



歌宋 · 忆龄

赛后分享会
南天尚音团队
11月4日

永远和
党
在一起

宋庆龄与中国福利会

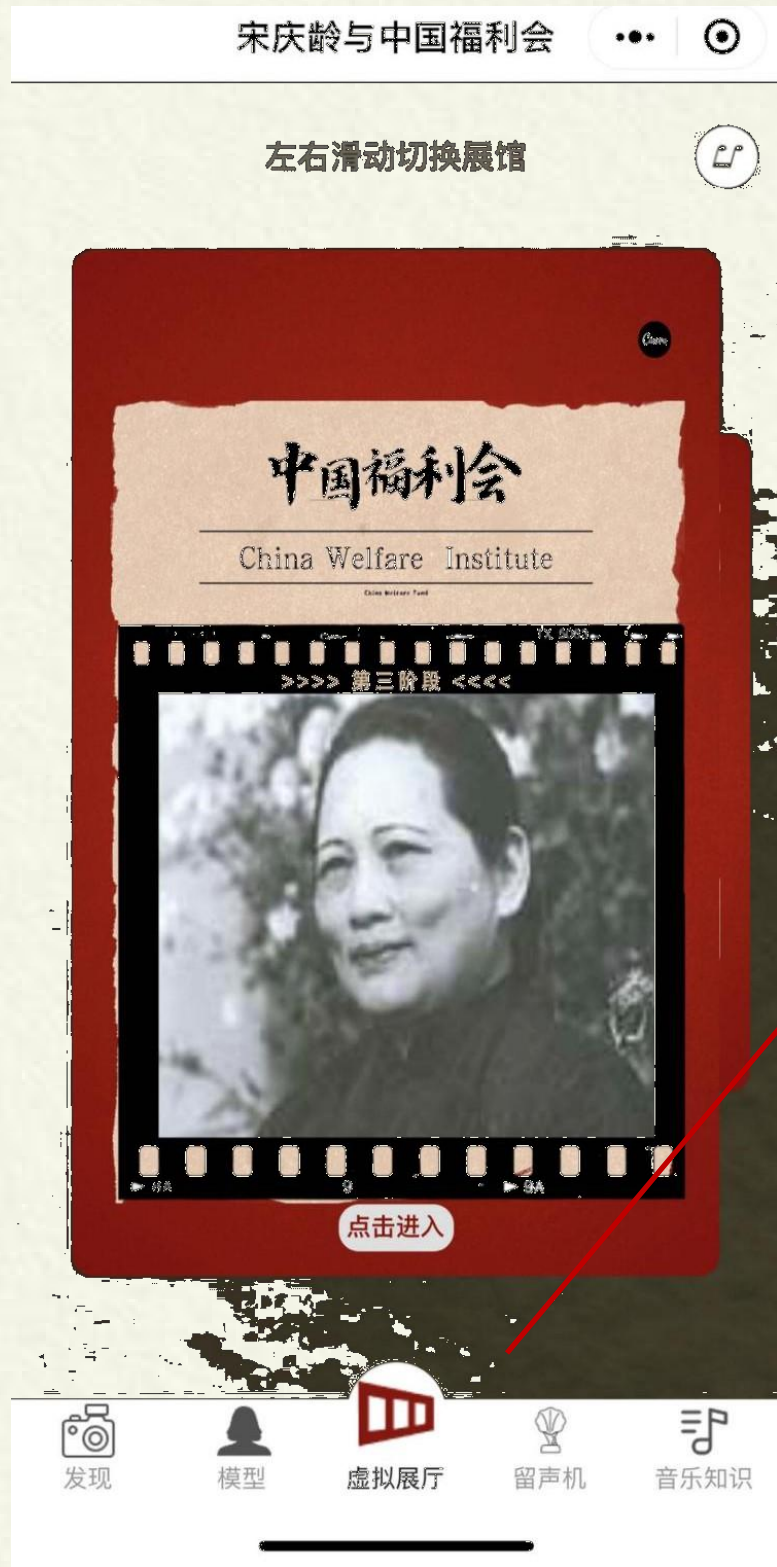
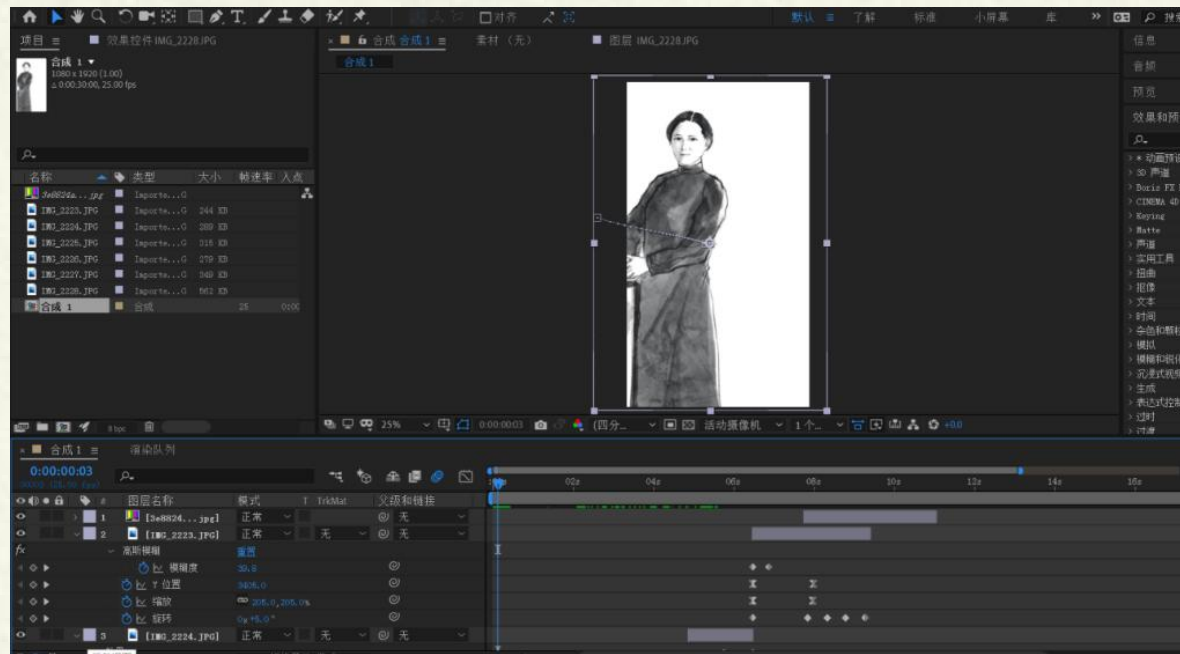
宋庆龄故居

