Yuwen Heng, PhD

y.heng@soton.ac.uk 18522708355 hengyuwen.com 123mutourener 0000-0003-3793-4811

Education

University of Southampton **Southampton, UK**
Doctor of Philosophy in Computer Science, Vision, Learning and Control Research Group 2020–Oct. 2023

University of Edinburgh **Edinburgh, UK**
Master of Science in Data Science, Distinction 2019–2020

University of Edinburgh **Edinburgh, UK**
Bachelor of Engineering with Honours in Electronics and Electrical Engineering, 2015–2017
Second Class, Division 1

North China Electric Power University **Beijing, China**
Bachelor of Engineering in Electrical Engineering and its Automation, GPA: 3.36 2013–2015

Research Interest

Computer Vision and its Applications, particularly material segmentation and panoramic depth estimation for immersive sound rendering, perception in autonomous driving (with a focus on sensor-fusion solutions), and the creation of perception datasets through data mining and corner-case generation.

Experience

Baidu ACG **Shanghai, China**
Internship in Autonomous Driving, Computer Vision, Leader of the Algorithm Team Dec 2021–Aug 2023

- Job Summary: Responsible for algorithm solution design, handling various business areas including autonomous driving simulation, data mining, automated annotation, automated quality assessment, and ground truth system. Also actively involved in innovative research in 3D point cloud detection and giant image retrieval model.
- Key Outputs of the Algorithm Team:
  1. Autonomous Driving Simulation: Developed a containerised simulation environment based on Carla and ROS using Docker. Integrated custom algorithms and loaded scenarios for autonomous driving simulation and evaluation. Collaborated with the platform development team to provide an early-stage autonomous driving simulation platform for external showcase in Baidu Cloud.
  2. Data Mining: Implemented data mining services based on 2D and 3D detection models, giant image retrieval models, and time-series vehicle sensor data. The service includes mining static scenes such as traffic lights, traffic participants, and barriers and dynamic scenes involving turns, cut-ins, cut-outs, and lane changes. The service shortened the data acquisition and annotation process.
  3. Automated Annotation: Utilised a large volume of annotated data within Baidu to train seven major models, including 3D detection and lane line recognition. These models were then validated by the internal annotation team, demonstrating an overall improvement of 30% in annotation efficiency.
  4. Automated Quality Assessment: Implemented an overfitting training approach to identify deviations from the overall annotation standards within labelled data. Successfully eliminated the need for one round of manual inspection process.
  5. Ground Truth System: Provided algorithmic support for Baidu Ground Truth System, including two application scenarios: cloud-based ground truth inference and on-vehicle ground truth visualisation. Main algorithms include 2D and 3D detection, lane line detection, multi-object tracking, trajectory reconstruction and prediction.

Beijing Overleap technology Limited **Beijing, China**
Co-founder and AR Engineer Jul 2017–May 2018

- Beijing Overleap was a start-up company founded by two of my classmates and me, aimed at developing cost-effective augmented reality (AR) hardware.
- Responsible for developing prototype AR software with the development toolkit, Vuforia in Java.
- Modified the render-related Java classes of Vuforia to apply affine translations to each frame so that the applications developed for mobile phones can run on our eyewear hardware without modifying the 3D resources and the application scene.
- Attended competitions and wrote the business plan to raise money for the company with my team. Successfully raised one million Chinese Yuan and built fully functional products for a science museum with 3D printed structures.

Projects

Dense Material Segmentation for Scene-Aware Spatial Audio Rendering Sep 2020–Now

- Research Description: This project aims to provide material labels for reconstructed 3D indoor scenes to synthesise spatial audio with multiple reflections and reverberation effects. Supervised by Dr. Hansung Kim and Dr. Srinandan Dasmahapatra at the University of Southampton.
- Research Background: Achieving accurate material segmentation for 3-channel RGB images is challenging due to the considerable variation in the appearance of a material. Moreover, the material datasets have sparse labels, where only a small region of the image is annotated. The sparse datasets present a significant challenge in recovering accurate material boundaries for lack of labelled samples near boundaries. To address these issues, multiple neural networks are proposed to enhance segmentation accuracy by incorporating material features along with contextual features related to objects and scenes.
- Research Achievements:
  1. Proposed and implemented a material labelling enhancement method based on semi-supervised learning (self-training).
  2. Introduced a semi-global contextual feature learning method utilising boundary loss.
  3. Designed a dynamic cross-resolution feature learning method based on transformer architecture.
  4. Developed a material feature learning method based on a camera model and hyperspectral image reconstruction.

  These contributions have resulted in achieving the highest accuracy of 88.34% on the test data. Furthermore, six papers have been published in conferences such as EUSIPCO, ICASSP, BMVC, and IEEE VR, and one chapter in the Springer CCIS journal. Additionally, one journal paper has been submitted to IEEE Trans. Image Process. and another conference paper has been submitted to AAAI.

Data Reduction for Supervised Training Feb 2020–Nov 2020

- This project is part of my MSc dissertation, supervised by Dr Cao Yang, at the Laboratory for Foundations of Computer Science (LFCS), University of Edinburgh.
- Surveyed on existing data reduction techniques for machine learning, focusing on instance selection algorithms and non- uniform sampling algorithms. Implemented the algorithms with scikit-learn and TensorFlow in Python.
- Analytically and empirically evaluated the relative accuracy of training over reduced datasets vs over original datasets.
- Proposed a new workflow to adapt the algorithms to work with image datasets and convolutional neural networks (CNN) by extracting image features first with pre-trained network parameters.
- Designed a new instance selection algorithm to weight the instances based on classification difficulty and the decision boundary in extracted feature space.

Publications

- 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

Patents

Contributed to the development of 1 invention patent, 2 utility model patents, and 1 design patent.

Awards

Overseas Students Pioneer Park; Creative Returnees Team	2017
2nd iCAN HongGuTan Cup VR/AR Innovation & Entrepreneurship Competition; Winning Team	2017
Hangzhou Overseas Returnees Innovation & Entrepreneurship Competition; Outstanding Young Returnees	2016

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 2020–2023.10

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

- 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 点云检测与图像检索大模型。
- 算法组重点产出：

  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。
  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。
  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。
  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。
  5. 真值系统：为百度云 roofbox 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。

研究经历

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 的半全局上下文特征学习方法；
  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 AAAI 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

创业经历

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

- 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；
- 工作简介：负责使用开发工具包 Vuforia 在 Java 环境中中开发 AR 软件；
- 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。

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