

# 衡誉文, 博士

✉ y.heng@soton.ac.uk 📞 18522708355 💻 hengyuwen.com 🐱 123mutourener 🆔 0000-0003-3793-4811

## 🎓 教育经历

南安普顿大学 (QS78)	南安普顿, 英国
计算机科学博士, Vision, Learning and Control Research Group	2020–2023.08
爱丁堡大学 (QS15)	爱丁堡, 英国
数据科学硕士, Distinction	2019–2020
爱丁堡大学	爱丁堡, 英国
电子电气工程荣誉学士, Second Class, Division 1	2015–2017
华北电力大学	北京, 中国
电气工程及其自动化学士, GPA: 3.36	2013–2015

## 🔬 研究方向

计算机视觉及其应用, 包括沉浸式声音渲染的材料分割、自动驾驶感知算法 (传感器融合方案)、自动驾驶数据集构建 (数据挖掘与 corner case 生成方案)、以及全景深度估计等。

## 💼 工作经历

🐾 百度 ACG	上海, 中国
自动驾驶感知算法-算法组组长	2021.12–至今
<ul style="list-style-type: none"><li>– 工作简介: 算法组业务方面, 负责百度众测算法组客户需求拆解、算法方案制定优化、工程落地项目管理等, 涉及业务包括自动驾驶仿真、数据挖掘、自动化标注、自动化验收、真值系统五个方向。算法研究方面, 主要负责 3D 点云检测与分割创新研究。</li><li>– 算法组重点产出:<ol style="list-style-type: none"><li>1. 自动驾驶仿真业务: 基于 carla 与 ros, 成功开发基于 docker 的容器化仿真环境, 能够调用自定义算法并加载场景进行自动驾驶仿真评测, 并配合平台开发, 为百度云提供了早期可以对外展示的自动驾驶仿真平台。</li><li>2. 数据挖掘业务: 基于 2D、3D 检测模型以及时序车辆传感器数据, 为客户提供包括红绿灯、交通参与者、闸机等静态场景和掉头、转弯等动态场景数据挖掘服务, 成功缩短数据采集标注流程, 并完成一项交付业务。</li><li>3. 自动化标注业务: 基于百度内部大量标注数据, 预训练包括 3D 点云检测、线识别在内的 7 类大模型, 并在内部标注组进行落地验证, 结果表明能够整体提升 30% 的标注效率。</li><li>4. 自动化验收业务: 基于过拟合训练, 在已经标注好的数据中, 找到小部分与整体标注规范不符合的标注结果, 成功减少一轮人工验收流程。</li><li>5. 真值系统: 为百度云 roofoox 真值系统提供算法支持, 包括云端真值推理以及车端真值可视化两个应用场景, 主要支持算法包括融合 3D 检测、多目标追踪、轨迹还原与预测等。</li><li>6. 算法研究: 在 3D 任务方面, 产出两项创新方案, 并计划分别投稿至 BMVC2023 与 3DV2023 两个会议。</li></ol></li></ul>	
北京穿越科技有限责任公司	北京, 中国
AR 技术联合创始人	2017.07–2018.05
<ul style="list-style-type: none"><li>– 公司简介: 北京 Overleap 是一家由我和我的两个同学创立的初创公司, 种子轮融资 100 万人民币, 旨在开发低成本的增强现实 (AR) 硬件;</li><li>– 工作简介: 负责使用开发工具包 Vuforia 在 Java 环境中开发 AR 软件;</li></ul>	

- 软件适配：修改了 Vuforia 的渲染相关的 Java 类，将仿射平移应用到渲染的每一帧中，这样为手机开发的应用程序就可以在我们的眼镜硬件上运行，而不需要修改 3D 资源和应用场景；
- 软件开发：参加比赛并撰写商业计划书，与我的团队一起为公司筹集资金，为一家科学博物馆用 3D 打印的结构建造了功能齐全的初代产品。

## 🔪 研究经历

### 自动驾驶感知系统—3D 融合检测与分割

2021.12–至今

- 研究描述：该研究项目诞生于百度 ACG 数据应用业务部业务困境，目的是为自动化标注、自动化验收以及真值系统提供超越当前水平的 3D 算法。
- 研究背景：当前的开源 3D 融合模型，存在远距离点云稀疏、体素化未考虑空间分布等问题，导致难以准确检测分割较远距离的 3D 物体。因此，我们通过分析点云分布，结合图片与点云数据特性，提出优化的融合算法，提升整体 3D 任务的表现。
- 研究成果：提出并实现
  1. 基于 cylinder3D，结合点云密度进行体素化；
  2. 基于图像深度极限，融合点云特征与图像特征；
  3. 基于点云投影，补全在体素化过程中被误删除的稀疏信息；
  4. 根据 3D 检测与分割的模型共性，提出统一的 3D 任务网络结构。计划投稿 BMVC2023 与 3DV2023 会议。

