

AI驱动下的XR内容生成

田第鸿 博士 (扩展现实研究中心主任)

刘 聪 (演讲人)

2021 年 07 月



关于AIRS

关于AIRS

- ▲ 深圳十大基础研究机构之一
- ▲ 香港中文大学（深圳）作为发起单位，联合多个世界顶级研究机构
- ▲ 人工智能与机器人领域的理论基础与应用基础研究的研究院
- ▲ 集政府引导、国际联合、大学依托、企业合作、创投融资、双创孵化等多方面于一体的研究院

AIRS的目标

- ▲ 建立一所国际性的开放型研究院，让世界各地的学者在AIRS享受研究
- ▲ 聚集有潜力的青年学者，培养未来的科研领导者
- ▲ 促进人工智能与机器人技术的产业化
- ▲ 创建全球人工智能与机器人技术的核心城市



扩展现实 (XR) 研究中心

- 扩展现实（包括虚拟现实、增强现实、混合现实等）是未来的个人计算与人机交互发展方向。扩展现实研究中心致力于研究和开发扩展现实中计算成像、显示与可视内容生成，以及端云协同处理的关键技术与系统，研究扩展现实与人工智能、计算机视觉技术的结合，开发扩展现实在智能制造、远程协作、教育培训领域的示范应用系统并推动其产业化。



AI驱动下的XR内容生成

高度逼真的三维可视内容是沉浸式可视交互的核心，而目前三维视觉内容的生成一方面依赖激光扫描等复杂的高精度三维获取设备，另一方面需要内容创作者们通过专业的三维内容设计工具进行大量的开发工作去构建高清晰的视觉内容。**本项目探索通过人工智能（AI）和沉浸式Web技术实现快速、便捷生成XR内容。**

云端实现AI与沉浸式计算+终端负责交互



云+端架构示意图

#未来五年这些行业值得关注#

数字经济重点产业

区块链

推动智能合约、多重共识算法、非对称加密算法、分布式容错机制等技术创新，以联盟链为重点发展区块链服务平台和金融科技、供应链管理、政务服务等领域应用方案，完善监管机制

人工智能

建设重点行业人工智能数据集，发展算法推理训练场景，推进智能医疗装备、智能运载工具、智能识别系统等智能产品设计与制造，推动通用化和行业性人工智能开放平台建设

虚拟现实和增强现实

推动三维图形生成、动态环境建模、实时动作捕捉、快速渲染处理等技术创新，发展虚拟现实整机、感知交互、内容采集制作等设备和开发工具软件、行业解决方案

数字社会建设

- 提供智慧便捷的公共服务。聚焦教育、医疗、养老、抚幼、就业、文体、助残等重点领域，推动数字化服务普惠应用
- 建设智慧城市和数字乡村。以数字化助推城乡发展和治理模式创新，全面提高运行效率和宜居度
- 推动购物消费、居家生活、旅游休闲、交通出行等各类场景数字化
- 将数字技术广泛应用于政府管理服务，推动政府治理流程再造和模式优化