目录

DIRECTORY

- 一 “歌宋·忆龄”作品介绍
- 二 核心技术与实现流程
- 三 解决的问题
- 四 备赛经验与参赛感想

— 小程序功能展示



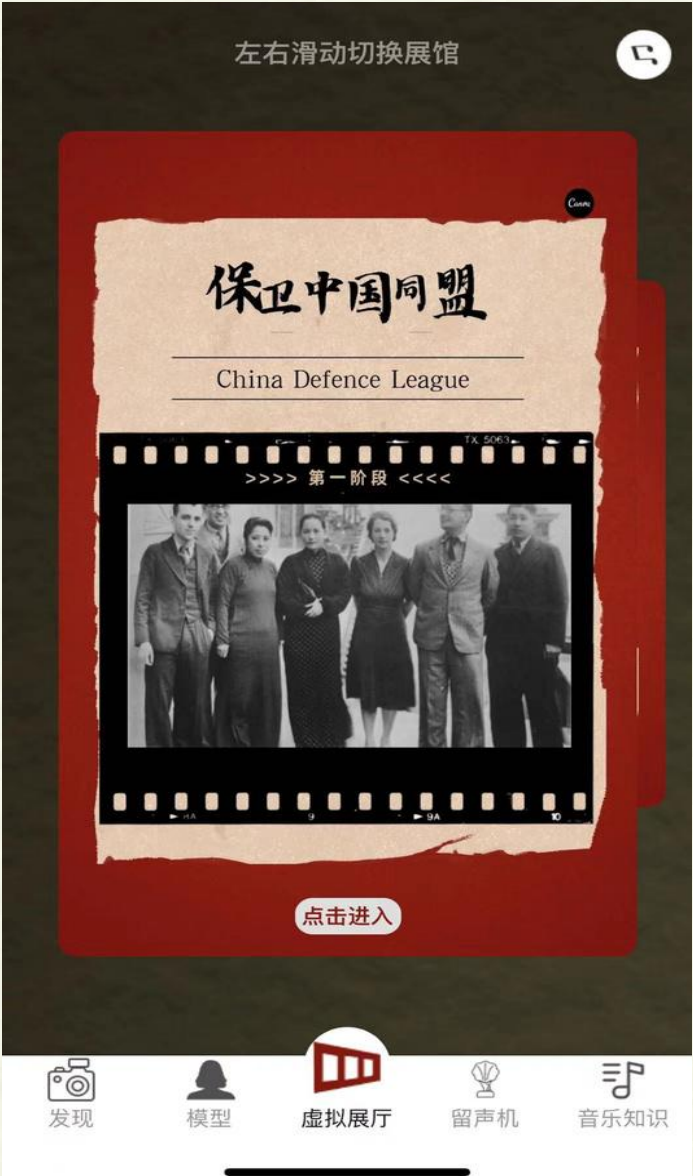
让记忆融入音乐
让音乐留存记忆

小程序五大功能版块

- 发现：历史推荐
- 模型（数字人）：原声再现
- 虚拟展厅：沉浸体验
- 留声机：原创音效
- 音乐知识：知识普及

— 小程序功能展示：VR虚拟展厅

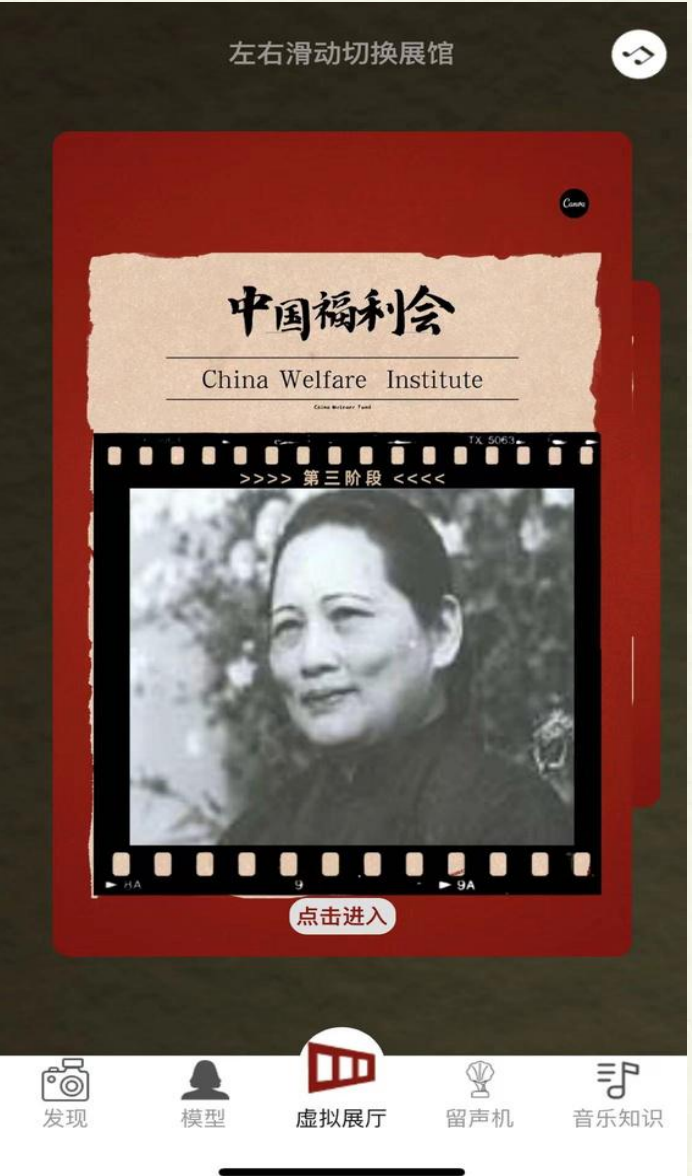
展厅一



展厅二



展厅三



— 小程序功能展示：无损全景声音效



- 选择场景，快速漫游
