

虚拟现实技术 2024 秋 - 小作业 2

满分: 100 (按比例转化为课程总分 15 分)

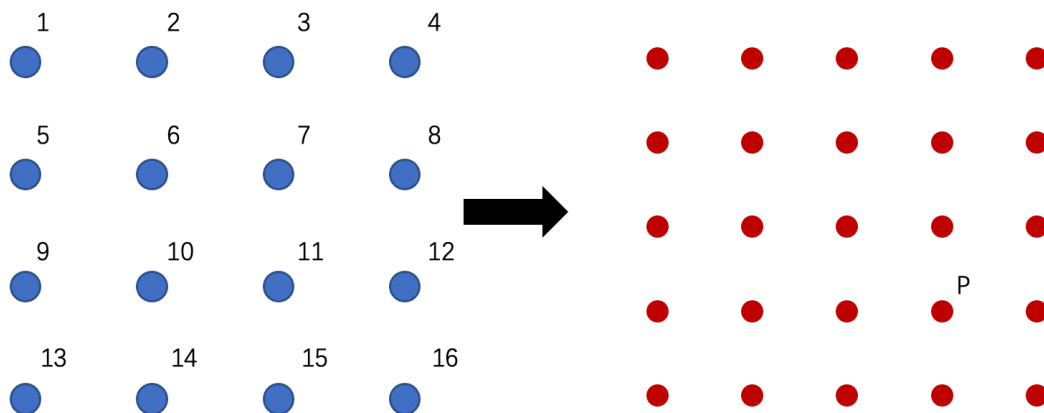
截止日期: 2024 年 10 月 27 日 23:59:59

1 书面部分 (40 分)

(1) 在 VR/AR 显示中, 视网膜级别的屏幕至少需要有多大的单眼视野和分辨率? (5 分) 目前消费级的 VR/AR 眼镜分别可以达到多大的单眼视野和分辨率? (5 分) 除了视野受限、分辨率较低外, 目前消费级的设备还存在着哪些不足? (5 分)

(2) 人可以从哪些视觉线索中感知深度? (5 分) 这些线索对提升 VR 环境的真实感有何启示? (5 分)

(3) 如下图, 一个 4×4 的图像被缩放为 5×5 , 颜色值在像素右上角标出, 使用**双线性插值**的方法计算 P 点的颜色值, 列出计算公式 (这里将像素看成点, 插值前后角点对齐) (10 分)



2 编程部分 (60 分)

本次作业中, 你将基于给定的代码框架, 使用 HRTF 实现空间立体声效果. 代码框架使用 Python 语言编写, 请依照下述流程进行环境配置, 并补全 main.py 中的 gen_360 函数.

2.1 环境配置

- 安装 Python ≥ 3.6
- 安装 Python 库依赖: `python3 -m pip install -r requirements.txt` (推荐使用虚拟环境)
- 运行主程序: `python3 main.py --demo` 运行框架给出的示例代码, 生成一个声源位置在左侧的双声道立体声音频 left-side.wav; `python3 main.py` 运行补全的代码.

2.2 具体要求

- `gen_360` 函数输入一个单声道 WAV 音频文件路径, 要求在该函数中生成一个双声道立体声 WAV 音频, 声源位置以某自定轨迹变化, 体现出空间立体声效果即可. 请在代码中以注释形式简要描述设计的轨迹及生成的思路 (或另附文件说明, 不必很长, 讲清楚即可).
- 实现中需要考虑对于不同长度音频的适配 (合理即可, 测试时不会采用极端测例).
- 实现的方式并不唯一, 只要结果体现立体效果并且符合设计的轨迹即可. 反之, 如果实现的效果较差或者效率过低, 可能会被扣分.
- 实现中可以调用 Python 标准库和 NumPy, 但 **不允许使用其他外部库**.
- 除了 `gen_360` 函数外, 你可以任意增加其他函数, 但不要改动已有的函数实现.

2.3 利用 HRTF 生成空间立体声

利用 HRTF 生成某一方位空间立体声的步骤如下:

1. 读取单声道音频数据, 通过快速傅里叶变换转换为频域信号 F ;
2. 读取某一方位左右耳的 HRIR (Head-Related Impulse Response) 数据, 通过快速傅里叶变换转换为左右耳的 HRTF (Head-Related Transfer Function), 分别记为 HL 和 HR;
3. 将 HL 和 HR 分别与 F 相乘, 通过快速傅里叶变换逆变换恢复为左右耳的时域信号, 保存为双声道音频.