两会 2021

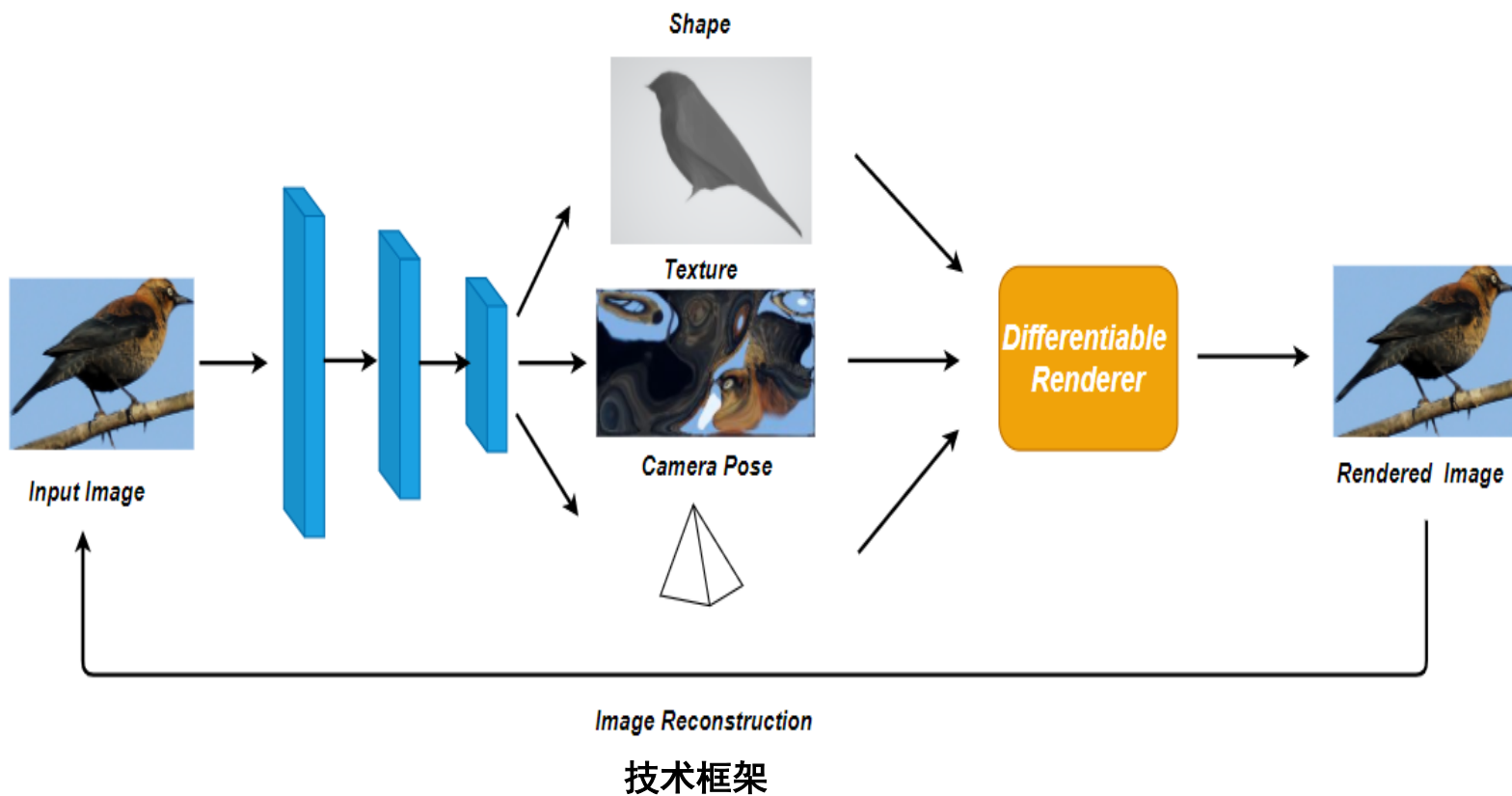
人民日报

AI驱动XR内容生成

- 语义辅助单视角/多视角三维模型生成
- 基于语义识别和搜索的XR内容生成
- 动作迁移与动态XR内容生成
- 扩展现实内容生成平台

基于语义辅助的三维模型生成

单视角图像



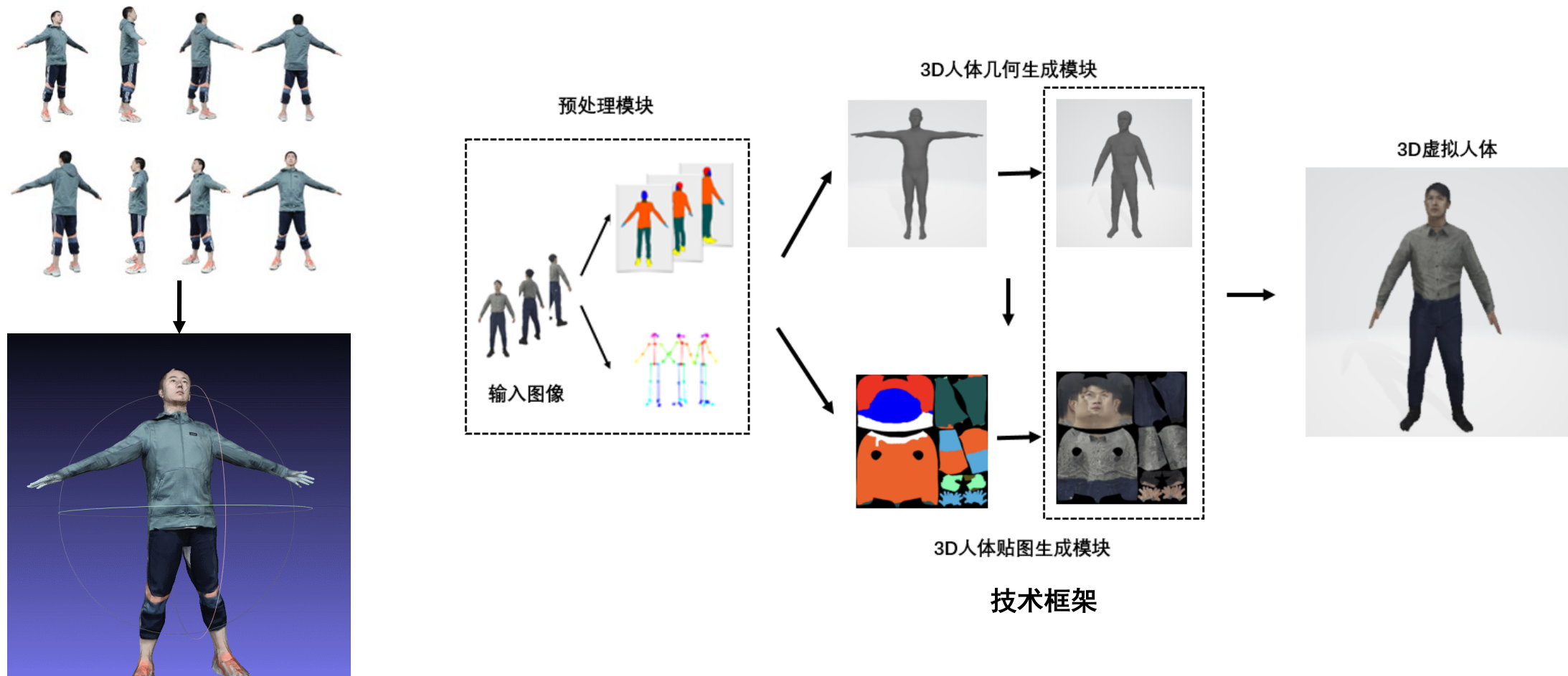
基于语义辅助的三维模型生成



实验结果

基于语义辅助的三维模型生成