- 按照指引，自主漫游
- 设计路线，自动漫游



根据场景为展厅中每张图片
自主创作独特的
无损、全景声音效



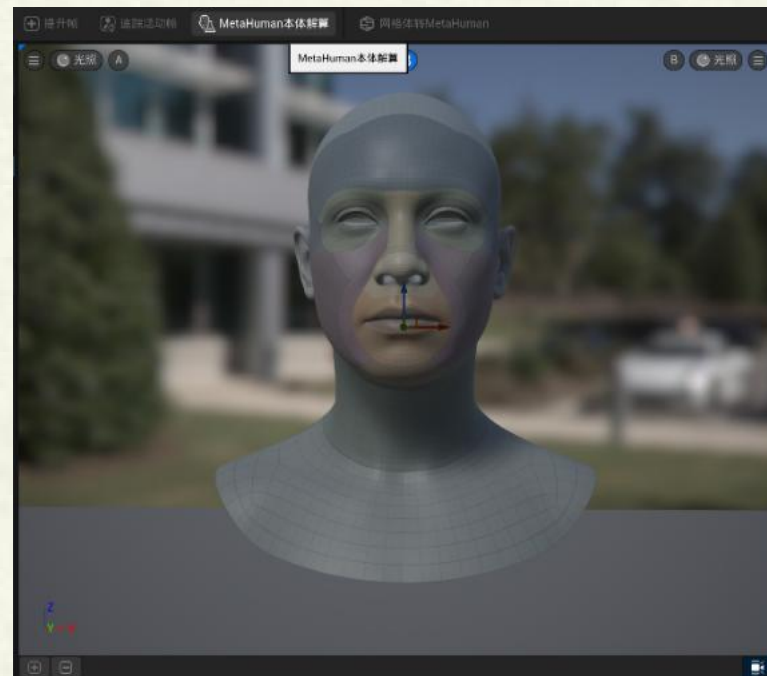
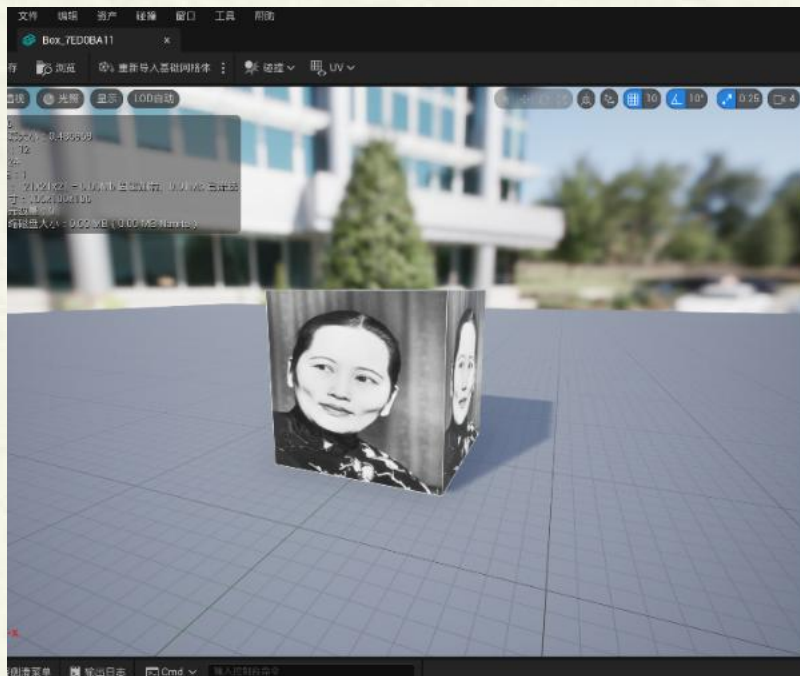
小程序功能展示：数字人

模型——宋庆龄女士的虚拟形象：原声再现

- 通过一张照片，在虚幻引擎5.0.3中进行了静态网格体转MetaHuman的处理，借助超人类本体解算技术追踪面部轮廓，并进行五官雕刻。
- 同时，使用MockingBird提取了宋庆龄女士说话声音的音色向量，结合声码器完成语音合成，实现了宋庆龄女士在虚拟世界中的重现。



