这周主要看了老师推荐的论文《Combining Temporal Features by Local Binary Pattern For Acoustic Scene Classification》，对比了之前看的关于声学场景的论文，发现在声学场景方面，MFCC是用得比较多的特征，以及HMM和GMM模型。

这篇论文提出了利用LBP描述符捕获帧级MFCC特征的时间演化，LBP特征可以编码这一重要信息，可以跟MFCC特征有效的结合以提高整个分类的性能。此外，一些互补的频谱特征被添加到MFCC和LBP特征提供的嵌入在频谱谱中的信息。此外，文章中采用了D3C分类器，它是一种基于k-均值聚类、动态选择和循环组合等的集成删减的混合模型。文章首先分析了LBP、邻域大小以及归一化方法对声波场景开发数据集的影响。然后，进行广泛的实验，探讨了补充频谱特征的影响，发现将特征向量加入到MFCC和LBP特征中可以得到80.3%的分类精度。

在这篇论文中提到的声学特征分为三类：MFCC特征、LBP特征、补充的频谱特征。

1. MFCC特征

梅尔频率倒谱系数被广泛的应用于声学信号的分析。尽管梅尔倒谱系数最开始是为语音处理设计的，他们在ASC领域仍然是作为提取特征的选择。 为了梅尔倒谱系数而使用的滤波器是用来模拟人类听觉系统的一些重要特性的，因此，一些可识别的特征可以被获得。在此，我们在[0 1k]、[1k 4k]和[4k 11k]的频率范围内用长度为23ms的时间窗提取了梅尔频率倒谱系数的次能带，然后这些次能带特征被集成在一起。中值和方差被当做基本统计特征。

B、LBP特征

利用LBP去捕捉由直接整合（用平均值或者方差）处理时间动态过程中遗失的信息。由于它的稳定的、单调的灰阶变化和低占用计算内存，LBP被广泛的用于图像识别和成功的在像素单位上赋予了局部模式稳健和波动的特征。

文中在局部循环邻域中定义结构T，将灰度图像的（P，R）作为采样点P（P>1）的灰度图像的联合分布的表示。与结构特征相关的信息可以被保存和表示为：

 (8)

对应着局部中心像素为（x，y）的灰度值和指的是间隔相等的像素灰度值在形成一个圆对称度量邻近集的半径为R（r＞0）的圆内。可以看出，LBP描述符对中心像素值的变化是不变的。

到目前为止，我们得到了一个高度辨识能力的结构，它记录了直方图中每个像素附近的各种模式的出现。此外，通过只考虑差异性而不是他们的精确度，我们可以实现灰度的刻度不变性：

(9)

是被给出的阈值函数：

(10)

因此，中心像素为(x,y) 的纹理信息可以通过特别的LBP数表示：

(11)

1）均匀的LBP（）：当位0 / 1变化的数目小于或等于2时，一个LBP描述符可以被视为是一个均匀模式。例如，模式11111111（0过渡），11100011（2过渡）是均匀的，而像01000100（4转换）、10010101（6转换）这样的模式是不一致的。在实践中，每个均匀LBP的出现都记录在一个独特的直方图里中，当所有非均匀LBP被分类为一个类别时，他们会被放入直方图中同一个库中。

1. 具有旋转不变性的均匀LBP（）：进一步来说，当图层旋转时，旋转不变的均匀LBP是用来避免LBP操作符的值的改变。通过旋转将相同的值转换到直方图中相同的存储夹过后，定义了这些LBP特征，给：

在对应的位转换数。



对应着逐位转换的数量

1. LBP特征的计算复杂度：对于每一个输入的音频声音而言，我们计算在第帧里所有的N维梅尔倒谱系数特征中的P点附近的LBP特征。计算所有LBP操作符的耗时为O()。因此，计算复杂度是具有恒定的MFCC维数和采样点的声音总长度的线性函数，这些特征可以实时计算。

C、补充频谱特征

各种各样的声谱特征可以用来识别频谱图中插入的信息。为了获得补充信息，我们从多个次能带中提取每一帧的声谱特征。一些常用的在频率轴上的测量在下文中被给出：

让我们定义一段音频信号的第i帧为，其频谱为；然后我们将频谱分为M个没有重叠的次能带。每一个次能带的频率变化范围是（，）。

频谱质心：被定义为所给次能带的加权平均频率，是计算每一个次能带的重心的好方法，它被给出：



1、频谱频带宽度：被定义为在一条次能带中，从频率到频谱质心的加权平均距离。这个特征展现了每一条次能带的相关分布。它由以下公式计算的出：



2、频谱带能：展现了每一条次能带的常规能量。SBE显示了环境声音的能量分布，因此，它说明了只要频率的范围，它由以下公式给出：



3、频谱平坦度：描述了在不同次能带中频率分布的频谱一致性。因此，类似噪音的声音趋于高值得频谱平坦度。我们可以用以下公式计算：



4、频谱顶端因子：提供了在每一条次能带中识别弱能量频谱的方法。对于不同的频谱顶端因子来说，类噪声的声音有着低的频谱顶端因子。它由以下公式给出:



5、香农熵：是一个选择每一条次能带的随机性的方法，我们可以根据以下公式给出：



6、Renyi信息熵：是描述每一条次能带的随机性的另一种方法，它被定义为：

