

小红书-社区技术部-智能创作组-科研实习生招聘

岗位：多模态/AIGC/虚拟数字人-研究型实习生

部门：小红书，社区技术部，智能创作组

坐标：北京/上海均可，优秀同学可支持远程

团队介绍

小红书-社区技术部-智能创作组，是小红书从事创新型项目研发的团队，主要从事计算机视觉、多模态相关方向研究和业务落地。目前团队共有 10 余位科研实习生，均为在读博士或硕士，分别是来自马普所、MBZUAI、莫那什、清华、北大、浙大、自动化所、华科、北航、北理工、南理工等知名高校的在读博士或硕士，从事虚拟数字人、多模态理解与生成方向的研究，并已有多篇投稿。

岗位职责

1. 利用前沿计算机视觉、计算机图形学算法，研究数字人或多模态相关的课题。
2. 在顶级会议（CVPR/ICCV/ECCV/SIGGRAPH/3DV 等）上发表学术论文。
3. 与德国马普所研究人员进行深度合作研究。

研究内容

1. 虚拟数字人方向（与马普所合作）

- a. 通用三维场景/物体生成 (<https://mime.is.tue.mpg.de/>)
- b. 三维场景与人体运动重建 (<https://mover.is.tue.mpg.de/>)
- c. 高保真三维数字人生成/重建 (<https://elicit3d.github.io/>)
- d. 有条件下三维人体运动生成和预测
 - i. 音乐-舞蹈生成: <https://www.mmlab-ntu.com/project/bailando/index.html>
 - ii. 语音-手势生成: <https://talkshow.is.tue.mpg.de/>
 - iii. 文本-动作生成: <https://guytevet.github.io/mdm-page/>
- e. [工程方向] 虚拟引擎/Unity/Unreal 驱动下人体运动动作生成

2. 多模态理解与生成（面向真实场景/多任务统一建模）

- a. 多模态（图像/视频/语言）基础模型、自监督/弱监督预训练
 - i. 高效多模态表征预训练 [BEiT-v3](#) [EVA](#) [FLIP](#)
 - ii. Instruct 跨模态交互学习 [InstructGPT](#) [ChatGPT](#)
- b. 可控的跨模态通用视觉内容生成
 - i. 知识增强的多模态生成预训练 [\[2210.15257\]](#) [ERNIE-ViLG 2.0](#)
 - ii. 文本到图像语义一致性增强 [Stable Diffusion v2.1](#)

岗位要求

1. 有较强的自我驱动力，热爱科研，熟悉以上任意领域的研究；
2. 熟练使用 Python 或 C++，熟练至少一种深度学习框架，如：PyTorch；
3. 有发表过相关顶级会议或期刊（CVPR/ICCV/ECCV/NeurIPS/ICML/ICLR/TPAMI 等）或者参与过知名开源项目的同学优先；
4. 实习时间大于六个月以上优先，北京/上海均可，优秀同学可支持远程；

投递方式

有意向同学可发送简历至：wanghaofan@xiaohongshu.com，简历评估通过后我们会在一周内联系，招聘长期有效