

# AI Demo



# 倾听你的声音

汤志远

# 文件夹结构

- ❑ AIDemo 虚拟机桌面上点击右键，选择“打开终端”。
- ❑ 用linux命令进入speechSeparation，并查看文件夹内容。

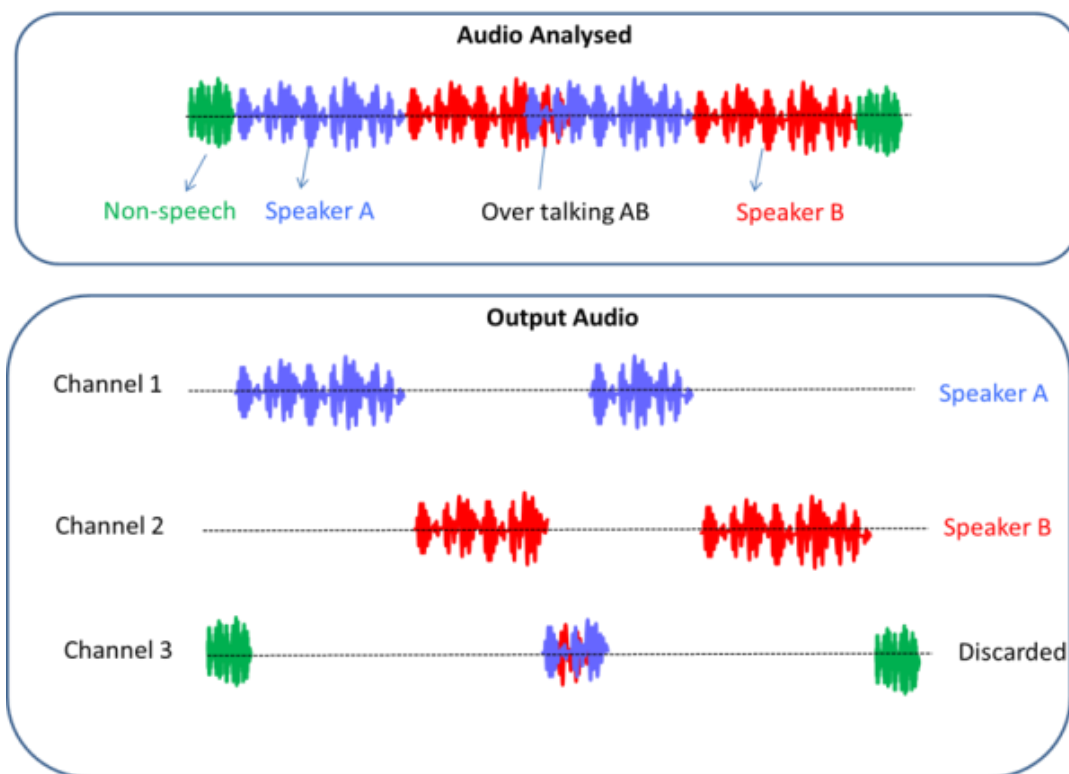
- doc: 样例程序文档说明
- code: 样例程序代码
- code/run.sh: 样例程序入口

```
[tutorial@localhost ~]$ ls
aibook  下载  公共  图片  文档  桌面  模板  视频  音乐
[tutorial@localhost ~]$ cd aibook/
[tutorial@localhost aibook]$ ls
demo
[tutorial@localhost aibook]$ cd demo/
[tutorial@localhost demo]$ ls
data  env  image  lang  mind  robot  speech
[tutorial@localhost demo]$ cd speech/
[tutorial@localhost speech]$ ls
speechSeparation
[tutorial@localhost speech]$ cd speechSeparation/
[tutorial@localhost speechSeparation]$ ls
code  data  doc
[tutorial@localhost speechSeparation]$
[tutorial@localhost speechSeparation]$
[tutorial@localhost speechSeparation]$
```