### 基于场景感知的空间音频渲染—密集材料分割

2020.09–至今

- 研究描述：该项目旨在重建 3D 室内场景，并为物体提供材料标签，以合成具有多次反射与混响效果的空间音频。由南安普顿大学 Hansung Kim 博士和 Srinandan Dasmahapatra 博士指导；
- 研究背景：对于密集材料分割而言，由于从图片上缺乏具备准确辨别能力的视觉特征，因此以往的方法只能取得 70%左右的准确度，且在分割结果上会混淆物体与材料，因此，我们提出通过人工神经网络学习材料特征，再结合物体、场景等上下文特征，达到提升分割准确度的目的。
- 研究成果：提出并实现
  1. 基于半监督学习（self-training）的材料标签增强方法；
  2. 基于 boundary loss 的材料特征学习方法；
  3. 基于 transformer 的动态多尺度特征学习方法；
  4. 基于相机模型与超光谱图像重建的材料特征学习方法。

在测试数据上将准确率提升至最高 88.34%的精确度，达到先进水平，并在 EUSIPCO、ICASSP、BMVC、IEEE VR 等会议发布六篇论文，并在 Springer CCIS 成刊发布一章节，一篇在投 IEEE Transactions on Image Processing 期刊论文与一篇在投 ICCV 会议论文。

### 监督学习的数据缩减

2020.02–2020.11

- 研究描述：本项目是我硕士论文的一部分，由爱丁堡大学计算机科学基础实验室（LFCS）的 Cao Yang 博士指导；
- 研究内容：调查了现有的机器学习的数据缩减技术，重点是实例选择算法和非均匀采样算法。在 Python 中用 scikit-learn 和 TensorFlow 实现了这些算法。从理论上和经验上评估了减少的数据集与原始数据集的训练的"相对准确性"；
- 研究成果：提出了一个新的工作流程，首先用预先训练的网络参数提取图像特征，使算法适用于图像数据集和卷积神经网络（CNN）。设计了一种新的实例选择算法，根据分类难度和提取的特征空间的决策边界对实例进行加权。在分类问题上，能够以 80%的数据实现原数据集 96%的精确度。

## 发表论文

- Yihong Wu, Yuwen Heng, Mahesan Niranjan, and Hansung Kim. Depth estimation for a single omnidirectional image with reversed-gradient warming-up thresholds discriminator. In *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023
- Yuwen Heng, Srinandan Dasmahapatra, and Hansung Kim. Material recognition for immersive interactions in virtual/augmented reality. In *2023 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pages 577–578. IEEE, 2023
- Yuwen Heng, Yihong Wu, Srinandan Dasmahapatra, and Hansung Kim. Enhancing material features using dynamic backward attention on cross-resolution patches. In *33rd British Machine Vision Conference 2022, BMVC 2022, London, UK, November 21-24, 2022*. BMVA Press, 2022a
- Yuwen Heng, Yihong Wu, Hansung Kim, and Srinandan Dasmahapatra. Cam-segnet: A context-aware dense material segmentation network for sparsely labelled datasets. In *17th International Conference on Computer Vision Theory and Applications (VISAPP)*, volume 5, pages 190–201, 2022b
- Alawadh Mona, Wu Yihong, Heng Yuwen, Niranjan Mahesan, and Kim Hansung. Room acoustic properties estimation from a single 360° photo. In *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022
- Yihong Wu, Yuwen Heng, Mahesan Niranjan, and Hansung Kim. Depth estimation from a single omnidirectional image using domain adaptation. In *European Conference on Visual Media Production (CVMP)*, pages 1–9, 2021

## 专利

工作与创业期间，共产出发明专利 1 项，实用新型专利 2 项，外观专利 1 项。

## 获奖经历

海淀创业园; 海淀创业潜力奖	2017
红谷滩杯 VR/AR 行业赛; 优胜奖	2017
杭州留学生创新创业大赛; 杰出归国青年奖	2016