# 语音实验2 使用voicebox进行说话人识别

## 实验准备

* 自行录制语音文件
* Matlab voicebox 工具箱，可以从QQ或者  
  <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html> 下载
* gmm\_estimate.m 包含用于训练说话人模型的函数gmm\_estimate ，可从QQ下载
* lmultigauss.m 包含用于计算测试语音似然率的函数lmultigauss，可从QQ下载

# 具体实施细则

每次朗读内容分为两个部分，切分后加起来一共6句：

1、登鹳雀楼 王之涣

2、白日依山尽

3、黄河入海流

4、欲穷千里目

5、更上一层楼

6、八千里路云和月

## 实验目的

本次实验初步介绍如何用GMM（Gaussian Mixture Model）的方法来进行说话人确认（Speaker Recognition）。

训练说话人模型时，使用第一周正常语速的6句

测试的时候，使用第一周其余的语句。

## 实验内容

### 配置Matlab工具箱

将实验准备中列出的工具箱解压放在一个文件夹下，使用 Matlab 的 File – Set Path 功能中的 Add with subfolders 按钮，将实验所需工具箱及相关添加到 Matlab 的 Path 中并使用 Save 按钮保存配置。

或者，更简单地，将依赖的文件放在和主程序相同的文件夹中即可。

### 参考代码

参考代码通过wavread以及melcepst读取.wav文件并提取特征train\_feature（12维MFCC），然后使用gmm\_estimate为说话人训练模型（16阶GMM），得到模型的3个参数[mu, sigma, c]。最后将被测特征test\_feature和要比对的说话人模型参数传入函数lmultigauss，即得到该被测特征与指定模型的比对得分lY。

代码中的 [YM,Y] = lmultigauss(x,mus,sigm,c) 试图计算模型匹配得分，得到YM, Y 两个参数。对Y求mean，即该段测试语音得分与对应模型的比对得分。

代码训练了1个说话人的模型，测试了13 组特征，最终score为得分数组，得分数组的第i行第j列的值代表第i个测试语音与第j个说话人模型的得分，分数高则表示更匹配。

使用本代码前请留意代码中的文件路径。

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%本代码适用于单人的说话人确认

clc;

clear all;

close all;

MFCC\_size=12;%mfcc的维数

GMMM\_component=16;%GMM component 个数

mu\_model=zeros(MFCC\_size,GMMM\_component);%高斯模型 分量 均值

sigma\_model=zeros(MFCC\_size,GMMM\_component);%高斯模型 分量 方差

weight\_model=zeros(GMMM\_component);%高斯模型 分量 权重

train\_file\_path='.\training\1\1\_';%模型训练文件路径

num\_train=6;%目标说话人模型训练文件的个数

test\_file\_path='.\testing\';%测试文件路径

num\_test=1;%测试情况下朗读的次数

num\_uttenance=6;%测试情况下每次朗读的句子的总数

all\_train\_feature=[];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%train model

%使用1\_1～1\_6训练

for i=1:num\_train

train\_file=[train\_file\_path num2str(i) '.wav'];

[wav\_data ,fs]=wavread(train\_file);

train\_feature=melcepst(wav\_data ,fs);

all\_train\_feature=[all\_train\_feature;train\_feature];

end

[mu\_model,sigma\_model,weight\_model]=gmm\_estimate(all\_train\_feature',GMMM\_component);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%test

for i=1:num\_test

for j=1:num\_uttenance

test\_file=[test\_file\_path num2str(i) '\_' num2str(j) '.wav'];

[wav\_data ,fs]=wavread(test\_file);

test\_feature=melcepst(wav\_data ,fs);

[lYM, lY] = lmultigauss(test\_feature', mu\_model, sigma\_model, weight\_model);

score(i) = mean(lY);

fprintf('Test：%d\_%d score:%f\n',i,j,score(i));

end

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%result

[max\_score,max\_id]=max(score);

[min\_score,min\_id]=min(score);

fprintf('Max score:%f\nMin score:%f\n',max\_score,min\_score);