绝大多数情况，运行如下命令运行样例程序：sh run.sh

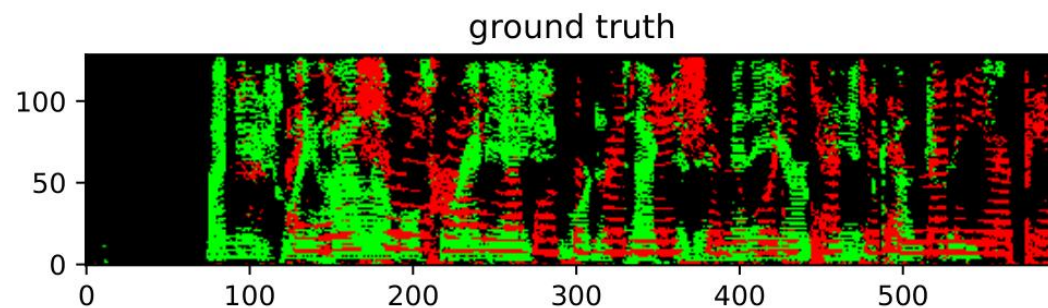
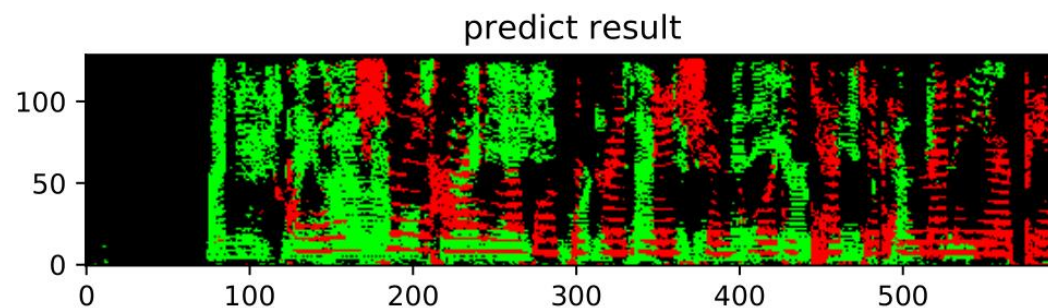
# 语音分离简介

语音分离可以简单理解为，输入一个多音源混杂的语音，将每个音源的声音分离出来。



# 实验一：运行缺省配置

打开终端窗口，进入speech/speechSeparation/code 目录，运行run.sh 程序，即可观察到基于缺省参数的DeepCluster 语音分离效果。



## 实验二：体验不同场景下的分离效果

前一个实验中的两个声源文件，01ua010a.wav 和20ra010a.wav，是两个性别不同的发音人发出的内容不同的语音，因此音源之间的差异性比较大，DeepCluster 比较容易实现对音源的分离。本实验将探讨在相对复杂的条件下，DeepCluster 方法的分离性能。下面的实验均需通过修改run.sh 完成。

# 实验二：体验不同场景下的分离效果

## (1) 观察说话内容对人声分离效果的影响

依两路音源的发音内容不同，分离的难度会有所不同。通过修改run.sh，比较DeepCluster对内容相同的发音和内容不同的发音的分离效果。相同内容：

```
python3 main.py model/speech_model.h5  
../data/speech_separate/01u/01ua010a.wav  
../data/speech_separate/20r/20ra010a.wav
```

不同内容：

```
python3 main.py model/speech_model.h5  
../data/speech_separate/01u/01ua010a.wav  
../data/speech_separate/20r/20ra010b.wav
```

# 实验二：体验不同场景下的分离效果

## (2) 观察发音人性别对人声分离效果的影响

依两路音源的发音人性别不同，分离的难度会有所不同。通过修改run.sh，比较DeepCluster对同性别混合发音和不同性别混合发音的分离效果。

相同性别：

```
python3 main.py model/speech_model.h5 ../data/speech_separate/01u/01ua010a.wav  
../data/speech_separate/027/027a010b.wav
```

不同性别：

```
python3 main.py model/speech_model.h5 ../data/speech_separate/01u/01ua010a.wav  
../data/speech_separate/20r/20ra010b.wav
```

# 实验二：体验不同场景下的分离效果

## (3) 观察音调匹配性对人声-背景音乐分离效果的影响

如果两路音源属于不同类的声音，则分离会相对容易，如音乐和人声。然而，分离效果可能受到音调匹配度的影响。下面的实验考察人声和音乐在音调匹配和不匹配两种情况下的分离效果。

音调匹配的人声和音乐分离：

```
python3 main.py model/music_model.h5  
../data/music_separate/0_accompaniment.wav  
../data/music_separate/0_vocals.wav
```

音调不匹配的人声和音乐分离：

```
python3 main.py model/music_model.h5  
../data/music_separate/0_accompaniment.wav  
../data/music_separate/1_vocals.wav
```



## 实验三：分离自己的声音

录制你和朋友的混合声音，上传到AIDemo虚拟机，对该混合声音进行音源分离。通过实验，可以发现你和哪位朋友的声音更相像（相像意味着更不容易被分离）。

The end