**音频语音处理技术 作业1：认识音频 报告**

姓名：\_\_\_\_\_\_\_\_\_\_\_ 学号：\_\_\_\_\_\_\_\_\_\_ 学院：\_\_\_\_\_\_\_\_专业方向:\_\_\_\_\_\_\_\_\_\_\_

【注意回答问题、提供代码、指令，贴波形和语图等。除本文档外，附带图原件文件和录音文件，请务必自行完成作业。】

1. 观察给定的两个音乐文件（music\_clip1&2, 一个是打击乐片段，另一个是弦乐片段）的采样率，量化位数和通道数；同时观察编码格式是什么。（提示：可以使用sox工具，指令:sox -V 目标wav -n）

答：

1. 录制一段自己的语音，内容为“音频语音处理技术真的很神奇啊”。提示：可以用电脑录音机程序或者专业录音软件audition或audicity（开源软件）录音，也可以使用手机录音机录制，采样率设置为44.1KHz，保存成wav格式。注意必须要录制自己的声音，压缩包中给出的语音样例（speech\_example）仅供参考。建议在安静的房间内，嘴巴离麦克风20-30cm，正常音量录音。

答：

1. 对上面的音频文件（语音和音乐）降采样到16K和8K，对比降采样前后的音质变化，写出主观感受。提示：用Python的librosa包进行wav读取和降采样，主要函数是librosa.load和librosa.resample，用pip安装librosa包即可。

答：

1. 对降采样到16K的语音和音乐文件转换到频域，画出语谱图，观察最高频率范围是多少，并写出语音和两段音乐在语谱上特点差异。提示：语谱图可以用python里的matplotlib.pyplot进行绘制，主要函数是plt.specgram，可以保存成pdf矢量图便于放大观察。进一步观察语谱细节可以用audition或audicity等工具。

答：

1. 对16K的语音波形和语谱图进行进一步观察，写出轻音和浊音在波形和语谱上的差别；同时测量一下几个典型元音的基频（F0），观察这些不同元音在共振峰上的差别，再观察下同一个元音在不同上下文时在语谱上的差异（语音识别建模要考虑发音上下文影响—即协同发音现象）。

答：

1. 对两个音频文件使用mp3进行压缩，观察压缩前后文件大小和音质的变化。提示：可以使用ffmpeg工具进行压缩，指令为【ffmpeg -i wav文件 -vn -ar wav文件采样率 -ac 2 -ab 码率 -f mp3 目标MP3文件】，同时可以观察不同码率对压缩质量的影响。

答：