2.4 代码说明

`gen_demo` 函数提供了一个使用 HRTF 生成空间立体声音频的示例. 该函数读取了左侧位置的 HRIR 数据, 通过上述生成步骤, 得到了一个双声道立体声音频 `left-side.wav`, 听上去好像声音从左侧发出来. 你需要补全 `gen_360` 函数, 实现声源位置在空间中不断变化的效果. 为实现这种效果, 你需要将音频数据切分为小的片段, 对每个片段应用不同位置的 HRTF. 为避免出现失真现象, 你可能需要将音频数据切分为稍长的、相互重叠的片段, 经过 HRTF 作用后取每个片段中间位置的一小段拼接得到最终的音频.

`main.py` 中, 提供了以下函数:

- `load_wav` 输入单声道 WAV 音频文件路径, 返回音频采样率和音频数据. 采样率为整数, 表示每秒的样本数. 音频数据为一维 NumPy 数组, 数据类型为 `int16`. 用音频数据的长度除以采样率即为音频以秒为单位的时长.
- `save_wav` 输入存储路径, 采样率和双声道音频数据, 将音频数据保存为双声道立体声 WAV 文件. 音频数据须为形状 $(N, 2)$ 的 NumPy 数组, 数据类型为 `int16`.
- `load_hrir` 输入 HRIR 数据文件路径, 返回左耳和右耳对应的 HRIR 数据. 左右耳的 HRIR 数据均为形状 $(25, 50, 200)$ 的 NumPy 数组, 数据类型为 `float32`. HRIR 数据具体的格式说明见后.

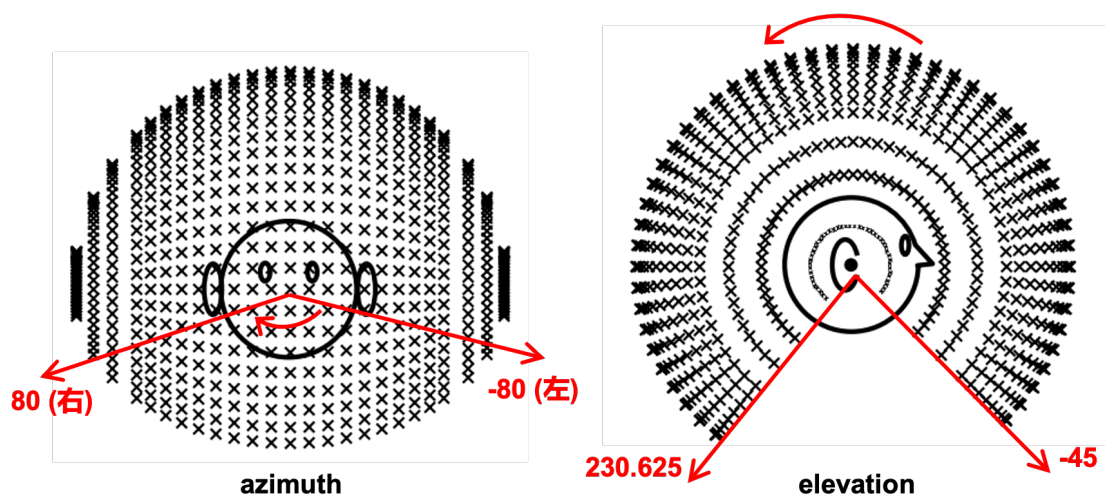
2.5 HRIR 数据格式

左右耳的 HRIR 数据均为形状 $(25, 50, 200)$ 的 NumPy 数组, 其中

- 第一维对应着 25 个 azimuth 角 (偏航角): $-80, -65, -55, -45, -40, -35, -30, -25, -20, -15, -10, -5, 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, 65, 80$
- 第二维对应着 50 个 elevation 角 (俯仰角): $-45, -45 + 5.625, -45 + 5.625 \times 2, -45 + 5.625 \times 3, \dots, -45 + 5.625 \times 49$

设左耳 HRIR 数据为 HL, 则 HL[0,0] 对应着 azimuth 角为-80 度、elevation 角为-45 度方位上测得的左耳 HRIR 数据.

azimuth 角和 elevation 角的一个简单示意如下图:



2.6 参考资料

- NumPy 基础教程 https://numpy.org/doc/stable/user/absolute_beginners.html
- 课件 05: 听觉、触力觉机制与装置

3 提交要求

将提交书面部分的作答 (电子版/手写拍照均可) 与代码打包上传至网络学堂. 样例文件结构如下所示, 顶层为以学号命名的文件夹, 将该文件夹压缩为学号.zip (如 2022011333.zip) 上传至网络学堂. 如果使用 Python 虚拟环境, 请注意**不要**将包含虚拟环境的文件夹上传上来.

```
2022011333
├── 书面部分.pdf
├── code/
│   ├── main.py
│   └── ...
```

如对本次作业有任何疑问, 请邮件联系助教或在微信群中提问. 助教联系方式:

- 徐天行 xutx21@mails.tsinghua.edu.cn
- 刘家宏 liujiaho23@mails.tsinghua.edu.cn