Report

实习工作总结

**Zhaohui Yin**

**Sat, Dec 24, 2020**

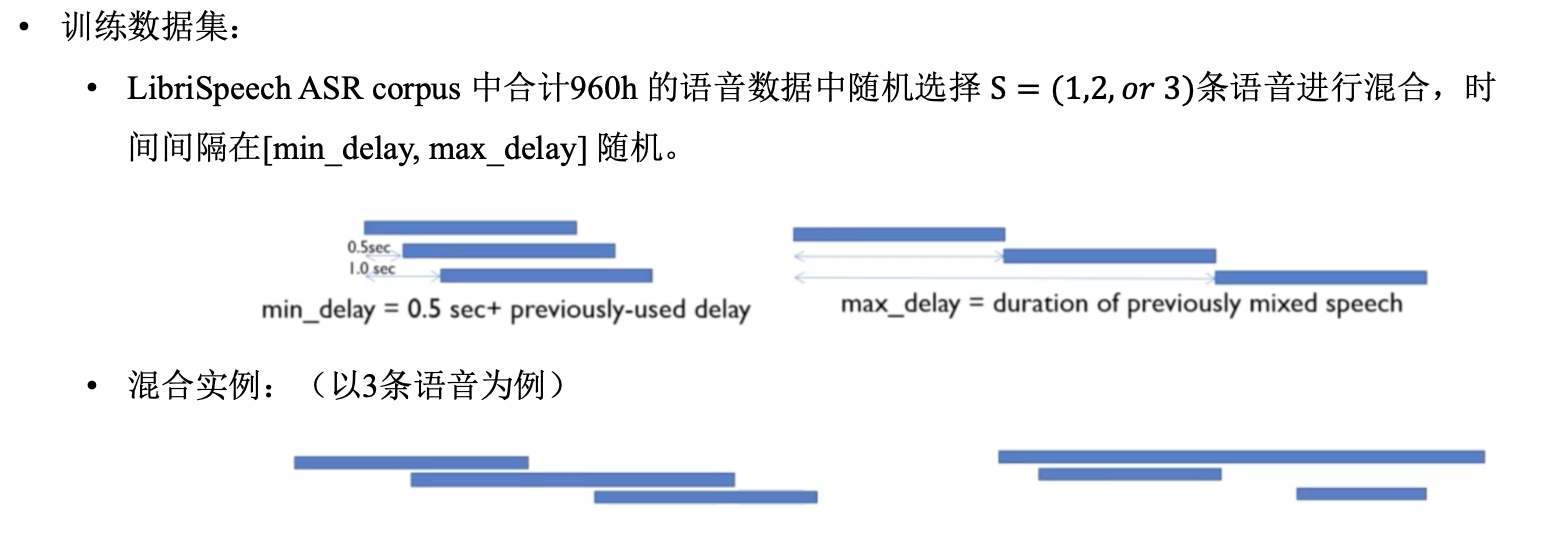
1. 计划安排表

1.1 十一月预定计划

|  |  |  |
| --- | --- | --- |
| 周 | 日期 | 工作内容 |
| WEEK 1  1.阅读新人资料  2. 初步搭建编程环境，  3. 初步确定工作大方向 | Tue Nov 3 | * + 报到，进行登记和账号激活等步骤   + 新人资料一至三   + 开会：杨磊-语音识别语言模型 |
| Wed Nov 4 | * + 上午：办公室人员及方向的初步认识和整理   + 下午：声纹识别研究综述的学习     1. 初步实验：认识语谱图 |
| Thu Nov 5 | * + 上午 & 下午：kalid的基础入门操作 (调通yesno识别小实验)   + 下午：胡老师、徐老师小组讨论，确定大方向：与重叠音相关的研究 |
| WEEK 2  1. 初步了解技术手段：重叠音检测（When），分离（Who），识别（What）  2. 传统模块化系统和端到端系统建模思路初步了解 | Tue Nov 10 | * + 李宏毅课程初步学习（为了弄清楚语音研究的各个大方向，以及重叠音研究是属于哪一个分支）   + 该课程后续需要持续学习 |
| Wed Nov 11 | * + CHiME6竞赛回顾， Review：初步了解各个技术在实际应用场景的功能：   + 基于CHiME\_2020\_paper\_du了解：   + 前端（数据增强，包括降噪和数据扩充）→ 语音分离 → 语音识别 |
| Thu Nov 12 | INTERSPEECH 2020 文献精读：着重了解端到端重叠音识别：  Serialized Output Training for End-to-End Overlapped Speech Recognition   * + 文献精读：Introduction & Related   + 关注问题：研究痛点，研究进展，建模一般流程。 |
| WEEK 3  1.进一步深入，选择端到端重叠音识别相关文献。深入了解其模型架构和技术细节。  2. 找到相关源码和实验数据 | Tue Nov 17 | * + 文献精读 ：AED-based single-speaker ASR   + 端到端的过程理解 （sequence to sequence）   + 注意力模型及实例操作（Tom chase Jerry）[理论理解，知乎] |
| Wed Nov 18 | * + Chorowski (2014)，注意力模型的理解（Attention )   + 初始值的分配问题（实际操作是怎样的？）（未解决，需要复现论文） |
| Thu Nov 19 | * + Gelabert2018，AED-based ASR模型详情。   + 文献模型复现（数据源：TIMIT，代码复现：Nabu【github】）(亟待解决) |
