

文本相关声纹识别

MFCC-GMM

★ Ref: 基于智能手机的语音认证技术的研究与实现

MFCC-DTW/VQ

★ Ref: 基于梅尔频率倒谱系数与动态时间规整的安卓声纹解锁系统

Code

GitHub, pierre-rouanet/dtw, MFCC+DTW

librosa

- ✓ librosa.load ○ 语音向量
- ✓ librosa.feature.mfcc ○ MFCC特征向量
- librosa.decompose.nn_filter ○ 降噪处理
- librosa.effects.trim ○ 去除首尾静音(较难有纯静音)

Python库

- ✓ dtw.dtw ○ 动态时间规划
- ✓ py-webrtcvad > pyvad, vad ○ 端点检测
- ✓ logmmse ○ 降噪处理
- ✓ scipy.io.wavfile.write ○ numpy向量保存为语音

算法步骤

- 降噪处理、端点检测
- 计算MFCC特征向量
- DTW计算特征向量最短距离
- 阈值判断

Look

- 采样设置: 采样率、声道数 (librosa进行重采样) (不同设备的影响)
- 样码语料库综合
- 端点检测(Voice Activity Detection (VAD))
- 降噪算法处理(Noise Reduction)

效果有待验证

GMM-UBM

★ Ref: 基于GMMUBM的快速说话人识别方法、声纹识别初探-以 GMM-UBM 为例

Code

商用; GMM的扩展

算法步骤

- MAP自适应算法相当于先进性一轮EM迭代得到新的参数,然后将新参数和旧参数整合
- 先使用大量的非目标用户数据训练UBM,然后使用MAP自适应算法和目标说话人数据来更新局部参数得到对应的GMM.

End2End

★ Ref: End-to-End Text-Dependent Speaker Verification