今，西雅图

- 大规模（城市级）室外众包（crowdsourced）图像的3D稀疏重建
  - 带领团队对100个城市中心区的Facebook街景图像进行了三位稀疏重建，覆盖范围达到4000km<sup>2</sup>以上。
  - 项目使用了多数据来源，包括Mapillary.com中的亿级别众包（即来自于人群，非专业采集）图像数据，第三方提供商数据，以及Facebook/Instagram用户图像等。项目旨在探索使用这些廉价数据源进行高精度三维重建的可行性以替代昂贵的大规模专业采集。在覆盖密集地区可达到有效地图全覆盖。
  - 3D重建系统以运动估计结构（Structure from Motion/SfM，如Colmap/OpenSfM）为核心，并辅以多项基于前沿神经网络模型的先验信息估计（如全景语义分割（Panoptic Segmentation）、重力估计、相机内参估计等等），以及融合了GPS鲁棒性估计、ICP和深度学习的3D地图向2D城市地图的对齐算法。
  - 3D地图向多个下游应用产品提供高质量输入。例如1) 智能眼镜导航/图像重定位，适用于GPS和磁力计高度不准确的地区，如市中心；2) AR眼镜的6自由度位姿估计，以进行增强现实渲染；3) 为机器学习模型提供带有绝对位姿（地理对齐）的高质量图片训练数据，如微调大模型增强空间感知能力。
  - 高效并行化的建图系统。我们将目标城市进行特定尺度的切块（tiling），并基于Facebook云计算基础设施开发了对于所有地块完全并行化的建图系统，将每一个城市的更新延时（latency）限制在24小时内。
- 室内稀疏3D重建，地图融合与重定位
  - 项目服务于Quest（VR）用户的室内地图的云端存储及更新，以方便用户快速进入游戏而不必重复进行室内环境的扫描。我负责了云端上的稀疏3D地图的融合，即新旧地图通过特征点匹配进行刚性对齐（rigid alignment）的部分。通过多种策略，在不伤害召回率的前提下将误融合率降至接近0%。
  - 融合算法还支持了将Quest SLAM地图和高精度激光雷达（LiDAR）扫描地图进行匹配，来为Quest图像的深度估计模型提供大量高质量训练数据。
- 室内稠密3D重建与布局估计
  - 项目服务于Quest用户和室内环境的交互式体验，包括对室内场景进行基于TSDF的稠密重建，以及对墙壁，物体，通道等布局元素的提取。我负责了对TSDF体信息的云端存储及更新模块，使得用户再次使用头显时可以完全跳过已经完成建模的区域，大幅度改善用户体验。

- 激光雷达约束下的3D地图优化
  - 在带有激光雷达专业采集设备 (pro rig) 的建图项目中, 我负责了对不同关键帧的激光雷达点云进行 ICP 匹配、生成高精度相对位姿约束、并进行位姿优化的部分。算法相对纯视觉闭环算法可生成更多高质量的闭环约束, 可大幅降低地图漂移 drift。

## Meta (Facebook) 软件工程师实习生 | 2020.5-2020.8, 纽约

- 基于轻量级前馈神经网络的移动信号的地点排序 (Place Ranking): 使用多任务多层感知器 (MLP) 来有效利用多种数据源的标注信息。准确率超越 Meta 的产品模型并大幅度降低 CPU 和内存开销, 可完全在移动端部署。

## 谷歌 (Google AI Healthcare) 软件工程师实习生 | 2019.5-2019.8, Palo Alto

- 基于 BERT 微调开发病理报告分类系统, 在五项多标签任务中取得优异的 mAP 指标; 在高置信度结果中系统准确率媲美人工, 成功降低 80%-98% 的人力成本。论文发表于 Neurips workshop。

## 谷歌 (Ads, Tools and Infrastructure) 软件工程实习生 | 2018.5-2018.8, Mountain View

- 通过微调神经机器翻译模型实现代码修复: 通过挖掘 Google 内部 Python 数据库中的历史修复模式 (Patterns) 来微调神经机器翻译模型。该模型能够自动将存在 Lint 问题的代码进行修复。

## ● 博士课题

### 结肠镜图像的三维重建 (SLAM)

我基于 SLAM 技术 (DSO 系统) 与深度学习实现了从肠镜影像中实时重建肠道表面, 用于辅助发现盲区并提供手术导航。我创新性地将神经网络预测的深度图作为 SLAM 流程的鲁棒关键帧深度, 并利用局部光束法平差 (Local Bundle Adjustment) 优化相机位移带来的漂移。这种双赢策略解决了传统 SLAM 在纹理匮乏的肠道图像中难以运行的痛点以及累积神经网络模型输出漂移过大的问题。通过深度图融合, 我成功实现了肠镜影像的实时 3D 重建。  
视频: <https://endoscopography.web.unc.edu/data-and-results/>

### 内窥镜视频中的图像检索 (Image Retrieval)

我开发了一套旨在辅助医生确定结肠镜进入位置及进度的手术导航系统。针对肠道影像纹理极度匮乏、导致传统词袋方法 BOW 失效的痛点, 我使用对比损失 (Contrastive Loss) 训练卷积神经网络来构建全局描述符, 为每帧图像生成更鲁棒的特征表达。贡献了行业内第一个结肠镜下定位的标注的性能评估数据集 (benchmark)。

