

SRT 中期汇报

计 32 黄世宇

计 32 余翔

计 32 陆禹

【前期工作】

a. 特征提取

完成对原始语音文件的处理，使用 matlab 对原始 wav 文件进行处理，提取 13 维 mfcc 特征，取一阶差分和二阶差分得到一共 39 维特征，根据第 1 维能量去掉静音部分。然后做一次 feature warping，最终得到 csv 特征文件。

b. 训练数据处理

Caffe 支持多种文件格式输入，我们选择了 HDF5 格式。利用 python 对提取出来的 csv 文件进行到 HDF5 的转换，并形成 label。

c. 训练

设计 CNN 网络，输入 HDF5，利用 Caffe 进行训练。

d. 结果提取

利用 Caffe，通过使用训练出来的模型对未知的输入语音进行分类。

【问题】

1. 目前训练得到的结果不好。
2. 还没建立起评估标准测试，例如 DET (Detection Error Trade-off) 和 EER (Equal Error Ratio) 以及 FA 和 FR。