点击
体验



— 小程序功能展示：其他



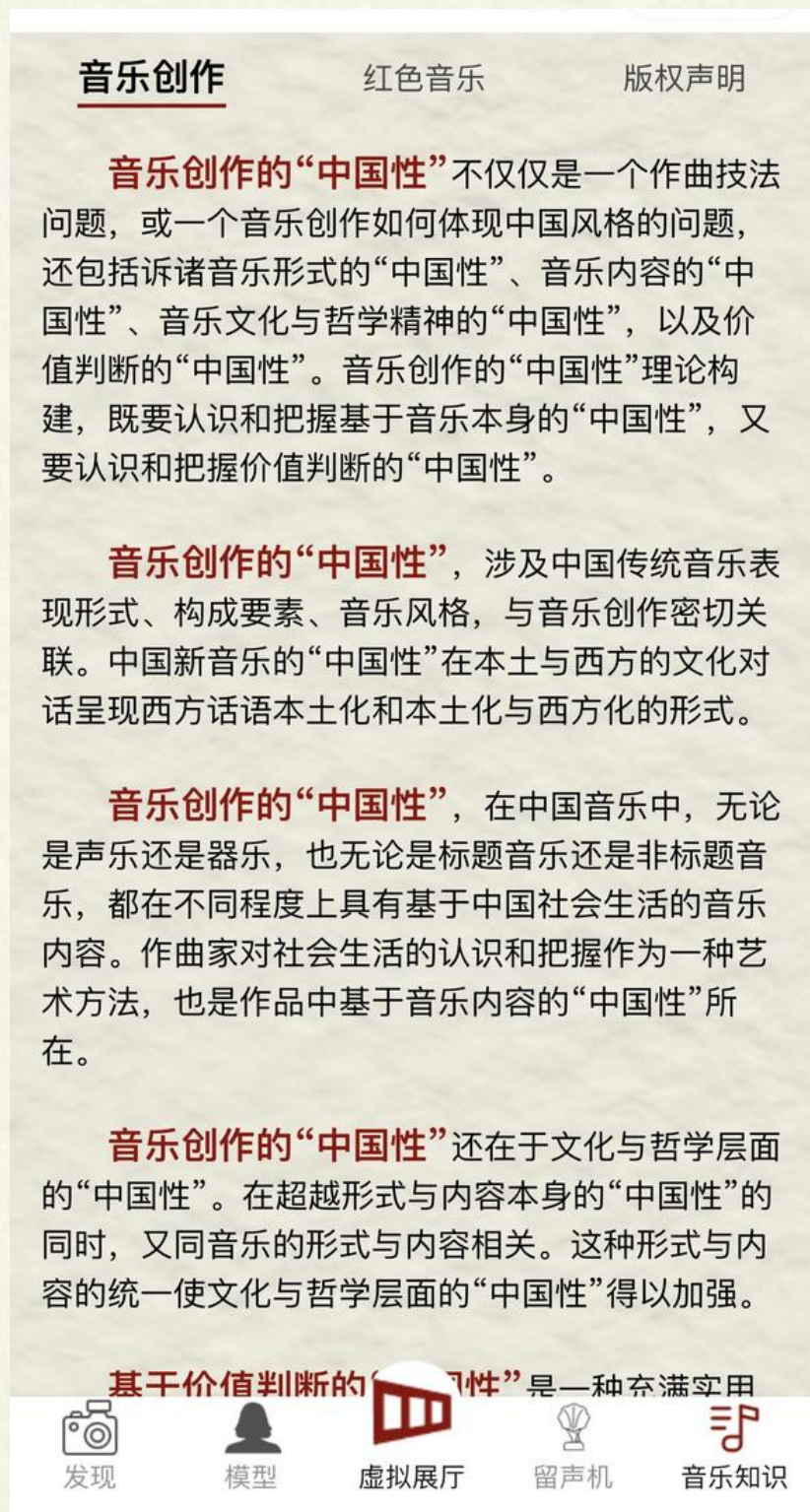
发现

历史上的今天、随机推荐



留声机

为展品图片量身定制音效与音乐



音乐知识

中国特色音乐创作和红色音乐

二 核心技术及实现流程：音乐设计

制作构思

构思整体交互框架制
作音频demo

解读图片

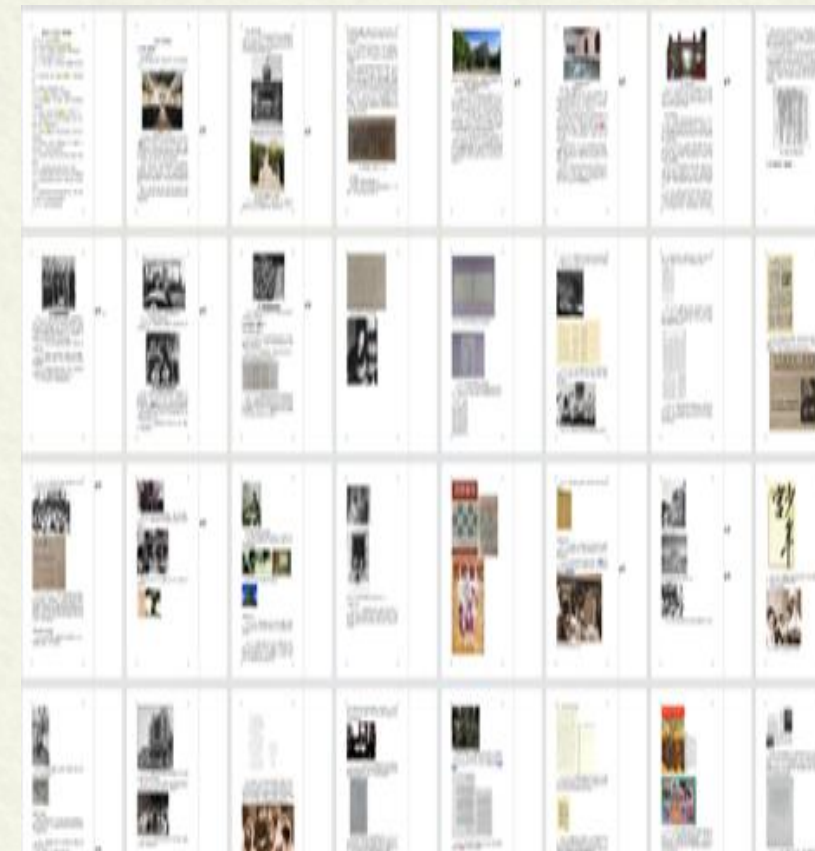
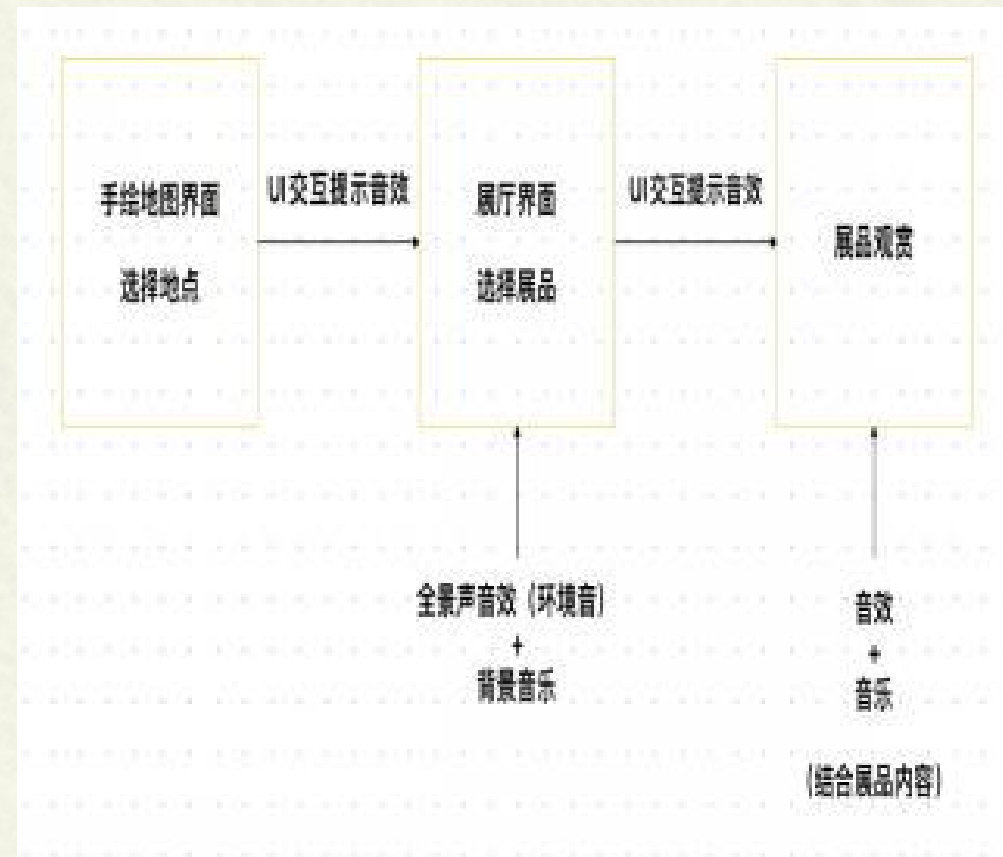
分析数据资料
解读图片