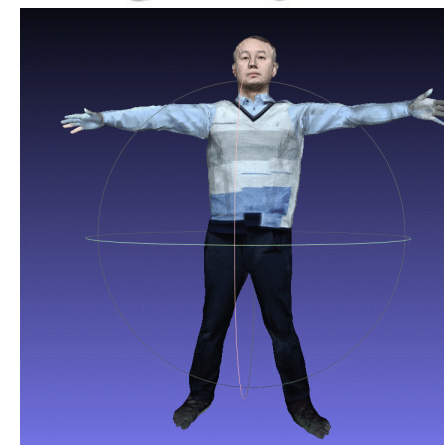
多视角图像驱动高精度三维人体模型生成



基于语义辅助的三维模型生成



生成效果



基于语义辅助的三维模型生成

单视角图像生成人体模型



输入单张图像



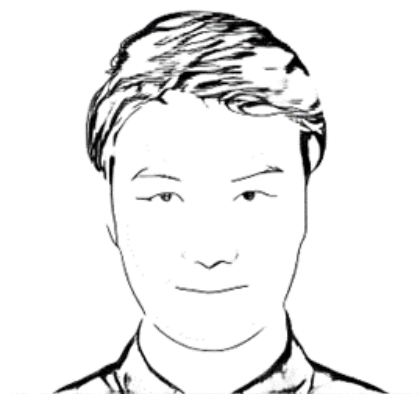
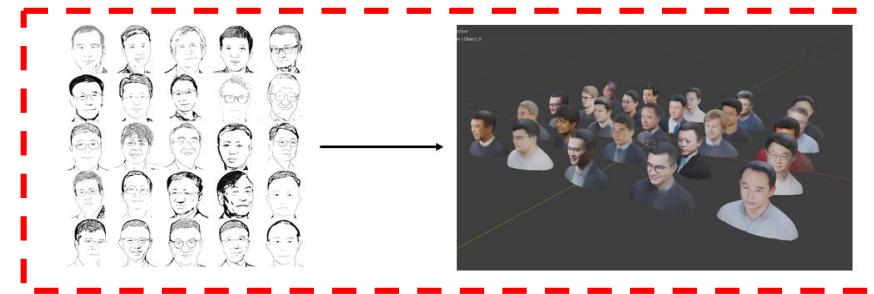
模型重建结果



AR展示结果

基于语义辅助的三维模型生成

Digital Avatar (可操纵的虚拟人像)



