聚类引擎开发测试环境

# 硬件要求

Intel(R) Xeon(R) CPU，内存64G。

# 软件要求

操作系统建议安装CentOS 7.x，gcc/g++ 4.8及以上。

# 测试运行

测试程序位于该路径下：

speakerCluster/Env/linux

|--cluster\_demo 测试程序

|--libClustering.so 聚类库

|--configure.cfg 配置文件

|--Decode.cfg 模型前向计算配置

|--feat.cfg 特征提取配置

|--spk\_role\_role1\_role2\_2390\_63\_60.bin d-vector特征提取模型文件

|--Relyon\_so 依赖的库文件

测试方法：

数据：语音及对应的识别结果\_sent.txt文件

./run.sh 1 wav.list configure.cfg

参数1：线程数

参数2：语音列表

参数3：配置文件

# 配置参数说明

## 4.1 配置参数说明

### 4.1.1 hybrid configure.cfg

<ClusterCfg>

#聚类目标人数

iDstClusterNum=2

iSingleMinTimeSec=0.1

iSingleOutTimeSec=10800

#聚类所使用的距离公式

ClusterDisType=GLR2

TimeSegMode=ORIGINAL

#KMeans聚类的次数

iKMeansTimes=10

#以下两个参数为特征提取的帧长和帧移

iFeatFramePeriodMSec=40

iFeatFrameStepPeriodMSec=20

#根据分词时间进行二次分段，分词间隔大于0.01s时进行二次切分

fCombinedTimeMSec=0.01

save\_blank=false

BreakPointDetect=false

WriteBreakPointLog=false

logTxt=log.txt

bpRuleCfg=bp\_rule.cfg

#是否开启高通滤波

HighPassPro=true

FC=26

dBGain=65

fClusBicAlpha=0.5

#短时过滤阈值，单位ms

shortTimeThreshold=600

#基频阈值

pitchThreshold=110

#能量阈值

energyThreshold=25

#该值为true时，上面两个参数才起作用，低于这些阈值的单独做KMeans

shortTimeFilterByKmeans=false

#特征是否使用基频特征

pitchFeatureEnable=true

#特征是否使用基频一阶差分特征

pitchFeature1OrderEnable=true

#特征是否使用能量特征

energyFeatureEnable=true

#电信系统前导音不参与聚类，add by xiaosujie

LeadingVoiceFilter=false

#前导音一般出现在开始处，如果开始的前几行不包含则不再继续判断，该参数可跟据sent文件配置

sentLineNum=5

LeadingVoiceSentence1=为-您-服务

LeadingVoiceSentence2=话-务-员,客服代表,客户-代表,号,为-您-服务

#ClusteringBuf模式时， shell 为离线客服shell接口，结果输出到语音目录下， distribute 苏宁分布式接口，输出到指定目录下

Mode=shell

<License>

iLicense=license.dat

### 4.1.2 e2e configure.cfg

<ClusterCfg>

#聚类目标人数

iDstClusterNum=2

iSingleMinTimeSec=0.1

iSingleOutTimeSec=10800

#聚类所使用的距离公式

#COSINE/EUC/GLR2

ClusterDisType=COSINE

TimeSegMode=ORIGINAL

#KMeans聚类的次数

iKMeansTimes=10

#以下两个参数e2e版本不起作用

iFeatFramePeriodMSec=40

iFeatFrameStepPeriodMSec=20

#根据分词时间进行二次分段，分词间隔大于0.5s时进行二次切分

#fCombinedTimeMSec=0.01

fCombinedTimeMSec=0.5

save\_blank=false

BreakPointDetect=false

WriteBreakPointLog=false

logTxt=log.txt

bpRuleCfg=bp\_rule.cfg

HighPassPro=false

FC=26

dBGain=65

fClusBicAlpha=0.5

#短时过滤阈值，单位ms

#short time filter ms

shortTimeThreshold=2500

#shortTimeThreshold=0

pitchThreshold=110

energyThreshold=30

shortTimeFilterByKmeans=false

pitchFeatureEnable=false

pitchFeature1OrderEnable=false

energyFeatureEnable=false

#是否做谱聚类以及谱聚类特征维度，可根据实际测试集进行调整

isDoSpectral=false

spectralClusteringDim=8

#e2e跳变点检测

#跳变点检测的窗长，单位帧，即窗长1s

winFrame=100

#跳变点检测的窗移

stepFrame=10

#skipFrame参数要同前向计算的跳帧参数同，见Decode.cfg

skipFrame=2

#相似度小于该值时，则为跳变点

consineThreshold=0.95

#ClusteringBuf模式时， shell 为离线客服shell接口，结果输出到语音目录下， distribute 苏宁分布式接口，输出到指定目录下

Mode=shell

#电信系统前导音不参与聚类，add by xiaosujie

LeadingVoiceFilter=false

#前导音一般出现在开始处，如果开始的前几行不包含则不再继续判断，该参数可跟据sent文件配置

sentLineNum=5

LeadingVoiceSentence1=为-您-服务

LeadingVoiceSentence2=话-务-员,客服代表,客户-代表,号,为-您-服务

<License>

iLicense=license.dat