收集声音素材

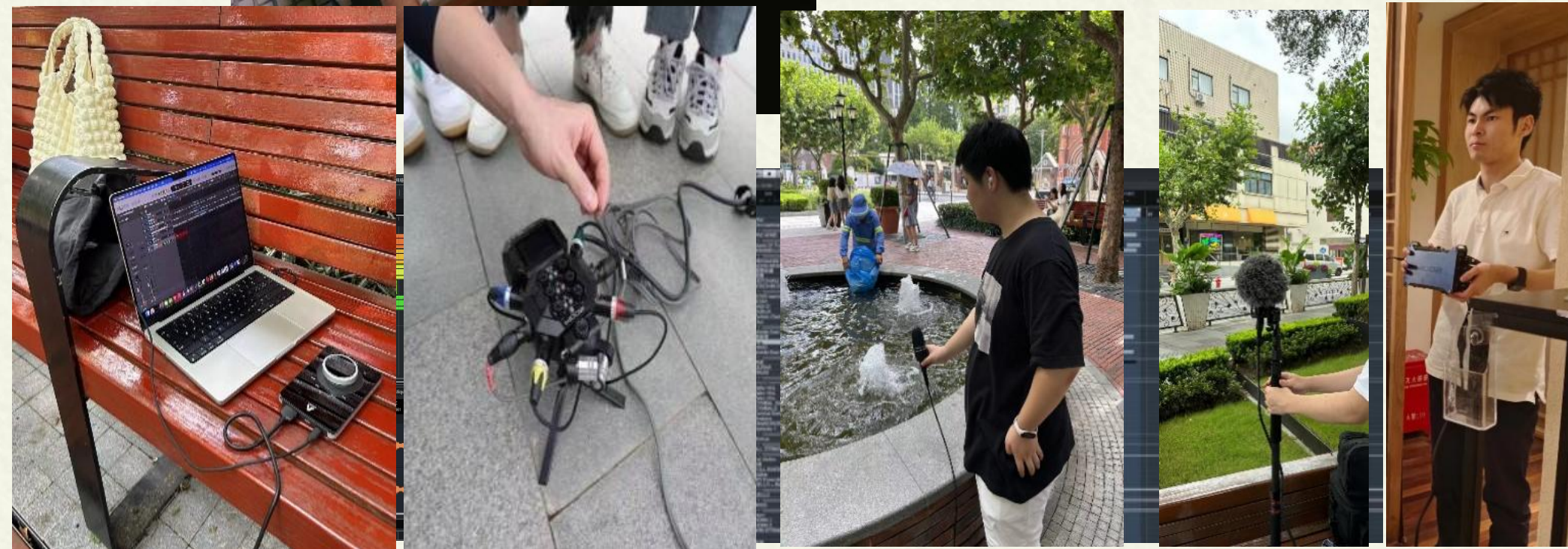
整合音效与音乐

剪辑与混音

导出为mp3、wav以及
全景声混音版三种版本



核心技术
全景声录音与混音



二

核心技术及实现流程：模拟人声制作流程

核心技术

Mockingbird

模拟人声



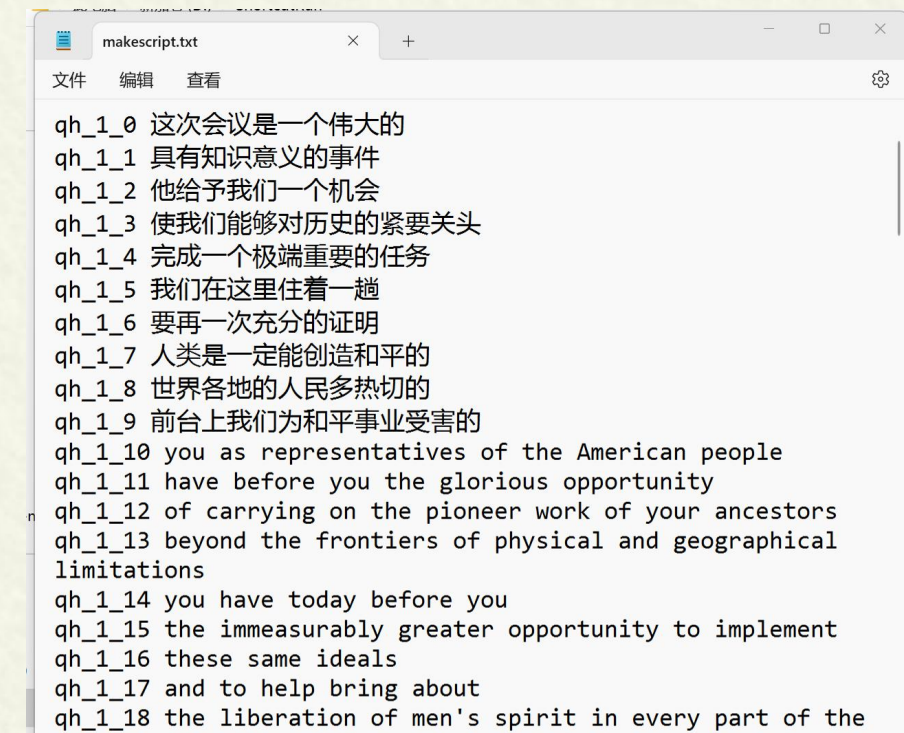
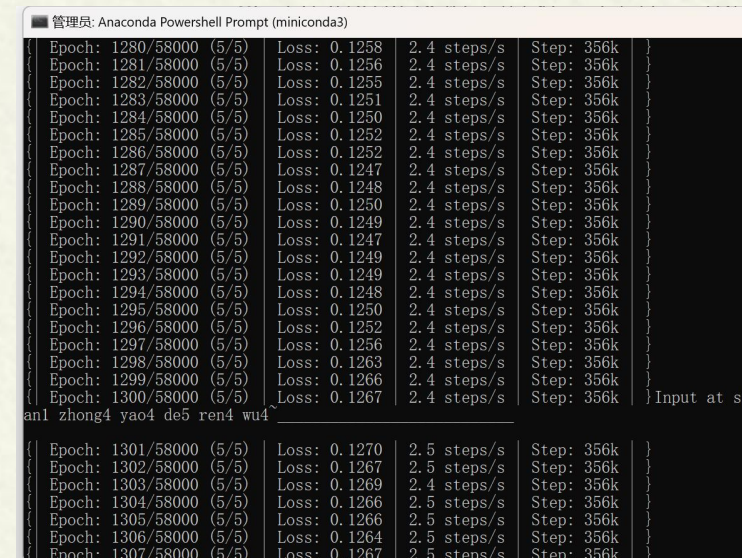