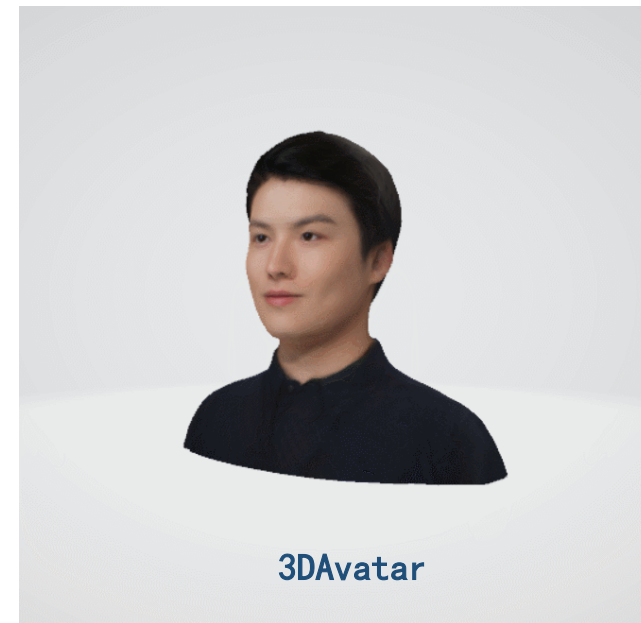
2DSketch

Step1: 生成2DAvatar



2DAvatar

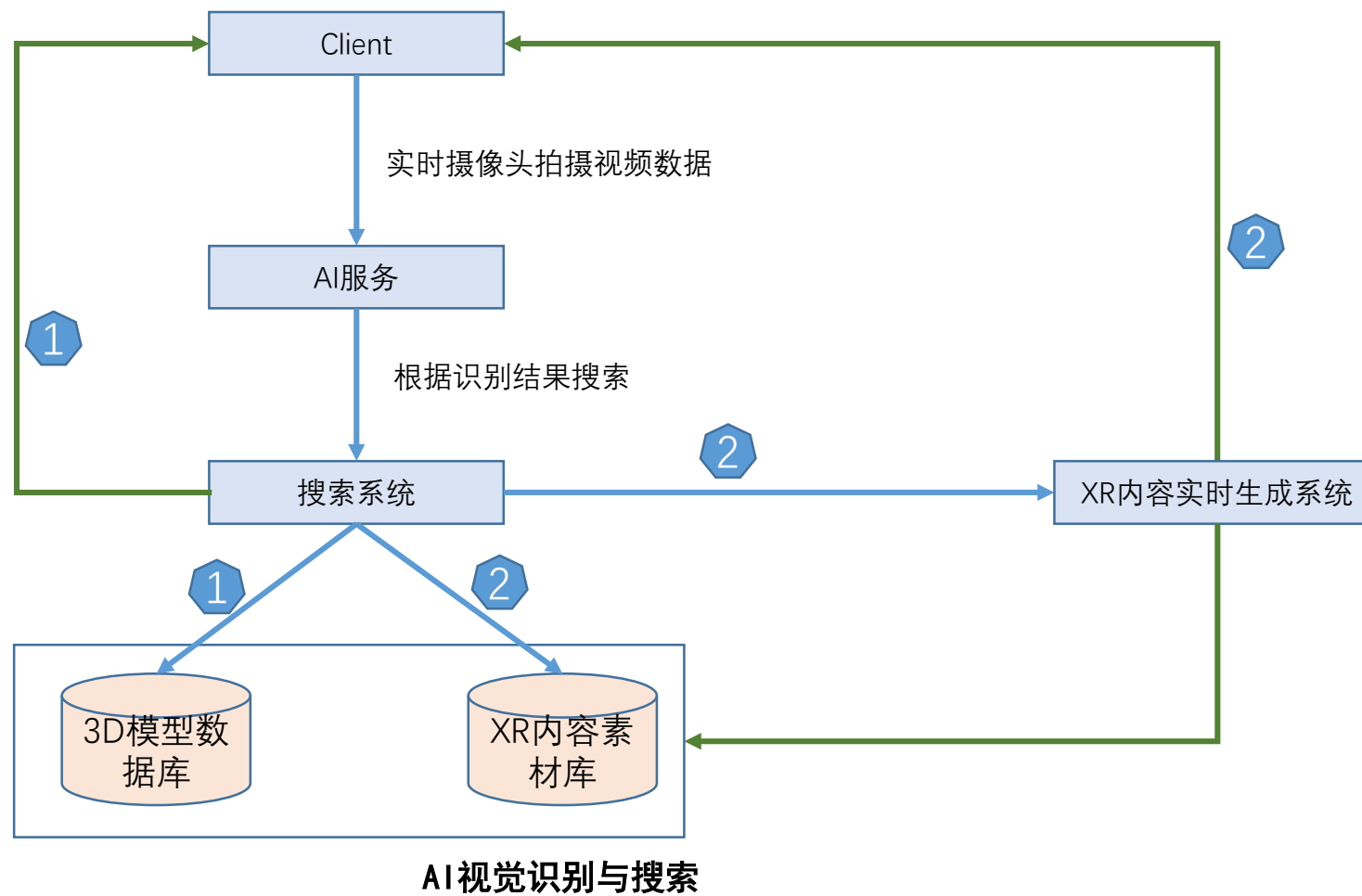
Step2: 生成3DAvatar



3DAvatar

可操纵虚拟人头生成流程（单视角）

基于语义识别和搜索的XR内容生成



基于视觉识别和搜索的XR内容生成

所见即所得



2020龙岗区优质文化项目（AR音乐会）

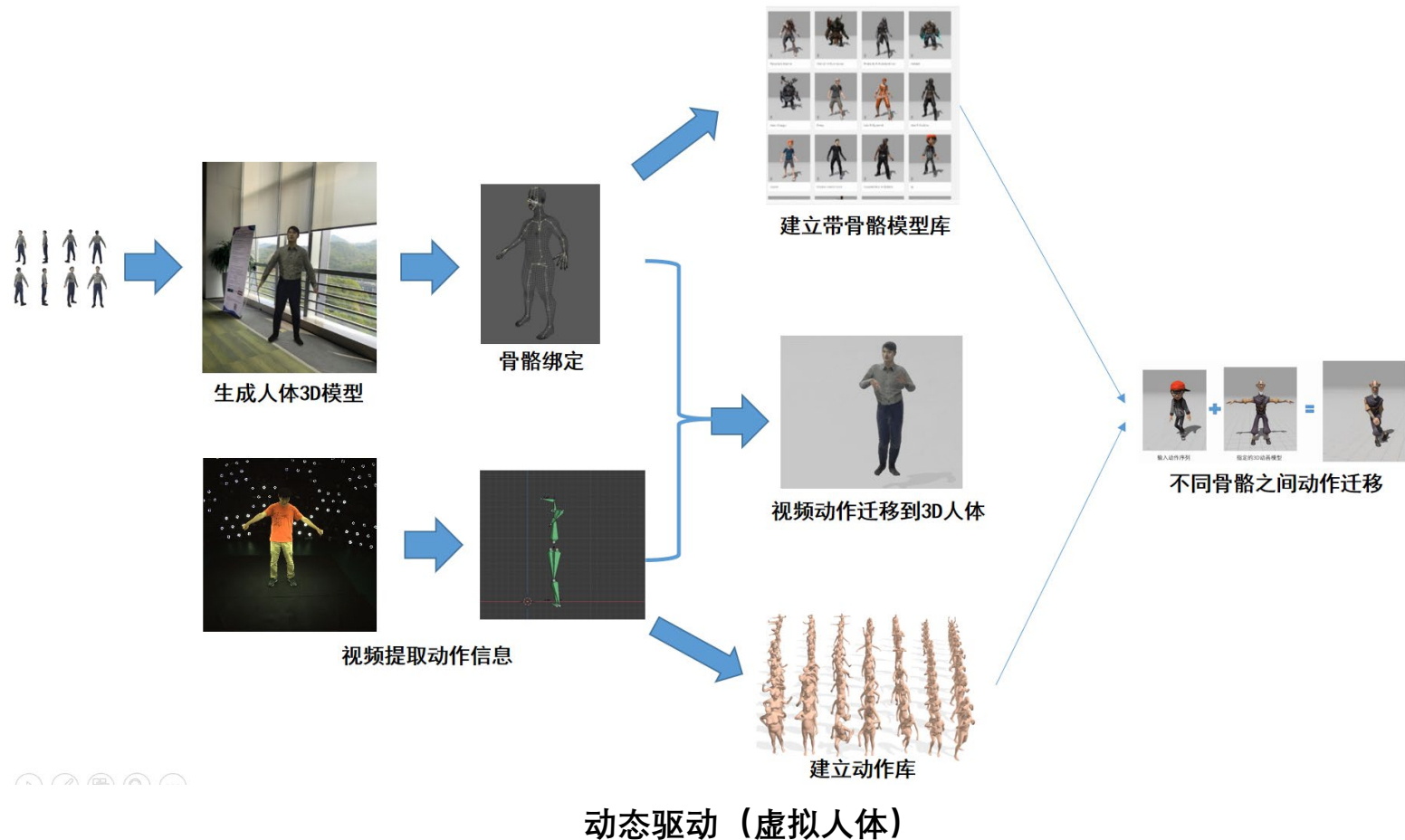


