第二章 语音信号的数字处理

2.1引言

实际的语音信号代表空气的密度随时间的变化，基本上是一个连续的函数，但是若要将此信号储存在计算机里，就必须先将此信号数字化。

为了给下一步语音增强做准备，需要先提取信号的特征参数，考虑到信号的短时平稳性，可以对信号进行短时处理以达到上述目的。另外，为了增加语音的高频分辨率及提取音频特征，克服采集信号是所带来的混叠，我们要先对信号进行预处理。

2.2语音信号预处理

2.2.1语音信号的采集

一般而言，当我们将声音储存为数字信号时，有一下几个参数需要考虑：

（1）采样频率（Sample Rate）：每秒种所取得的声音资料点数，以赫兹（Hz）为单位，采样频率越高声音品质越好，但是数据量越大。常用采样频率如下：

1. 8KHz：电话及一般玩具内语音芯片的音质

2. 16KHz：语音识别系统的音质

3. 44.1KHz：CD音质

实际采集音频时使用的频率依照奈奎斯特采样定理来确定。

（2）采样解析度（Bit Resolution）：存储采样后的数字信号时使用的位元数，常用的数值如下：

1. 8bit：可表示的数值范围为0-255 或-128-127

2. 16bit：可表示的数值范围为-32768-32767

换句话说，每个取样点的数值都是整数，以方便储存。但是在MATLAB 中，通常把音频信号的值正规化到[-1, 1] 范围内的浮点数，因此若要转回原先的整数值，就必须再乘上，其中nbits是采样解析度。

（3）声道：一般只分单声道（Mono）或立体声（Stereo），立体音即是双声道。