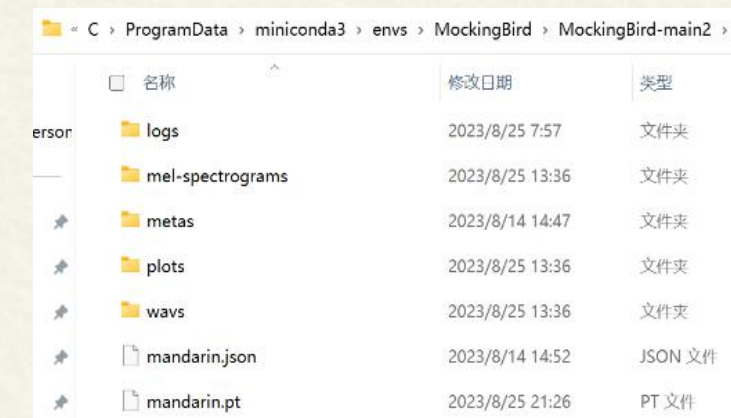
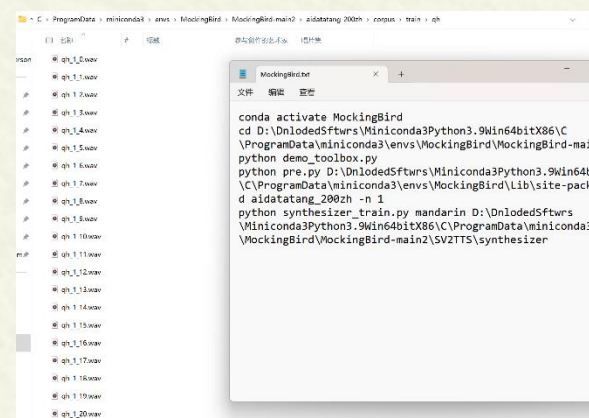
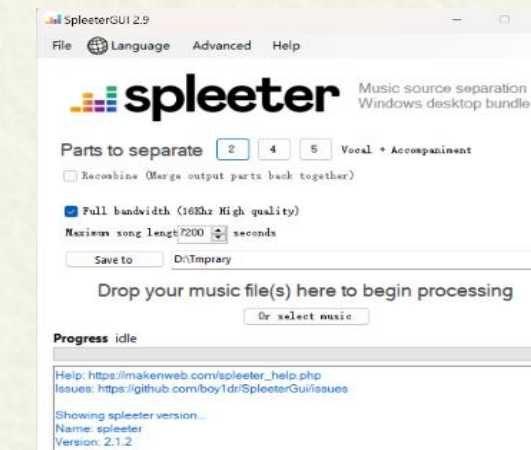
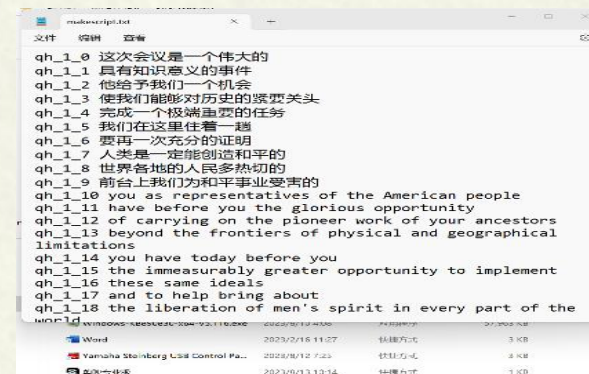
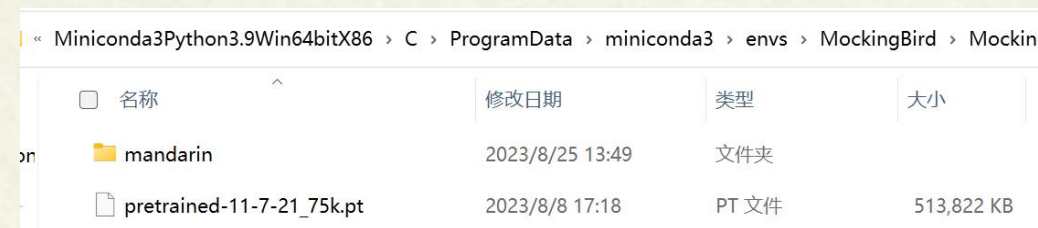
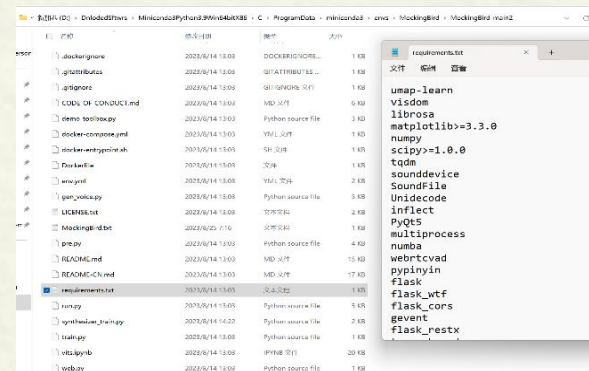
使用Anaconda和Python3.9搭建环境