成果演示

AI驱动下的动态XR内容生成

动态驱动

- 视频提取动作信息
- 不同骨骼间动作迁移
- AI驱动人形模型骨骼绑定
- 动画驱动（人物模型库与动作库）
- 实时动画驱动

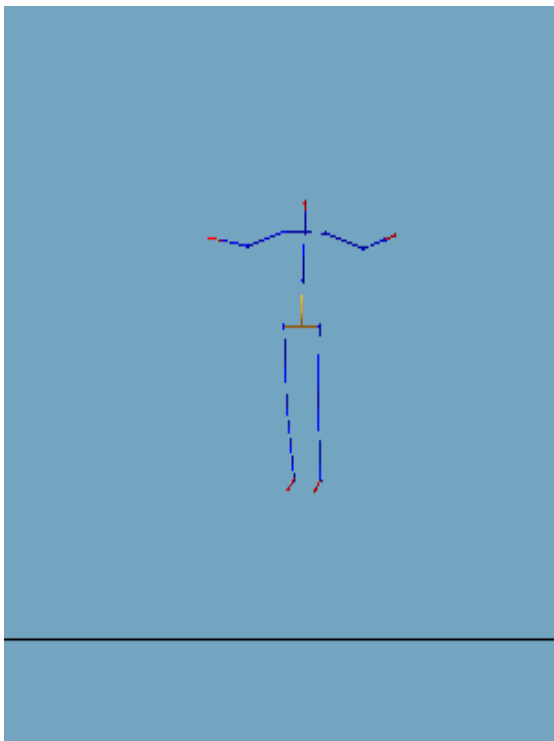


AI驱动下的动态XR内容生成

动态驱动（视频提取动作信息）



输入视频



提取的动作

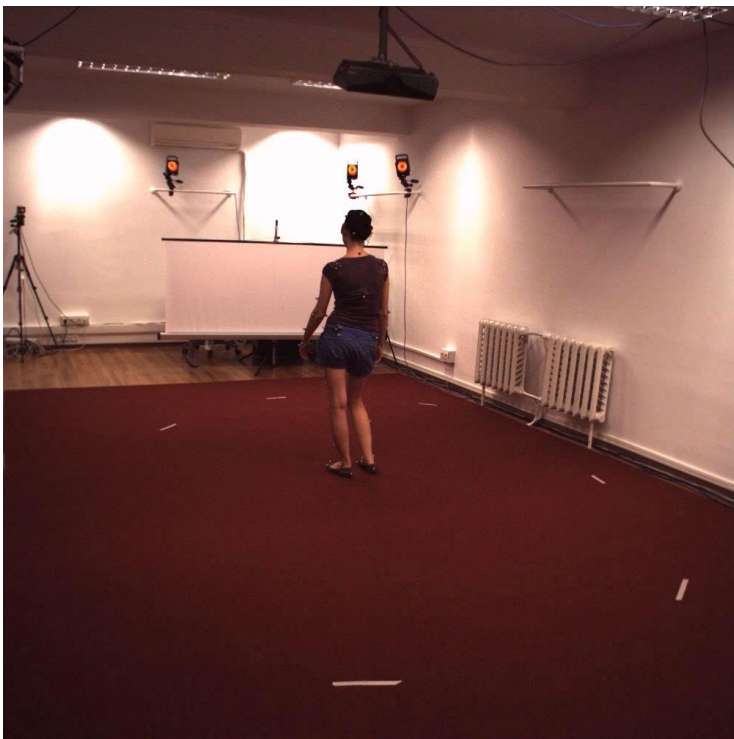


渲染结果

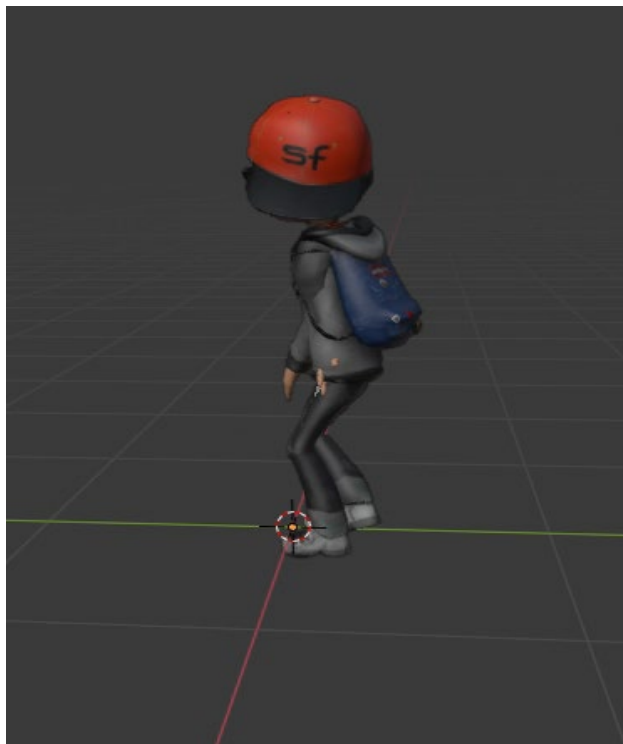
