BN特征提取

安装使用文档



# 操作系统

Linux 内核系统，推荐ubuntu

# 环境安装

1. 安装[kaldi](http://kaldi-asr.org/doc/install.html)，然后进行环境创建配置

# 下载kaldi

git clone https://github.com/kaldi-asr/kaldi.git kaldi --origin upstream

cd kaldi

# 编译kaldi

cd tools #进入kaldi目录下的tools文件夹

./extras/check\_dependencies.sh #检查系统依赖包要求是否满足，按照提示进行安装系统依赖包

# 安装kaldi

cd src #进入kaldi目录下的src文件夹

# 以下脚本同样存在于INSTALL文件中

./configure --shared

make depend -j 8

make -j 8

# 数据准备

批量化特征提取按照一定格式准备好需要转换为bottleneck特征的音频文件。这一节会简单介绍几个数据准备文件的生成过程。

# 解压提供的bn\_extractor.zip

unzip bn\_extractor.zip

cd bn\_extractor

# 在data目录下创建自己的数据目录，其中里面的sample为示例目录，比如目录名为synthesis, 创建完目录后，在目录下准备3个文件，分别是wav.scp，text和utt2spk。

mkdir data/synthesis

cd data/synthesis

### wav.scp 准备

wav.scp指定每条语音的地址，提取工具会根据文件里的音频地址提取特征，语音文件需要转为16k，16 bit，单通道的wav文件。文件中的格式为 <音频名>（空格）<音频地址>。音频名字会作为键值后面索引bottleneck特征使用。可参考data/sample/wav.scp文件：

### utt2spk 准备

utt2spk里描述每一条语音对应的说话人，如果同批数据里有相同的说话人，那么提取器会利用同一个说话人的所有音频来提取更准确bottleneck特征。文件的格式为<音频名>（空格）<说话人>。音频名得与wav.scp里的一样。可参考data/sample/utt2spk文件：

p225\_001 /NASdata/zhangchx/datasets/VCTK/wavs/p225/p225\_001.wav

p225\_002 /NASdata/zhangchx/datasets/VCTK/wavs/p225/p225\_002.wav

p225\_003 /NASdata/zhangchx/datasets/VCTK/wavs/p225/p225\_003.wav

p225\_004 /NASdata/zhangchx/datasets/VCTK/wavs/p225/p225\_004.wav

### text准备

文件内容和utt2spk相同即可

cp utt2spk text

p225\_001 p225

p225\_002 p225

p225\_003 p225

p225\_004 p225

# 特征提取

在bn\_extractor目录下更改path.sh中的环境变量路径，将其中的KALDI\_ROOT路径指向机器中kaldi的安装目录，假设kaldi目录为/home/kaldi，那么path.sh文件中kaldi根目录地址为：

export KALDI\_ROOT="/home/kaldi"

最后用脚本./extract\_btlnfea.sh提取特征即可。下面以sample目录为例：

./extract\_btlnfea.sh data/sample data/sample/ivector data/sample/bnf

# 脚本需要3个参数，分别是数据准备目录，说话人特征保存目录，bottleneck特征保存目录，这里我们分别设置为data/sample，data/sample/ivector和data/sample/bnf。最后提取完的特征会保存在data/sample/bnf里，具体特征地址索引会保存在data/sample/bnf/feats.scp文件里。

# 脚本还有3个参数分配机器的计算资源，越大提取速度越快。其中nj是提取bottleneck特征的并行实例数。取决于机器的gpu数量和数据的说话人个数，取两者最小值。比如可用gpu数量为8，数据说话人数量为3，那么nj最大可以设置为3. 另外一个是mfcc\_nj，其最大值为cpu数量和音频数量的最小值。而ivector\_nj的最大值为cpu数量和说话人数量的最小值。