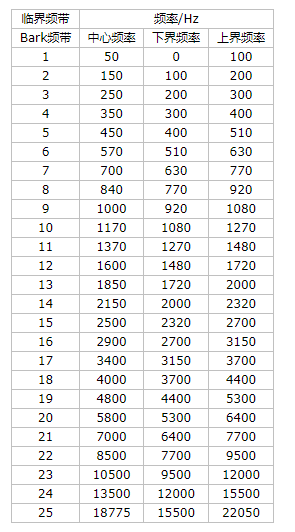
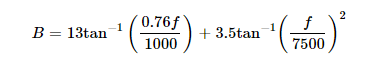
由于人耳对声音的感知（如：频率、音调）是非线性的，为了对声音的感知进行度量，产生了一系列的尺度（如：十二平均律），这里重点说下Bark尺度与Mel尺度。刚开始的时候，我自己也没弄明白这两个尺度的区别。后来才逐渐的明白他们的思考出发点，这里简单分享出来。

Bark（巴克）频率尺度是以Hz为单位，把频率映射到心理声学的24个临界频带上，第25个临界频带占据约：16K~20kHz的频率，1个临界频带的宽度等于一个Bark，简单的说，Bark尺度是把物理频率转换到心理声学的频率。Bark尺度频率的中心频率与临界带宽边界频率如下表所示：



我发现有不少式子试图对上表进行建模，用的比较多的一个式子是（Zwicker，Terhardt 1980）：



上式中频率f代表中心频率，我把上面的式子在Matlab中代入进行计算，发现前5个Bark频带与计算出来的出入比较大，尚不知道这是什么原因。

　Mel频率尺度也是一种频率映射感知模型，它描述的是音高感知的非线性映射，函数表示如下：



人类对不同频率语音有不同的感知能力：对1kHz以下，与频率成线性关系，对1kHz以上，与频率成对数关系。频率越高，感知能力就越差了。因此，在应用中常常只使用低频MFCC，而丢弃中高频MFCC。

在Mel频域内，人对音调的感知能力为线性关系，如果两段语音的Mel频率差两倍，则人在感知上也差两倍。

要注意的一点是，这里的频率f单位是1kHz，也就是说，1kHz是Mel频率与以赫兹为单位的真实频率之间的参考点（1kHz=1000mel），另一方面，Mel一词来源于音乐术语melody，是旋律的频率分量与与音高感知之间距离的度量。