➤ 通过AI方式提取视频中人物动作信息（骨架动作信息）

AI驱动下的动态XR内容生成

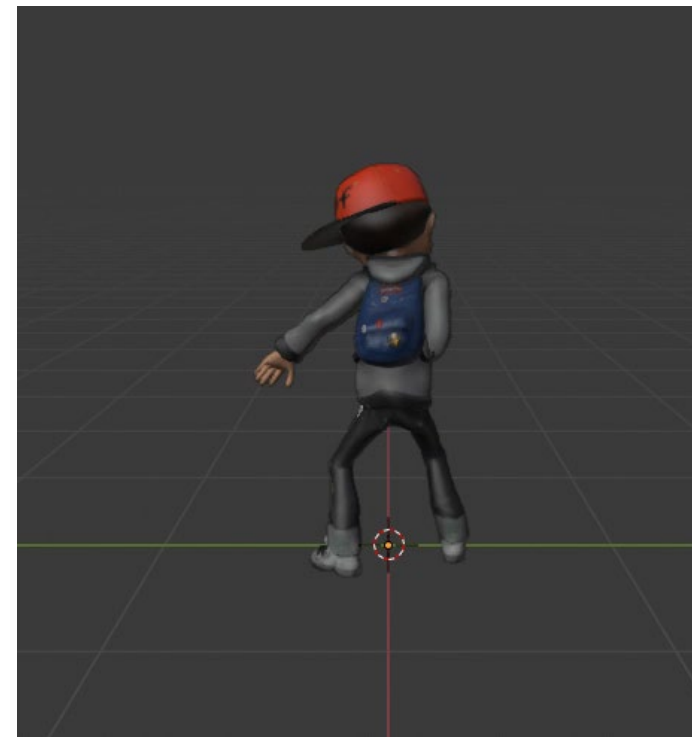
动态驱动（不同骨骼间动作迁移）



输入视频



原始网络渲染结果

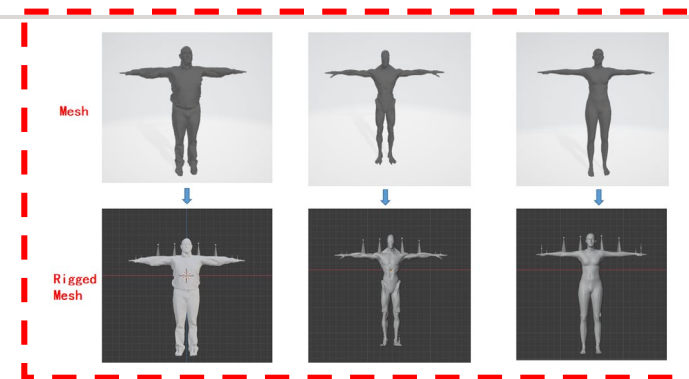


自适应池化渲染结果

- 通过自适应池化算法解决不同骨骼模型之间的动作迁移

AI驱动下的动态XR内容生成

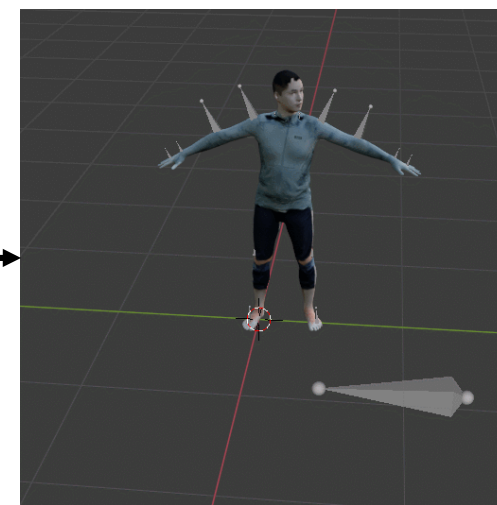
动态驱动 (AI驱动骨骼绑定)