将宋庆龄生前音视频用spleeter去噪后生成文本

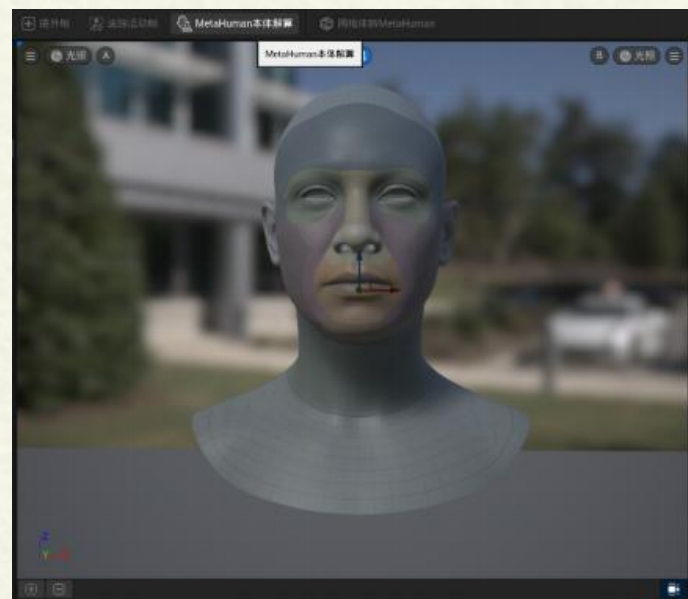
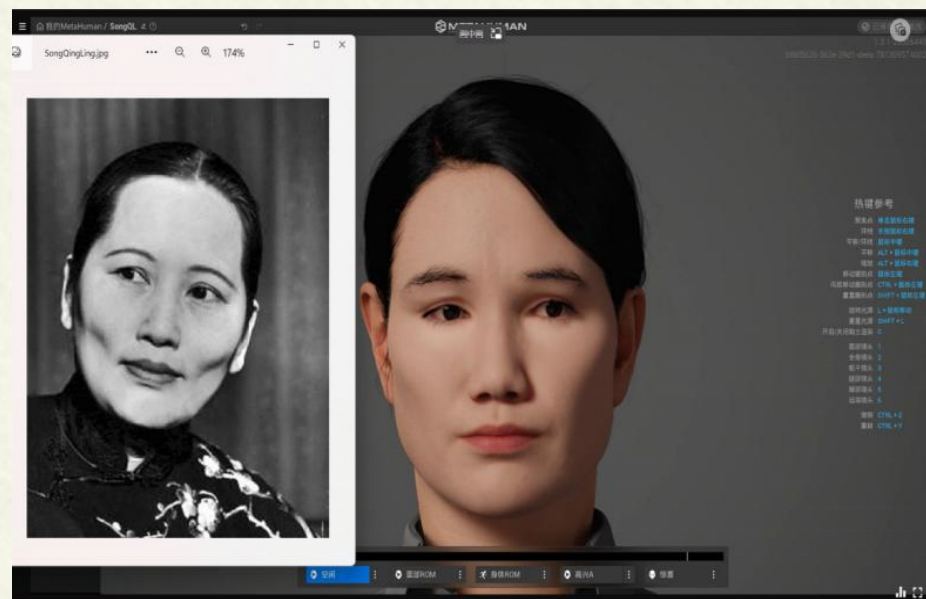
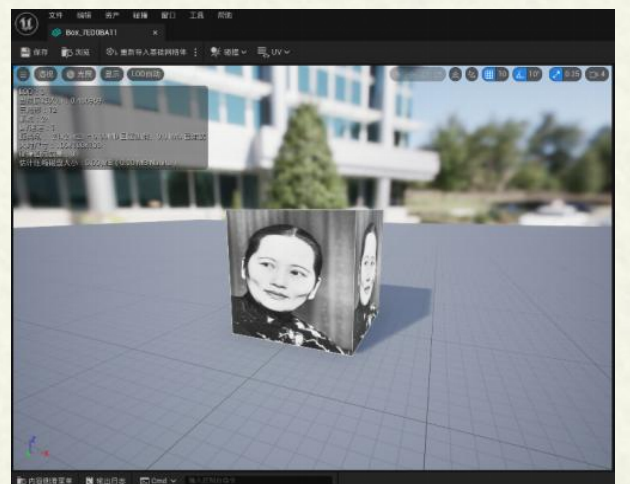
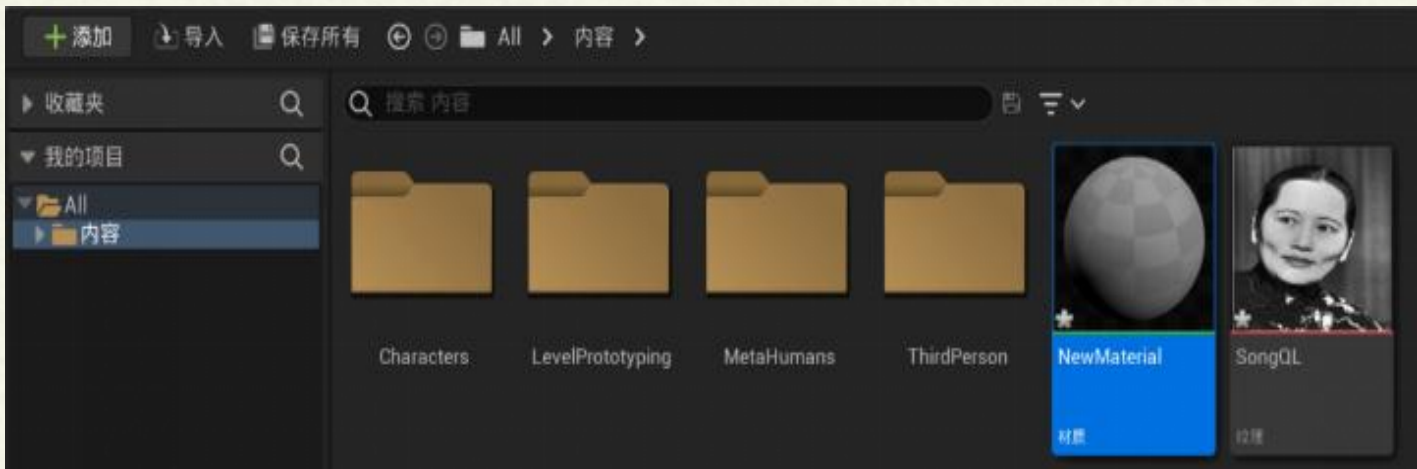
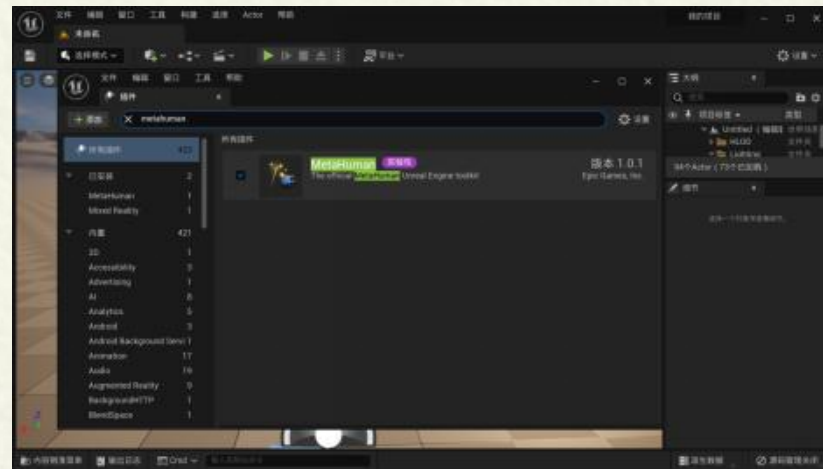
得到训练用音频集，在MockingBird环境中使用普通话语言模型开始训练