### 广义柱体的表面形变 (图形学)

我攻克了将扭曲的广义柱面三角网格进行无褶皱最小畸变拉直的图形学问题。我将形状分析的数学理论 (骨架表达) 与有限元网格变形方法跨学科结合, 完成了核心算法。该系统采用 C++ 开发, 基于 Qt Creator 和 VTK (Visualization Toolkit) 实现了图形处理与可视化, 解决了复杂几何体在变形过程中的扭曲失真问题。

## ● 其他专业经历

### 讲师: 北卡罗来纳大学教堂山分校计算机系课程 (COMP776) 《三维世界的计算机视觉》

- 主讲人 | 研究生课程 | 课程内容涉及 3D 计算机视觉, 包括: 像机模型, 投影, 成像, 图像变换及处理, 透视变换 (单应矩阵 Homography), 鲁棒性估计 (RANSAC), 对极几何 (epipolar geometry), 多视角几何 (multiview geometry), 立体深度估计 (stereo), 从运动估计结构 (Structure from Motion, SfM), 卷积神经网络基础介绍。

### 本科暑期实习: 布莱根妇女医院, 哈佛医学院 | 2015.7-2015.9

- 医学影像可视化软件《3D Slicer》中的模块开发: 该模块实现了妇科近距离放射治疗中 MRI 图像管片的自动分割。在此基础上, 我贡献了一个迭代误差检测与修正框架, 相关合作论文已发表于医学影像顶级期刊 Medical Image Analysis (Media)。

### 清华大学本科毕业设计: 《基于深度学习的废弃物视频中的可回收物体检测》

- 通过微调 FasterRCNN 实现了一个能够在废弃物传送带视频中有效实时检测可回收垃圾的 Python 程序。

- **R. Ma** and R. Wang, S. Pizer, J. Rosenman, S. McGill, and J. Frahm. Real-time 3D reconstruction of colonoscopic surfaces for determining missing regions. In Medical Image Computing and Computer Assisted Intervention (MICCAI) (2019), Shenzhen, China. **Best Presentation Award**
- **R. Ma** and R. Wang, S. Pizer, J. Rosenman, S. McGill, and J. Frahm. RNNSLAM: Reconstructing the 3D Colon to Visualize Missing Regions during a Colonoscopy. In Medical Image Analysis - MICCAI 2019 Special Issue
- **R. Ma**, S. McGill, R. Wang, J. Rosenman, J. Frahm, Y. Zhang, and S. Pizer. Colon10K: A Benchmark for Place Recognition in Colonoscopy. In IEEE International Symposium on Biomedical Imaging (2021)
- **R. Ma**, C. Chen, G. Li, W. Weng, A. Lin, K. Gadepalli, and Y. Cai. Human-centric Metric for Accelerating Pathology Reports Annotation. In Neurips 2019 Workshop of Machine Learning for Health, Vancouver.
- **R. Ma**, Q. Zhao, R. Wang, J. Damon, J. Rosenman, and S. Pizer. Deforming generalized cylinders without self-intersection by means of a parametric center curve. In Computational Visual Media Journal (2018).
- **R. Ma**, Q. Zhao, R. Wang, J. Damon, J. Rosenman, and S. Pizer. Skeleton-based generalized cylinder deformation under the relative curvature condition. In Pacific Graphics (2018), Hong Kong, China.
- Y. Zhang, S. Wang, **R. Ma**, S. McGill, J. Rosenman, and S. Pizer. Lighting Enhancement Aids Reconstruction of Colonoscopic Surfaces. In Information Processing in Medical Imaging (IPMI) (2021), Bornholm, Denmark.
- S. McGill, J. Rosenman, R. Wang, **R. Ma**, J. Frahm, S. Pizer. Artificial Intelligence Identifies and Quantifies Colonoscopy Blind Spots. Endoscopy (2021)
- S. Zhang, L. Zhao, S. Huang, **R. Ma**, B. Hu, Q. Hao. 3D Reconstruction of Deformable Colon Structures based on Preoperative Model and Deep Neural Network. In IEEE International Conference on Robotics and Automation (ICRA) (2021), Xi'an, China
- F. Chen, **R. Ma**, J. Liu, M. Zhu, and H. Liao. Lumen and media-adventitia border detection in IVUS images using texture enhanced deformable model. In Computerized Medical Imaging and Graphics (2018) 66.
- A. Mastmeyer, G. Pernelle, **R. Ma**, L. Barber, and T. Kapur. Accurate model-based segmentation of gynecologic brachytherapy catheter collections in MRI-images. In Medical Image Analysis (2017).
- [专利] **R. Ma**, R. Wang, S. Pizer, J. Frahm, J. Rosenman, and S. McGill. Methods, Systems and Computer Readable Media for Three Dimensional (3D) Reconstruction of Colonoscopic Surfaces for Determining Missing Regions. United States Patent NO. US010682108B1.