

基于生成式先验的人像视频编辑

汪兆辰

November 18th , 2024

总体设想

基于文章 Portait Video Editing Empowered by MultiModal Generative Priors 提出的方法，开发一个 Web 客户端实现对用户上传的人像视频的编辑，包括：

- 文字驱动的视频编辑
- 图像驱动的视频编辑
- 改变全局光照

- 更进一步地，该工作使用了多种先验模型：在对人体建模时使用了 SMPL-X；而由于文章中提出的视频生成策略为使用 2D 生成模型迭代更新视频帧数据集，再调整 3DGS 渲染出结果，因此可以考虑替换效果更好的先验模型，以提高视频编辑的效率和质量。
- 另一方面，在文章的演示中，该方法主要被应用于风格化处理视频和衣着替换，可以考虑开发更多的功能。

目前进展

- 正在学习 MIT Web Development 课程，预计本周完成，之后将着手搭建网站基本架构。
- 调研了一些相关最新文献，考虑在后续添加新功能

问题