- Mockingbird训练至**88%**接近原声
- Step从75k训练到**98k**

训练完成，运行demo toolbox.py或web.py，导入训练用音频与指定文字，试听效果后导出



二 核心技术及实现流程：数字人设计流程



三 解决的问题

1、提高图像的质感与动感

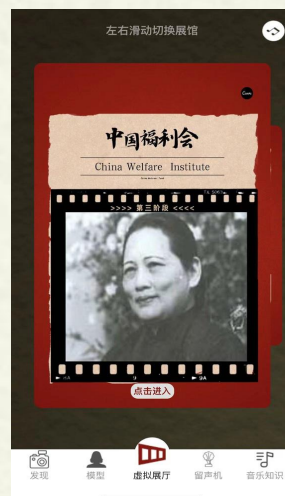
从理论到作品落地，充分验证了音乐与图像多模态交叉交融的可能性。

2、再现宋庆龄人物原声

构建宋庆龄数字人，使用Mockingbird进行声音模型训练，再次聆听伟人声音。

3、让记忆从平面映到立体

虚拟展厅与全景色、无损音效的结合，使档案资源、图像资源从平面的维度延伸到无限的立体空间。



四 备赛经验与参赛感想

量变引起质变

- 时间分配 1（构思）：1（资料收集）：3（实施与测试）6月初报名 8月底提交作品
- 图片资料收集与处理总量 >1.4GB；音效与音频总量 >1GB 赛事培训 学习各类开放数据的使用
- 集中研讨会总时长>30h 每周集中讨论进展

团队成员

中山大学



王美贤



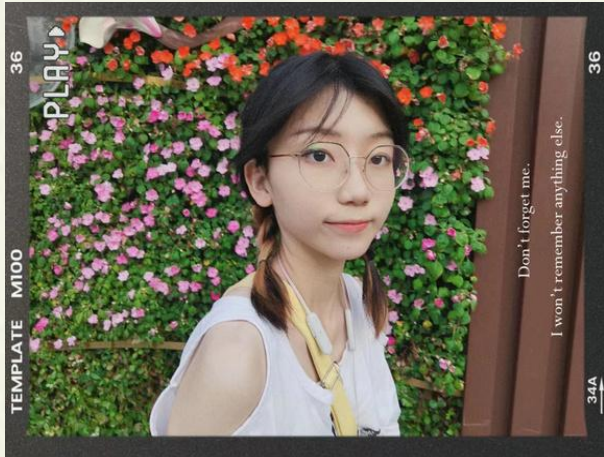
王心雨



颜欣杰



林纪元

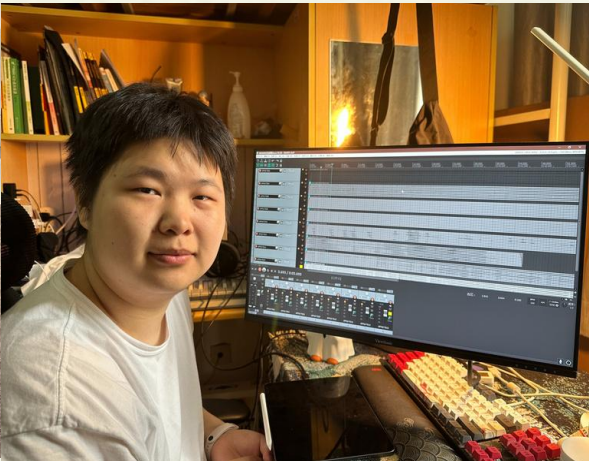


余心语

上海音乐学院



程宸



崔祺杰



黄倩雯



李赓



潘家云

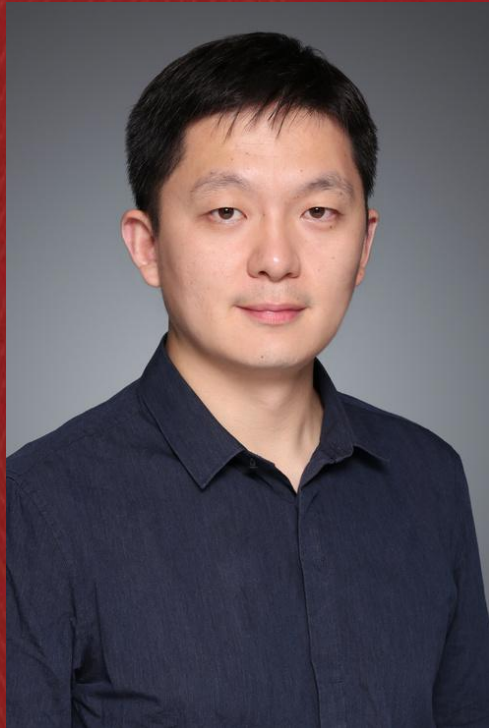


谢凤祺

指导老师



陈涛



陈世哲



张龙飞