

一个音调”的对数-对数图就展现出来了,如图7-12a所示。图中曲线的给出直接结果是:人们并没有必要对那些功率在可听阈值之下的频率编码。例如,在图7-12a中,如果100Hz的功率是20dB,那么在输出上就可以忽略掉,而且不会感觉到声音质量的损失,因为在100Hz处20dB是低于可听水平的。

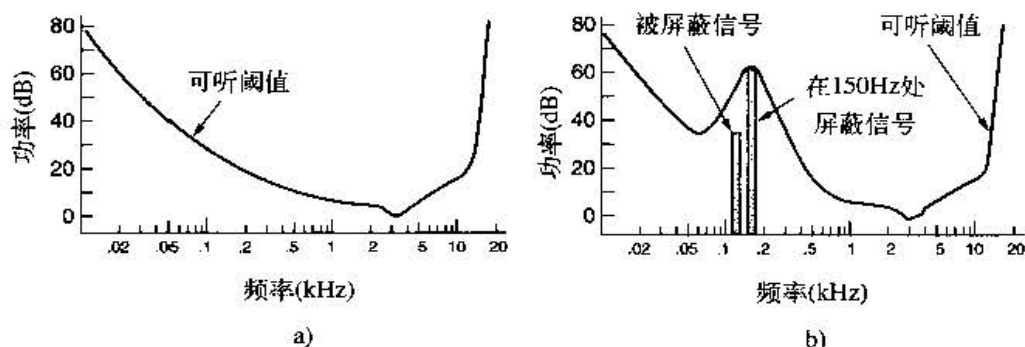


图7-12 a) 作为频率函数的可听阈值; b) 屏蔽效应

现在考虑实验2。计算机再次运行实验1,但是这次却一有个大约150赫兹的等幅正弦波叠加在实验频率上。我们发现,在150Hz频率附近的可听阈值上升了,如图7-12b所示。

这一新实验的结果表明:通过跟踪那些被附近频段能量更强的信号所屏蔽的信号,可以省略越来越多的编码信号中的频率,以此来节约二进制位。在图7-12中,125Hz信号的输出是可以完全忽略掉的,并且没有人能够听出其中的不同。甚至当某个频段中的一个强大信号停止后,出于对暂时屏蔽这一知识的了解,也会让我们在耳朵恢复期的时间段内省略掉那些被屏蔽的频率。MP3编码的实质就是对声音做傅立叶变换从而得到每个频率的能量,之后只传递那些不被屏蔽掉的频率,并且用尽可能少的二进制位数来编码这些频率。

有了这些信息作为背景,现在来考察有关编码是如何完成的。通过抽取32kHz、44.1kHz或者48 kHz的波形,完成声音压缩。第一个和最后一个都是四舍五入的整数。44.1kHz是用于Audio CD的,因为这个值能很好地捕获人耳可听到的所有音频信息。可以在以下四个配置中任选一个,用一或两个通道完成抽样:

- 1) 单声道(一个输入流)。
- 2) 双声道(例如,一个英语的和一個日语的音軌)。
- 3) 分立立体声(每个通道分开压缩)。
- 4) 联合立体声(完全利用通道间的冗余)。

首先,选择输出的比特率。MP3可以将摇滚CD的立体声降低到96kbps,并且在质量上几乎没有任何失真,甚至连摇滚迷都听不出差别。而对于一场钢琴音乐会,至少需要128kbps。造成这样不同的原因是因为摇滚的信噪比要比一场钢琴音乐会要高得多(至少从工程角度上看)。也可以选择稍低一点的输出比率,接受质量上的少许失真。

然后将这些样本处理成1152(大概26ms)的一些组,每组首先通过32个数字滤波器,获得32个频率波段。同时,将输入放进一个心理声学的模型中,测定被屏蔽的频率。接下来,进一步转换32频率波段中的每一个,以提供一个更精确的频谱解决方案。

接着,将现有的二进制位分配到各个波段中,大部分二进制位分配给拥有多数频谱能量的未屏蔽波段,小部分二进制位分配给拥有较少频谱能量的未屏蔽波段,已屏蔽的波段不分配二进制位。最后,用霍夫曼编码来对这些二进制位进行编码,它可以将经常出现的数字赋予较短的代码,而对不常出现的数字赋予较长的代码。

实际的工作过程更复杂。为了减少噪音,消除混淆,以及利用通道间冗余,需要各种各样的技术,不过这些内容超出了本书的范围。

## 7.5 多媒体进程调度

支持多媒体的操作系统与传统的操作系统在三个主要的方面有所区别:进程调度、文件系统和磁盘调度。本节中我们开始讨论进程调度,在后面的各节中接着讨论其他主题。