

第一周周报

2020.10.08

报告人：何干越，其他组员：米唯实，安若鼎，张颖，杨恺

| 项目名称 | Python实现下的含噪语音分割以及数字-说话人识别 |
|--------------|---|
| 当前项目所处阶段以及目标 | <ul style="list-style-type: none"> - 准备阶段，开始学习使用Python处理语音信号，按照项目初步规划以及调研结果开展准备工作 - 目标：熟悉Python相关库的使用，语音信号处理入门 |
| 本周工作计划 | <p>各小组成员对自己所分配到的工作进行熟悉，掌握库的基本使用方法</p> <ul style="list-style-type: none"> - 何干越：使用librosa库对语音信号进行重采样，不明显失真的情况下得到数据量更小的信号段。尝试添加白噪声并使用AR模型配合Kalman滤波的方式消去噪声 - 米唯实：调研并尝试从i-vector的方式对不同说话人的信号进行处理 - 张颖：学习语音预处理相关知识，包括分帧加窗技术以及VAD静音检测（分割语音实质内容段） - 安若鼎 & 杨恺：对语音信号进行频域变换，熟悉Python & Matlab 语音信号处理库，进行频域上的处理 |
| 本周工作情况 | <ul style="list-style-type: none"> - AR 模型 Kalman滤波实现完成，可以对某一语音信号以8000Hz的频率重采样，进行分帧后使用16阶AR模型进行系数预测，互相迭代进行滤波 - 建立Github库 开始项目管理，小组成员学习Git相关知识 - 对实施方案进行了调研以及初步的设计，安排了人员工作，设计方案请见网页链接调研结果与设计方案(初步) <p>组内其他成员进行了Python或是Matlab库的熟悉</p> |
| 项目中遇到的问题 | <ul style="list-style-type: none"> - Python librosa库不支持输出LPC系数预测协方差，需要重新实现 - 小组内只有部分成员会使用Python，需要他们先行使用Matlab实现，再将思路由会Python的组员进行移植，这将会存在可能的不兼容或者库功能不同的问题，虽然个人认为Python库非常齐全，完全可以对Matlab相关函数进行替换，但是潜在的问题不容忽视 - 小组内还没有在实验思路的基础上提出更好的实现方案，这一点有待进一步讨论 |
| 下周工作计划 | <ul style="list-style-type: none"> - 小组所有成员对相关库函数已经比较熟悉，并且已经动手实践过语音信号处理 - 完成时域法加噪滤波降噪，进行实质内容分割，使用传统分类器进行分类 |
| 小组成员安排（初步） | <ul style="list-style-type: none"> - 何干越：时域法概率滤波器建模，时域法分类 - 米唯实：说话人识别以及分类器选择 - 张颖：VAD静音识别，语音分割以及DTW时域法分类 - 安若鼎 & 杨恺：实现完整的频域法处理 |

附加说明

- 其他可能加入的想法均在**调研结果与设计方案(初步)**中提到
- 由于我们认为Python灵活性较强，易于基于此语言进行实现，故小组内所有使用Matlab的代码最终都会被移植（重写）到Python框架下
- 本次周报没有完全使用模板，出于需要增加项目以及方便书写的考虑