

实验题目：语音标注与识别

实验目标

1. 理解语音信号的基本特性
2. 掌握语音信号的基本处理与标注方法。
3. 学会语音特征提取与可视化。
4. 体验语音识别的完整流程

实验准备

1. 安装工具：
 - a. 录音与标注工具：[Audacity](#)
 - b. Python 环境：安装 Anaconda 或 Python 3.8+
 - c. Python 库：`pip install librosa matplotlib`
2. 录制语音数据集：
 - a. 用手机或电脑麦克风录制 10 条语音，内容为简单中文短句（例如：“丽红老师真好看”）。
 - b. 每条语音长度控制在 3-5 秒，背景尽量安静。
 - c. 保存为 .wav 格式，命名规则：`audio_1.wav, audio_2.wav`，采样率选 16000Hz，单声道。

实验内容

1. 语音标注与基础分析
 - a. 用 **Audacity** 标注语音
 - i. 打开 Audacity，拖入音频文件 `audio_1.wav`。
 - ii. 标注文本内容：
 1. 播放音频，在波形图上选中语音段（鼠标左键拖动）。
 2. 点击顶部菜单→轨道→添加标签轨→在选中区域输入对应文本（例如“丽红老师真好看”）。
 - iii. 标注静音段：观察波形图中最接近零的部分（开头/结尾的空白），用相同方法标注为“静音”。
 - iv. 导出标注文件：文件→导出→导出标签为文本文件→保存为 `audio_1_labels.txt`。
 - b. 基础分析

- i. 查看语音时长：Audacity 界面左下角显示总时长。
- ii. 统计静音段占比：用标签中的“静音”段时长/总时长*100%（例如：总时长 5s，静音段 1s→占比 20%）。

2. 语音特征提取与可视化

- a. 参考以下代码，提取 MFCC 特征：

```
import librosa
import librosa.display
import matplotlib.pyplot as plt

# 1. 加载语音文件
audio_path = "audio_1.wav" # 修改为你的文件路径
audio, sr = librosa.load(audio_path, sr=16000)

# 2. 绘制波形图
plt.figure(figsize=(12, 4))
librosa.display.waveshow(audio, sr=sr)
plt.title("语音波形图(音量随时间变化)")
plt.xlabel("时间(秒)")
plt.ylabel("振幅")
plt.show()

# 3. 提取 MFCC 特征
mfcc = librosa.feature.mfcc(y=audio, sr=sr,
n_mfcc=13)

# 4. 绘制 MFCC 热力图
plt.figure(figsize=(12, 4))
librosa.display.specshow(mfcc, x_axis="time",
sr=sr)
plt.colorbar(label="MFCC 系数值")
plt.title("MFCC 特征(不同颜色代表不同频率的能量)")
plt.xlabel("时间(秒)")
plt.ylabel("MFCC 系数序号")
plt.show()
```

注：

•

率。

- `waveshow()`: 显示声音的波形（横轴是时间，纵轴是音量大小）。
- `mfcc()`: 计算 MFCC 特征，将声音转换为 13 维的数字矩阵。
- `specshow()`: 用颜色深浅表示 MFCC 系数大小，横轴是时间，纵轴是不同 MFCC 维度。

b. 观察与记录

- i. 对比不同语音的波形图和 MFCC 图（例如“李红老师真好看” vs “李立耀怎么这么帅”）。
- ii. 回答思考题：MFCC 图中颜色深浅代表什么？

实验要求

1. 提交材料:

- a. 提交录音数据集（包含 10 条录音文件 `audio_1.wav` ... `audio_10.wav`）、标注数据集（对应 10 条录音文件的 10 个标注文件 `audio_1_labels.txt` ... `audio_10_labels.txt`）。
- b. 实验报告，重命名为“**学号姓名.docx**”。
- c. 截止日期：2025 年 4 月 13 日 23:55。

2. 实验报告内容:

采集的每一条语音数据都应该包括:

- a. 步骤 1 中的基础分析部分两问。
- b. 步骤 2 中的观察与记录部分两问（附上每一条数据的波形图和 MFCC 图，并回答问题）。