生成的模型



AI驱动骨骼模型生成



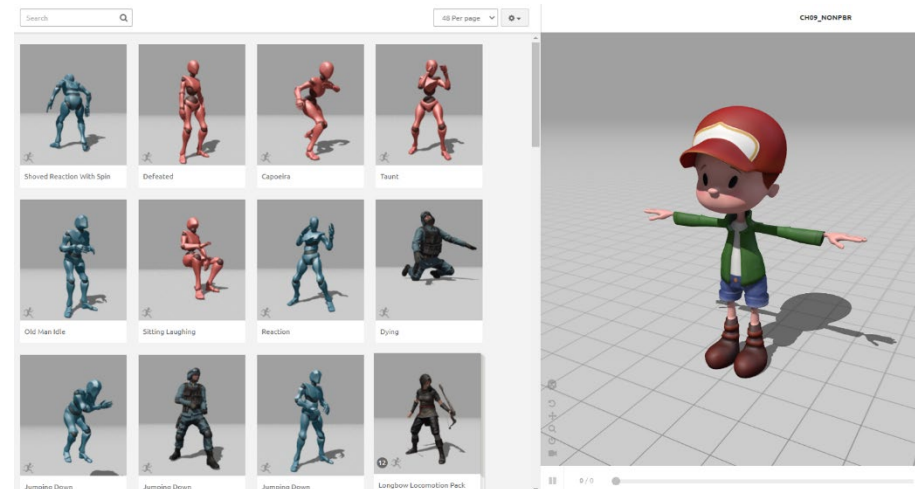
模型带动画

AI驱动下的动态XR内容生成

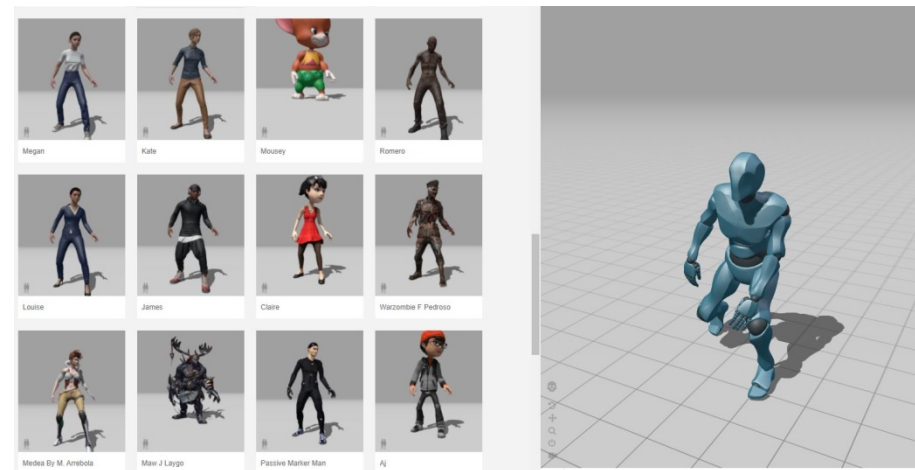
动态驱动



动态驱动生成模型



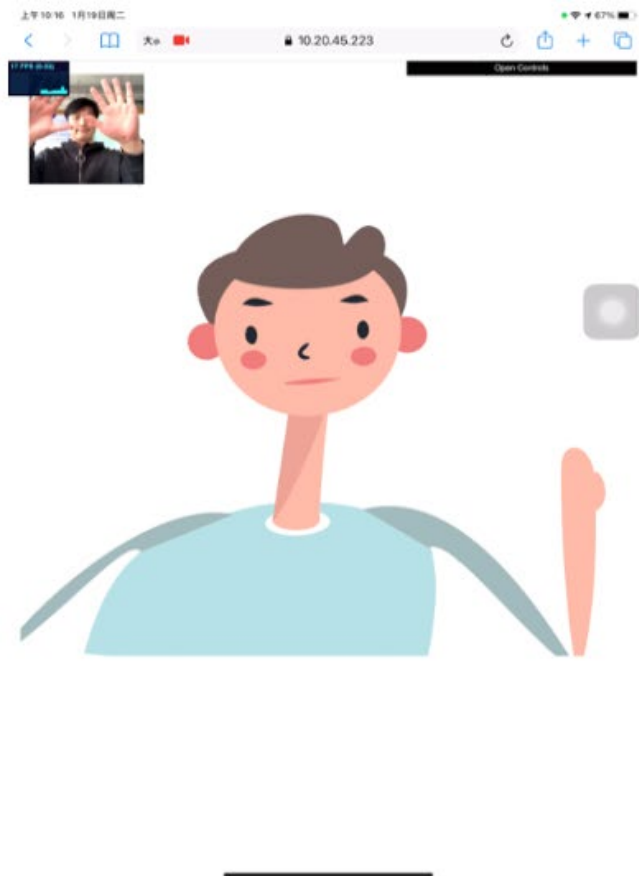
动作信息库



动画模型库

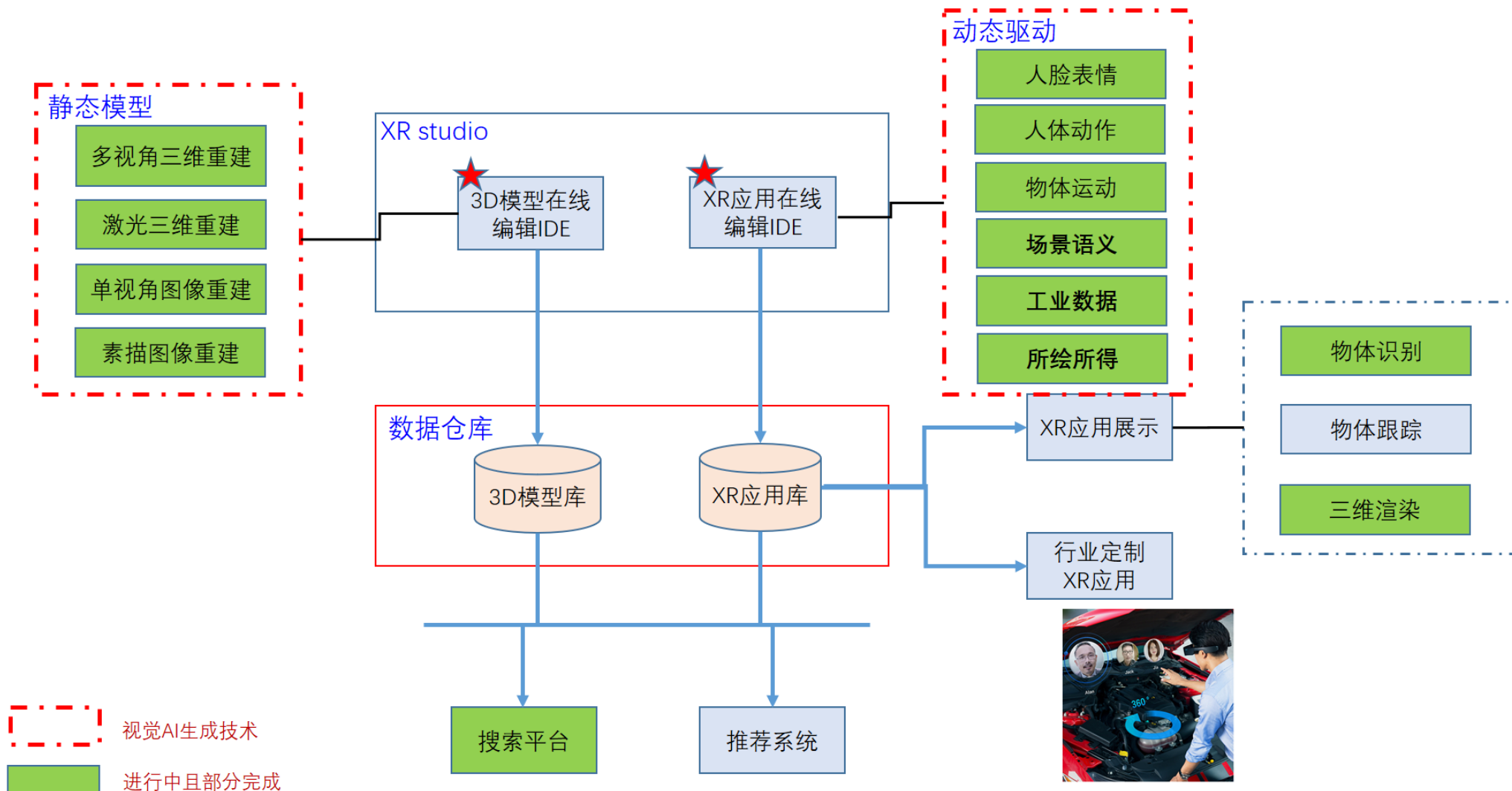
AI驱动下的动态XR内容生成

动态驱动（实时）



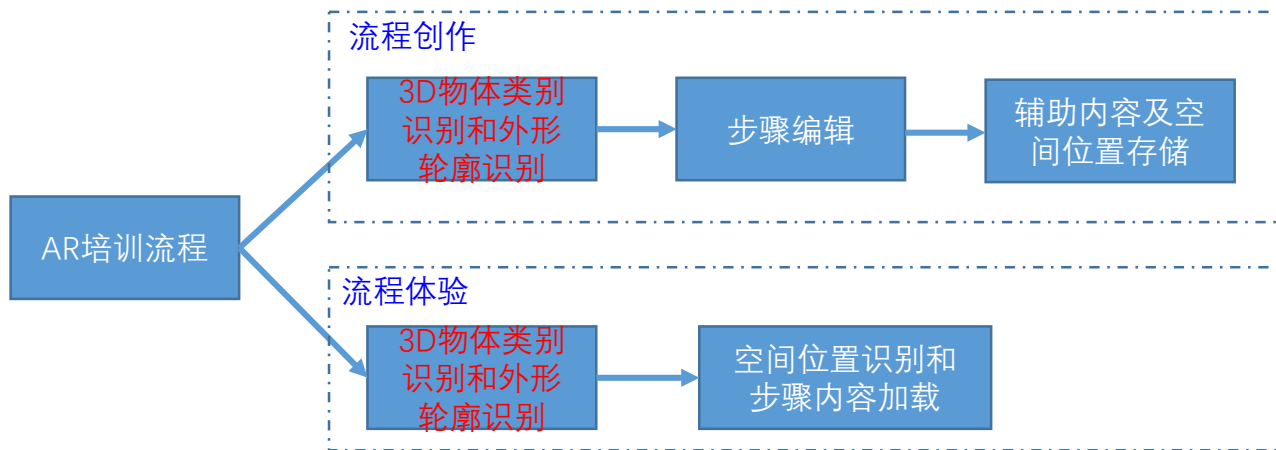
人的实时动作捕捉（2D/3D）

扩展现实内容平台



平台应用场景

交互编辑



基于WebXR的交互编辑（咖啡机3D说明书）

