语法

模型

字典

语料库

标注

提取参数

测试语料库

提取参数

识别结果

HMM三要素

1. 建立speech文件夹，存储原始语音数据。
2. 建立task文件夹，在/task/下建语法文件wordnet.task ：

内容：

$digit=YI|ER|SAN|SI|WU|LIU|QI|BA|JIU|LING|YAO;

({SENT-START} {$digit} {SENT\_END})

用括号{}括住 START\_SIL和 END\_SIL表示其可不存在或者重复多次（允许在单

词之前或之后长时间的停顿，或者根本没有停顿）。

括号[]括住表示零个或一次出现（如果没有单词，可能只是识别停顿）

1. 建立hmm文件夹，建立proto的马尔科夫原始模型：

内容：

~o <HMMSetId> 0\_9\_mono

<VecSize> 39 <MFCC\_0\_D\_A\_Z>

~h "proto"

<BeginHMM>

<NumStates> 5

<State> 2

<Mean> 39

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

<Variance> 39

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

<State> 3

<Mean> 39

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

<Variance> 39

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

<State> 4

<Mean> 39

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

<Variance> 39

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

<TransP> 5

0.0 1.0 0.0 0.0 0.0

0.0 0.6 0.4 0.0 0.0

0.0 0.0 0.6 0.4 0.0

0.0 0.0 0.0 0.7 0.3

0.0 0.0 0.0 0.0 0.0

<EndHMM>

1. 建立文件dictionary，在dictionary下建立pronuciate.dic的发音字典文件：

内容：BA b a sp

ER er sp

JIU j iu sp

LING l ing sp

LIU l iu sp

QI q i sp

SAN s an sp

SENT-END [] sil

SENT-START [] sil

SI s i sp

WU w u sp

YAO y ao sp

YI y i sp

1. 建立mlf文件夹，在mlf文件夹下建立train\_word\_level0.mlf的训练词的主标注文件：

内容：

#!MLF!#

"\*/CHcctI2.17.lab"

YI

.

"\*/CHcctI2.16.lab"

ER

.

"\*/CHcctI2.11.lab"

SAN

.

"\*/CHcctI2.12.lab"

SI

.

"\*/CHcctI2.13.lab"

WU

.

"\*/CHcctI2.1.lab"

LIU

.

Etc

1. 生成底层wordnet.slf任务网络文件。

HParse -A -D -V -T 1 task/wordnet.gram task/wordnet.slf

1. 建立config文件夹，在config文件夹下建立hcopy.conf的配置文件

内容：

# Coding parameters

SOURCEKIND = WAVEFORM

SOURCEFORMAT = WAV

#SOURCERATE=625.0

SOURCERATE=1250.0 #单位为100纳秒 采样率8k

TARGETKIND = MFCC\_0\_D\_A\_Z #二阶差分并作均值规整的MFCC系数

TARGETRATE = 100000.0 # 帧移80 # 10ms frame

SAVECOMPRESSED = T #输出压缩

SAVEWITHCRC = T #输出进行校验和检查SAVEWITHCRC is set true then a checksum is #appended to the output

WINDOWSIZE = 200000.0 # 帧长160

USEHAMMING = T #汉明窗

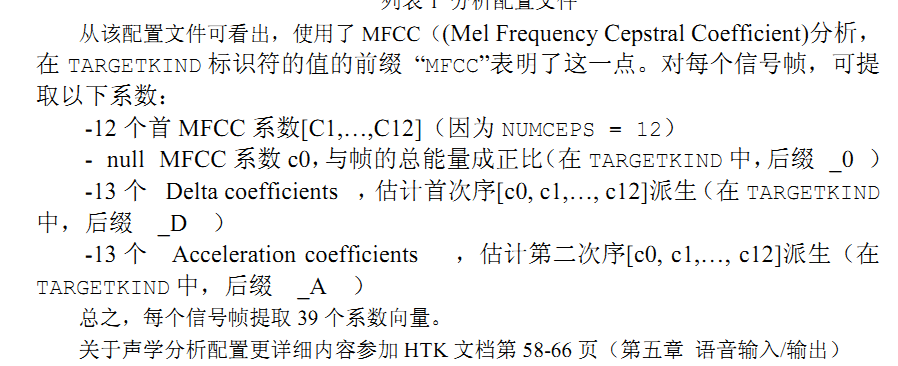
PREEMCOEF = 0.97 #预加重

NUMCHANS = 26 #滤波器通道的数目 Number of filterbank channels

CEPLIFTER = 22 #Length of cepstral liftering

NUMCEPS = 12 # compute c1 to c12

ENORMALISE = F



1. 按配置问价进行训练语音特征提取，建立scp的文件夹，在scp的文件夹下建立train\_speechtofeature.scp的文件，内容如下，建立feature文件夹执行HCopy -A -D -V -T 1 -C config/hcopy.cong -S scp/train\_speechtofeature.scp提取特征

./speech/CHcctI2.17.wav ./feature/CHcctI2.17.mfc

./speech/CHcctI2.16.wav ./feature/CHcctI2.16.mfc

./speech/CHcctI2.11.wav ./feature/CHcctI2.11.mfc

./speech/CHcctI2.12.wav ./feature/CHcctI2.12.mfc

./speech/CHcctI2.13.wav ./feature/CHcctI2.13.mfc

./speech/CHcctI2.1.wav ./feature/CHcctI2.1.mfc

Etc

1. 初始化模型，计算所有训练语音参数的全局均值和方差，用来初始化模型，在scp文件夹下建立train\_feature.scp的文件，内容如下，在hmm模型下建立hmm0文件夹，执行hcompv -A -D -V -T 1 -f 0.01 -m -S scp/train\_feature.scp -M hmm/hmm0 hmm/proto（与原型不同）

内容：./feature/CHcctI2.17.mfc

./feature/CHcctI2.16.mfc

./feature/CHcctI2.11.mfc

./feature/CHcctI2.12.mfc

./feature/CHcctI2.13.mfc

./feature/CHcctI2.1.mfc

./feature/CHcctI2.14.mfc

./feature/CHcctI2.10.mfc

./feature/CHcctI2.23.mfc

./feature/CHcctI2.19.mfc

./feature/CHcctI2.3.mfc

1. 构造hmm/hmm0目录下的hmmdefs模型定义文件和macros全局宏定义文件，创建macros文件 。
2. 无sp模型下训练三遍。
3. 修正静音模型，并添加sil和sp的状态邦定。（参考htkbook 35页 有状态图）
4. 继续训练两遍。
5. 离线识别。