本词条由[“科普中国”科学百科词条编写与应用工作项目](https://baike.baidu.com/science) 审核 。

在声音处理领域中，**梅尔频率倒谱**(Mel-Frequency Cepstrum)是基于声音频率的非线性[梅尔刻度](https://baike.baidu.com/item/%E6%A2%85%E5%B0%94%E5%88%BB%E5%BA%A6" \t "_blank)(mel scale)的[对数](https://baike.baidu.com/item/%E5%AF%B9%E6%95%B0" \t "_blank)能量频谱的线性变换。

**梅尔频率倒谱系数** (**Mel-Frequency Cepstral Coefficients，MFCCs**)就是组成梅尔频率倒谱的系数。它衍生自音讯片段的[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1)(cepstrum)。倒谱和梅尔频率倒谱的区别在于，梅尔频率倒谱的频带划分是在[梅尔刻度](https://baike.baidu.com/item/%E6%A2%85%E5%B0%94%E5%88%BB%E5%BA%A6" \t "_blank)上等距划分的，它比用于正常的对数[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1)中的线性间隔的频带更能近似人类的听觉系统。 这样的非线性表示，可以在多个领域中使声音信号有更好的表示。例如在音讯压缩中。

在声音处理领域中，**梅尔频率倒谱**(Mel-Frequency Cepstrum)是基于声音频率的非线性梅尔刻度(mel scale)的[对数](https://baike.baidu.com/item/%E5%AF%B9%E6%95%B0" \t "_blank)能量频谱的线性变换。

**梅尔频率倒谱系数**(**Mel-Frequency Cepstral Coefficients，MFCCs**)就是组成梅尔频率倒谱的系数。它衍生自音讯片段的[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1" \t "_blank)(cepstrum)。倒谱和梅尔频率倒谱的区别在于，梅尔频率倒谱的频带划分是在梅尔刻度上等距划分的，它比用于正常的对数[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1" \t "_blank)中的线性间隔的频带更能近似人类的听觉系统。 这样的非线性表示，可以在多个领域中使声音信号有更好的表示。例如在音讯压缩中。

梅尔频率倒谱系数（MFCC）广泛被应用于[语音识别](https://baike.baidu.com/item/%E8%AF%AD%E9%9F%B3%E8%AF%86%E5%88%AB" \t "_blank)的功能。他们由Davis和Mermelstein在1980年代提出，并在其后持续是最先进的技术之一。在MFCC之前，线性预测系数（LPCS）和线性预测倒谱系数（LPCCs）是[自动语音识别](https://baike.baidu.com/item/%E8%87%AA%E5%8A%A8%E8%AF%AD%E9%9F%B3%E8%AF%86%E5%88%AB/5807980" \t "_blank)的的主流方法。

　　MFCC通常有以下之过程:

1. 将一段语音信号分解为多个讯框。
2. 将语音信号预强化，通过一个[高通](https://baike.baidu.com/item/%E9%AB%98%E9%80%9A)[滤波器](https://baike.baidu.com/item/%E6%BB%A4%E6%B3%A2%E5%99%A8)。
3. 进行[傅立叶变换](https://baike.baidu.com/item/%E5%82%85%E7%AB%8B%E5%8F%B6%E5%8F%98%E6%8D%A2" \t "_blank)，将信号变换至频域。
4. 将每个讯框获得的频谱通过梅尔滤波器(三角重叠窗口)，得到梅尔刻度。
5. 在每个梅尔刻度上提取[对数](https://baike.baidu.com/item/%E5%AF%B9%E6%95%B0)能量。
6. 对上面获得的结果进行[离散傅里叶反变换](https://baike.baidu.com/item/%E7%A6%BB%E6%95%A3%E5%82%85%E9%87%8C%E5%8F%B6%E5%8F%8D%E5%8F%98%E6%8D%A2" \t "_blank)，变换到[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1)域。
7. MFCC就是这个[倒频谱](https://baike.baidu.com/item/%E5%80%92%E9%A2%91%E8%B0%B1" \t "_blank)图的幅度(amplitudes)。一般使用12个系数，与讯框能量叠加得13维的系数。
8. MFCC的原理
9. [编辑 语音](javascript:;)
10. 声音信号是连续变化的，为了将连续变化信号简化，我们假设在一个短时间尺度内，音频信号不发生改变。因此将信号以多个采样点集合成一个单位，称为'''讯框'''。一个讯框多为20-40毫秒，如果讯框长度更短，那每个讯框内的采样点将不足以做出可靠的频谱计算，但若长度太长，则每个讯框信号会变化太大。
11. **预强化**的目的就是为了消除发声过程中，声带和嘴唇造成的效应，来补偿语音信号受到发音系统所压抑的高频部分。并且能突显高频的共振峰。
12. 由于信号在时域上的变化通常很难看出信号的特性，所以通常透过[傅里叶变换](https://baike.baidu.com/item/%E5%82%85%E9%87%8C%E5%8F%B6%E5%8F%98%E6%8D%A2" \t "_blank)将它变换成频域上的能量分布来观察，不同的能量分布，就能代表不同语音的特性。
13. 由于能量频谱中还存在大量的无用讯息，尤其人耳无法分辨高频的频率变化，因此让频谱通过梅尔滤波器。**梅尔滤波器**，也就是一组20个非线性分布的**三角带通滤波器（Triangular Bandpass Filters）**，能求得每一个滤波器输出的对数能量。必须注意的是：这 20 个三角带通滤波器在['''梅尔刻度'''](https://baike.baidu.com/item/%27%27%27%E6%A2%85%E5%B0%94%E5%88%BB%E5%BA%A6%27%27%27" \t "_blank)的频率上是平均分布的。 梅尔频率代表一般人耳对于频率的感受度，由此也可以看出人耳对于频率 f 的感受是呈对数变化的。[1]
14. 噪声的敏感性
15. [编辑 语音](javascript:;)
16. MFCC特征在加性噪声的情况下并不稳定，因此在[语音识别系统](https://baike.baidu.com/item/%E8%AF%AD%E9%9F%B3%E8%AF%86%E5%88%AB%E7%B3%BB%E7%BB%9F/5277928" \t "_blank)中通常要对其进行归一化处理(normalise)以降低噪声的影响。一些研究人员对MFCC算法进行修改以提升其鲁棒性，如在进行DCT之前将log-mel-amplitudes提升到一个合适的能量(2到3之间)，以此来降低低能量成分的影响。[1]