平台应用场景

工业远程操纵（平行扫描与机器试校场景）

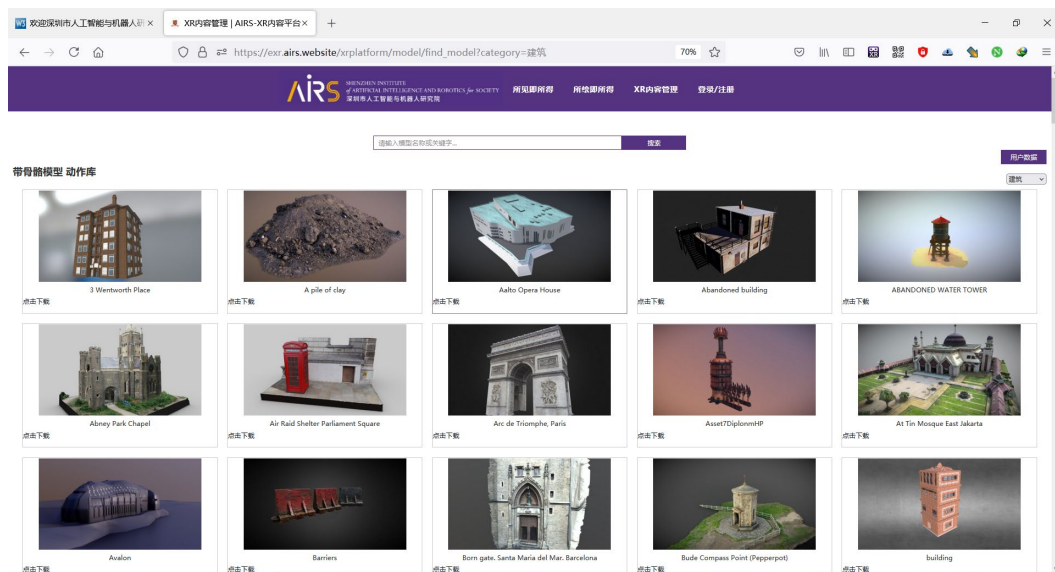


平台应用场景

工业多机协同与远程协助场景



XR + 沉浸式Web



<https://exr.airs.website/xrplatform>



<https://www.chinaw3c.org/archives/2586/>

- 基于AI生成技术搭建了一套的扩展现实内容生成的Web平台
- AIRS正成为**W3C会员单位**及**粤港澳大湾区代表处**，将联合并汇集扩展现实沉浸式网络领域的产业各方推动围绕**WebXR**、**WebML**等**Web技术标准化**工作

欢迎合作!

liucong@cuhk.edu.cn

Thanks!

AIRS