

# 王文彬

## 联系信息

手机: (+86) 166-1995-3902

邮箱: wenbin.wang@vipl.ict.ac.cn

个人主页: kennethwong.tech

联系地址: 北京市海淀区科学院南路 6 号, 中国科学院计算技术研究所, 100190

## 研究兴趣

计算机视觉, 主要研究方向为场景理解、场景图生成/视觉关系检测、物体检测、图像标题生成等

## 教育背景

中国科学院计算技术研究所, 计算机应用技术, 在读博士研究生 2017.09 - 2022.06 (预计)

导师: 陈熙霖研究员, 王瑞平研究员

南开大学, 计算机科学与技术, 工学学士 2013.09 - 2017.06

## 发表文章

- **Wenbin Wang**, Ruiping Wang, Shiguang Shan, Xilin Chen, Exploring Relationship Context and Visual Pattern for Scene Graph Generation, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8188–8197, Long Beach, CA, USA, 2019. (CCF-A, 一作)
- **Wenbin Wang**, Ruiping Wang, Shiguang Shan, Xilin Chen, Sketching Image Gist: Human-Mimetic Hierarchical Scene Graph Generation, In: Proceedings of the European Conference on Computer Vision (ECCV), LNCS 12358, pp. 222–239, 2020. (一作)
- **Wenbin Wang**, Ruiping Wang, Xilin Chen, Topic Scene Graph Generation by Attention Distillation from Caption, In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 15900–15910, 2021 (CCF-A, accepted, 一作).
- **Wenbin Wang**, Ruiping Wang, Xilin Chen, Sketching Image Gist: Importance-First Scene Graph Generation, submitted to IJCV.

## 项目经历

中国科学院计算技术研究所 | 博士研究生 2020.01-2020.09

- **基于移动机器人真实采集环境数据的三维场景理解与语义理解 (一期)**。具身智能日渐受到领域内关注, 并有潜力成为人工智能领域新的学习范式。而通往具身智能的第一步, 就是需要在真实环境中运动起来, 而不仅仅是靠在静态数据集学习。本项目内容是将组内研究的一系列场景理解算法在移动机器人上实现, 以实时采集的真实场景数据为输入。本人担任团队负责人, 同时具体负责利用前端感知数据进行场景语义理解的过程。具体完成: 对于 RGBD 视频帧根据模糊程度进行筛选并提取关键帧, 生成单帧场景图, 利用深度信息消除冗余节点, 合并单帧场景图, 生成全局的环境语义图。最终实现在真实环境下的涉及约 100 类物体、50 种关系的环境语义图生成。
- **开源场景图生成算法框架的开发与维护**。场景图生成领域缺乏统一的算法框架, 不利于多个方法的公平评测。基于物体检测系统, 开发与维护场景图生成算法框架, 将场景图生成的各个关键步骤抽象为若干可等效替换的关键模块, 目前已集成约十种主流场景图生成算法。同时集成一些可能用到

的图像标题生成、显著性物体检测等其他相关领域算法以及可视化的 API，为实现端到端的连续任务流、可视化分析等提供便利。

## 中国科学院计算技术研究所 | 实习生

2017.3-2017.5

- **场景标注工具开发及场景类别、属性数据集构建。**在场景理解任务中，各个子任务（如物体检测、场景分类、场景属性识别等）所需的数据集各自为政，缺乏统一的满足多个子任务的单一数据集，不利于开展统一的评测。为了满足全方位场景理解的需求，开展基于 MSCOCO 数据的场景类别和场景属性的标注采集工作。独立开发基于 MFC 的可视化标注工具客户端以及基于 PyQt 的可视化校验工具。利用这些工具，构建了涵盖 137 类场景、102 类场景属性的场景类别与场景属性数据集，包含约 60,000 张图像数据。该数据集与 MSCOCO 原有的物体检测等数据标注一同被应用于后续的场景理解相关研究中。

## 技术能力

---

- **编程语言:** Python, C/C++, Java
- **深度学习架构:** 熟练使用 pytorch, 了解并曾使用 Caffe、TensorFlow 等平台

## 奖项荣誉

---

- 中国科学院大学三好学生标兵, 2021 年
- 中国科学院大学学业奖学金一等奖, 2020 年
- 中国科学院计算技术研究所所长优秀奖, 2019 年
- 南开大学综合一等奖学金, 2016 年
- 北美数学建模大赛 M 奖, 2015 年

## 学术服务/其他

---

- Pattern Recognition (PR), Neural Computing & Applications (NCAA) 期刊审稿人
- CVPR 2022, ICCV 2021, ICME 2021, ICIG 2021, CICA 2021 会议审稿人。
- 博客 <https://www.cnblogs.com/Kenneth-Wong/>