| WEEK 4  1. 技术分享，整理之前的工作内容，周六汇报。 | Tue Nov 24 | * + 整理已读文献内容，并制作PPT，周六进行文献分享 |
| Thu Nov 26 |
| Sat Nov 28 | 小组分享 |

1.2 十二月预定计划

技术分享会着重分享端到端重叠音识别相关文献，会上胡老师提出以下两点建议：1）深入了解重叠音分离的方法原理；2）着手制作人工合成的重叠音实验数据，如下所示：



|  |  |  |
| --- | --- | --- |
| 周 | 日期 | 工作内容 |
| Dec-WEEK 1  1. 深入了解语音分离技术及常用模型架构：  **Deep Clustering** && **PIT** | Tue Dec 1 | * + CHiME 5 文献阅读，kanda （2019）提出了GSS方法，值得一读，可以稳定降低1.5%的词误率   + 李宏毅语音分离部分详细学习，[Section 12 && 13]   + Deep Clustering && PIT 模型架构，以及TasNet模型技术细节 |
| Wed Dec 2 |
| Dec-WEEK2  1.着手制作人工合成的重叠音实验数据**Libri2Mix**  (即2个语音随机重叠，不添加噪声的合成语音数据) | Tue Dec 8 | * + 服务器配置环境变量，/cpu2/home/yinzhaohui/.bashrc   + 外网电脑pycharm安装和配置   + 人工合成重叠音的源码学习：   **1）相关文献：Cosentino(2005),LibriMIx:An Open-Source Dataset for Generalizable Speech Separation.**  **2）Github：https://github.com/JorisCos/LibriMix** |
| Wed Dec 9 (0.5 day) | * + 源码学习，调试程序先利用LibriSpeech的**test\_clean数据**作为程序测试样本。   + 进度：调试程序 ing. |
| Thu Dec 10 |

|  |  |  |
| --- | --- | --- |
| 周 | 日期 | 工作内容 |
| Dec-WEEK 3  调试程序 | Wed Dec 16 | 成果1: 利用LibriSpeech-test\_clean（346M）对程序进行测试，成功获取2-speaker混合语音数据及过程变量元文件，语音重叠发生时间为0。 |
| Thu Dec 17 |
| Dec-WEEK4  调试程序 | Mon Dec 21  (0.5 day) | 成果2: 利用LibriSpeech-test\_clean（346M）对程序进行测试，成功获取语音重叠开始时间随机的2-speaker混合语音数据及过程变量元文件。 |
| Tue Dec 22 |
| Wed Dec 23 | 成果3:获取语音重叠开始时间随机的多人说话数据及其过程变量元文件。 |