# 实验四 MFCC特征提取及分析

## 实验目的：

### 1.1学习与掌握MFCC特征提取原理与方法

### 1.2提取一帧浊音信号的MFCC特征（13维的倒谱参数）

### 1.3对提取的MFCC特征做PCA聚类分析

## 实验设备

应用软件Matlab2018a，浊音音频一段。

## 实验原理

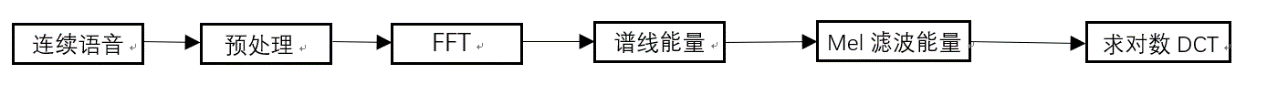
### 3.1梅尔倒谱系数（MFCC）

根据人耳听觉机理的研究发现，人耳对不同频率的声波有不同的听觉敏感度。两个响度不等的声音作用于人耳时，则响度较高的频率成分的存在会影响到对响度较低的频率成分的感受，使其变得不易察觉，这种现象称为掩蔽效应。一般来说，低音容易掩蔽高音，而高音掩蔽低音较困难。在低频处的声音掩蔽的临界带宽较高频要小。所以，从低频到高频这一段频带内按临界带宽的大小由密到疏通过一组带通滤波器，对输入信号进行滤波。将每个带通滤波器输出的信号能量作为信号的基本特征，对此特征经过进一步处理后就可以作为语音的输入特征。

梅尔倒谱系数（Mel-scale Frequency Cepstral Coefficients，简称MFCC）是在Mel标度频率域提取出来的倒谱参数，Mel标度描述了人耳频率的非线性特性，它与频率的关系可用下式近似表示：

### 3.2语音特征参数MFCC提取过程

基本流程：



3.2.1预处理

预加重：

预加重的目的是提升高频部分，使信号的频谱变得平坦，保持在低频到高频的整个频带中，能用同样的信噪比求频谱。

分帧、加窗：

将每一帧乘以汉明窗，以增加帧的左端和右端连续性。

3.2.2快速傅里叶变换（FFT）

由于信号在时域上的变换通常很难看出信号的特性，所以通常将它转换为频域上的能量分布来观察，不同的能量分布，代表不同语音的特性。所以在乘上汉明窗后，每帧还必须再经过快速傅里叶变换以得到在频谱上的能量分布。

3.2.3谱线能量

对语音信号的频谱取模平方得到语音信号的谱线能量。

3.2.4计算通过Mel滤波器的能量

将能量谱通过一组Mel尺度的三角形滤波器组，定义一个拥有M个滤波器的滤波器组，采用的滤波器为三角滤波器，中心频率为f(m)。

三角滤波器的频率响应定义为:

计算每个滤波器组输出的对数能量为 ：

3.2.5计算DCT倒谱

经离散余弦变换（DCT）得到MFCC系数 :

将上述的对数能量带入离散余弦变换，求出L阶的Mel参数。L阶指MFCC系数阶数，通常取12-16,本实验中取13，M是三角滤波器个数。

### 3.3主成分分析方法(PCA)

PCA的主要思想是将n维特征映射到k维上，这k维是全新的正交特征也被称为主成分，是在原有n维特征的基础上重新构造出来的k维特征。PCA的工作就是从原始的空间中顺序地找一组相互正交的坐标轴，新的坐标轴的选择与数据本身是密切相关的。其中，第一个新坐标轴选择是原始数据中方差最大且协方差为0的方向，第二个新坐标轴选取是与第一个坐标轴正交的平面中使得方差最大协方差为0的，第三个轴是与第1,2个轴正交的平面中方差最大协方差为0的。依次类推，可以得到n个这样的坐标轴。通过这种方式获得的新的坐标轴，我们发现，大部分方差都包含在前面k个坐标轴中，后面的坐标轴所含的方差几乎为0。于是，可以忽略余下的坐标轴，只保留前面k个含有绝大部分方差的坐标轴。事实上，这相当于只保留包含绝大部分方差的维度特征，而忽略包含方差几乎为0的特征维度，实现对数据特征的降维处理。

### 3.4PCA算法两种实现方法

3.4.1 基于特征值分解协方差矩阵实现PCA算法

输入：数据集，需要降到k维。

3.4.1.1 去平均值(即去中心化)，即每一位特征减去各自的平均值。

3.4.1.2 计算协方差矩阵 。

3.4.1.3 用特征值分解方法求协方差矩阵的特征值与特征向量。

3.4.1.4 对特征值从大到小排序，选择其中最大的k个。然后将其对应的k个特征向量分别作为行向量组成特征向量矩阵P。

3.4.1.5 将数据转换到k个特征向量构建的新空间中，即Y=PX。

3.4.2 基于SVD分解协方差矩阵实现PCA算法

输入：数据集，需要降到k维。

3.4.2.1 去平均值，即每一位特征减去各自的平均值。

3.4.2.2 计算协方差矩阵。

3.4.2.3 通过SVD计算协方差矩阵的特征值与特征向量。

3.4.2.4 对特征值从大到小排序，选择其中最大的k个。然后将其对应的k个特征向量分别作为列向量组成特征向量矩阵。

3.4.2.5 将数据转换到k个特征向量构建的新空间中。

## 实验内容

### 4.1提取一帧浊音信号的MFCC特征（13维的倒谱参数）

### 4.2对提取的MFCC特征做PCA聚类分析

## 实验结果

### 5.1提取语音信号的MFCC特征（13维的倒谱参数）

代码：

clear all;

