Fbank特征代码参考：<https://blog.csdn.net/wxb1553725576/article/details/78048546>

关于Fbank与MFCC的不同可以参考这篇博客：

https://blog.csdn.net/sun\_\_\_shy/article/details/82994455

1. pre\_process.m程序：

函数功能：该函数是实现了对sx438.wav的文件的一些预处理操作，并加入了一些随机噪声

Figure1:

上面这幅图是原始信号波形，下面的图是将每个幅值加了300的直流分量，放大看可以看到前面直流分量在300处，直流分量dc offset



Figure2:

该音频文件的信息为使用了这个程序获得，并且由于只有347帧，所以在后续画图的时候是从51处开始画，plot(frames(51,:))，这个从51处开始画，不知道为何？？？

%frames(51,:)选取的是第51帧信号作为example 并不是从51帧开始画！！！

三张图自上而下的解释是第一张图是没有去除整帧信号的均值的幅值图，第二张图是去除了mean值的幅值图，第三张图是将使用了一个后一个值减去前一个值\*0.97的操作，是属于预加重操作

见下页



Figure3：画出汉明窗函数图

hamming = zeros(1, num\_of\_per\_frame);

for i = 1:num\_of\_per\_frame

hamming(1,i) = 0.54 - 0.46\*cos(2\*pi\*(i-1) / (num\_of\_per\_frame-1));

end

figure;

plot(hamming); title('Hamming Window Function'); hold on;



Figure4:画出预加重后的波形图(上图)，和加窗后的预加重波形图



1. pre\_process\_fbank.m函数

主要是包含之前的几幅图，以及以下两幅图，细致的描述了信号与mel域三角滤波器，以及两者相乘也就是时域卷积的结果





1. pre\_process\_fbank\_mfcc.m

这个图就是最后s438.wav的mfcc特征，全选mfcc\_frames矩阵，然后选择绘图中的imagesc即可显示下图，mfcc\_frames是347\*40的矩阵，347帧，40列，也就是40个





看了下该函数里的fbank\_frames也就是fbank特征，也是347\*20 double类型数组，其imagesc图如下图所示：



其中挑选出该条语音信号的一帧mfcc特征，40代表的是40个特征，不是frames，plot如下图所示：

