这周结合着论文《Comparison of techniques for environmental》，开始着手研究代码的编写，但主要力度放在了论文上。

我们知道声学场景分类可以分为两个阶段：特征提取以及分类。在每个阶段选择怎样的方法会决定我们得到结果的精准度。这篇论文主要写了特征提取技术以及分类技术的选择以及实现，并且对所提到的技术进行了比较。

与以往看到的论文不同的是，这篇文章中指出特征提取可以分为平稳和非平稳。平稳特征提取产生一个整体的结果，详细说明了整段信号里所包含的频率。在平稳特征提取中，信号中的频率位置没有区别。相反，非平稳特征提取将信号分割成离散时间单元，这允许频率被识别为发生在信号的特定区域，帮助我们理解信号。

平稳的特征提取方法有：

·固有频率提取（音乐和语音）

·同态倒谱系数

·梅尔频率倒谱系数

·线性预测倒谱系数

·梅尔频率线性预测系数

·Bark频率倒谱系数

·Bark频率线性预测系数

·感知线性预测特征

我们实现的同态倒谱系数（HCC）算法是基于听觉工具箱中的MFCC算法。该算法被修改，以产生HCC，而不是消除与梅尔频率滤波器的卷积。

与Mel频率滤波器的卷积。

同态倒谱系数：为了应用这种方法，我们首先用汉明窗分割信号。然后我们为每个加窗部分计算倒谱（）。倒谱是对数频谱的傅立叶变换。一旦我们这样做了，我们可以用以下计算倒谱系数：

 

其中



是加窗部分被操纵的长度。我们选取了由这一式子产生的前13个系数。然后用向量表示法表示这些特征，这是一个更好的训练网络的办法。与MFCC一样，特别注意删除向量中的第一个标量，它代表总信号功率，因此对信号幅度太敏感

非平稳的特征提取方法有：

•短时傅里叶变换（STFT）

•快（离散）小波变换（FWT）

•连续小波变换（CWT）

•Wigner-Ville分布（WVD）

短时傅立叶变换基本思想是将信号加滑动时间窗，并对窗内信号做傅立叶变换，得到信号的时变频谱。因而它的时间分辨率和频率分辨率受Heisenberg测不准原理约束，一旦窗函数选定，时频分辨率便确定下来。这就使它对突变信号和非平稳信号的分析存在局限性，因而不是一种动态的分析方法, 不能敏感地反映信号的突变，只适用于对缓变信号的分析。  
  Wigner一Ville分布定义为信号中心协方差函数的傅立叶变换，它具有许多优良的性能，如对称性、时移性、组合性、复共扼关系等，不会损失信号的幅值与相位信息，对瞬时频率和群延时有清晰的概念。  
其不足是不能保证非负性，尤其是对多分量信号或具有复杂调制规律的信号会产生严重的交叉项干扰，这是二次型时频分布的固有结果，大量的交叉项会淹没或严重干扰信号的自项，模糊信号的原始特征。后续的有人对Cohen类中的核函数进行改造，提出了伪winger—ville分布、修正平滑伪Winger—Ville分布等各种各样的新型时频分布，对交叉项干扰的抑制起了较大的作用，但是不含有交叉项干扰且具有Winger—Ville分布聚集性的时频分布是不存在的。  
  小波变换通过伸缩和平移运算对信号进行多尺度分解,能够有效地从信号中获取各种时频信息, 它在时域和频域同时具有良好的局部化性质，具有多分辨率分析特性。  
但小波分解的结果依赖小波基函数，而各小波基函数的适用范围很不一致，这就造成了小波基选择问题，如果母小波选择不当，则应用效果会大受影响；小波分析不具有自适应性，一旦选择了小波基和分解尺度，则用它来分析具多频率成分的数据时，所得结果只能反映某一固定频带内的信号,所以要选择不同的小波基。

分类方法：

·动态时间规整（DTW）

·隐藏马尔可夫模型（HMM）

·学习矢量量化（LVQ）

·自组织地图（SOM）

·隐马尔可夫模型的各个状态

·人工神经网络（ANN）

·长期统计

·最大似然估计（MLE）

·高斯混合模型（GMM）

·支持向量机（SVM）

另外，文章中还提出了一种采样方法——刀切法，虽说没有提及具体内容，这里我去百度了一下：1949年，Quenouille提出了刀切法，这是近代重采样方法的标志，以后，由Quenouille(1949，1956)和Tukey(1958)不断完善，重采样方法成为统计学的重要方法之一。刀切法的原始动机是降低估计的偏差。常用做法是：每次从样本集中删除一个或者几个样本，剩余的样本成为“刀切”样本，由一系列这样的刀切样本计算统计量的估计值。从这一批估计值，不但可以得到算法的稳定性衡量(方差)，还可以减少算法的偏差。这个方法暗示，刀切法的样本集需要事先给定，即，它的重采样过程是在给定样本集上的采样过程。