[x,fs]=audioread('D:\浊音b(2).mp3'); %浊音的获取

x=x(:,1);

x=filter([1-0.9375],[1],x); %预处理高频预加重，抵消频谱倾斜

wlen=960; %设置窗长

win=hamming(wlen);

inc=400; %设置帧移

xf=enframe(x,wlen,inc)'; %分帧、加汉明窗

n3=length(xf);

n4=fix(wlen/2)+1;

p=24;

bank=v\_melbankm(p,wlen,fs,0,0.5,'m'); %Mel滤波器

bank=full(bank);

bank=bank/max(bank(:)); %归一化Mel滤波器组系数

for k=1:13 %DCT系数

n=0:p-1;

detcoef(k,:)=cos((2\*n+1)\*k\*pi/(2\*p)); %p为滤波器个数

end

for i=1:size(xf,2) %计算每帧mfcc参数

v1=xf(i,:);

s=v1';

t=abs(fft(s));

t=t.^2;

c1=detcoef\*log(bank\*t(1:n4));

m1(i,:)=c1';

end

---------------------------------------------------------------------

[y,fs]=audioread('D:\清音f(2).mp3'); %浊音的获取

y=y(:,1);

y=filter([1-0.9375],[1],y); %预处理 高频预加重，抵消频谱倾斜

wlen=960; %设置窗长

win=hamming(wlen);

inc=400; %设置帧移

yf=enframe(y,wlen,inc)'; %分帧、加汉明窗

n3=length(yf);

n4=fix(wlen/2)+1;

p=24;

bank=v\_melbankm(p,wlen,fs,0,0.5,'m'); %Mel滤波器

bank=full(bank);

bank=bank/max(bank(:)); %归一化Mel滤波器组系数

for k=1:13 %DCT系数

n=0:p-1;

detcoef(k,:)=cos((2\*n+1)\*k\*pi/(2\*p)); %p为滤波器个数

end

w = size(yf,2)

for i=1:size(yf,2) %计算每帧mfcc参数

v2=yf(i,:);

s=v2';

t=abs(fft(s));

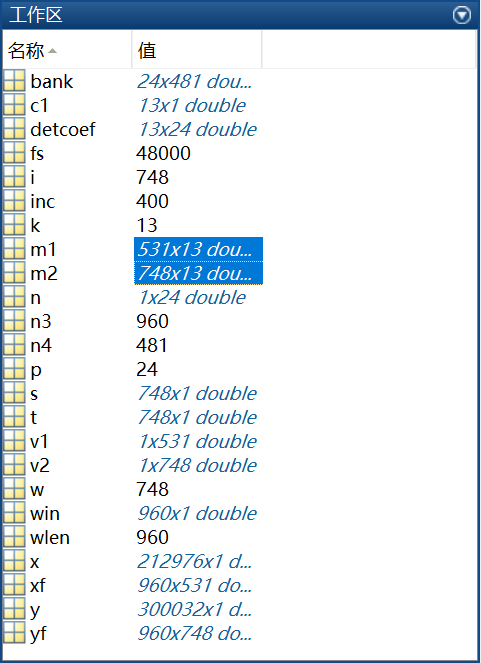
t=t.^2;

c1=detcoef\*log(bank\*t(1:n4));

m2(i,:)=c1';

end

运行结果：



### 5.2对提取的MFCC特征做PCA聚类分析

代码：

[pcaData1,COEFF3] = fastPCA(m1, 2 );

subplot(2,1,1);

plot(pcaData1(:,1),pcaData1(:,2),'b.');

title('降为2维后浊音的mfcc')

[pcaData2,COEFF3] = fastPCA(m2, 2 );

subplot(2,1,2);

plot(pcaData2(:,1),pcaData2(:,2),'b.');

title('降为2维后清音的mfcc')

%函数说明

function [pcaA,V] = fastPCA( A, k )

% 快速PCA

% 输入：A --- 样本矩阵，每行为一个样本

% k --- 降维至 k 维

% 输出：pcaA --- 降维后的 k 维样本特征向量组成的矩阵，每行一个样本，列数 k 为降维后的样本特征维数

% V --- 主成分向量

[r,c] = size(A);

meanVec = mean(A); % 样本均值

Z = (A-repmat(meanVec, r, 1));

covMatT = Z \* Z'; % 计算协方差矩阵的转置 covMatT

[V,D] = eigs(covMatT, k); % 计算 covMatT 的前 k 个本征值和本征向量

V = Z' \* V; % 得到协方差矩阵 (covMatT)' 的本征向量

for i=1:k % 本征向量归一化为单位本征向量

V(:,i)=V(:,i)/norm(V(:,i));

end

pcaA = Z \* V; % 线性变换（投影）降维至 k 维

End % 保存变换矩阵 V 和变换原点 meanVec

运行结果：

