

基于 MATLAB 语言实现 语音信号的采集与分析

邵铭浩 大连民族学院 116600

【文章摘要】

通过语音传递信息是人类最重要、最有效且最为方便的交换信息形式。语音信号的采集与分析技术是一门涉及面很广的综合性交叉科学,其应用与语音学、声音测量学以及数字信号处理等诸多学科有着不可分割的紧密联系。本文简要介绍了简短语音信号的采集与分析方法,并通过 PC 机录制一段声音,最后用 MATLAB 语言进行仿真和分析。实现简短语音信号较为直观和形象地展现。

【关键词】

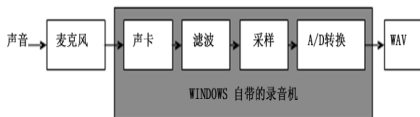
语音信号;采集与分析;MATLAB

一、引言

语言是人类特有的功能。由于语言和语音具有最大的信息容量和最高的智能水平,一直是人们构成思想沟通和感情交流的最主要的途径。人类现已开始进入了信息化时代,用现代化手段研究语音信号,使人们能更加有效地产生、传输、存储、获取和应用语音信息,对语音信号的研究也是一项极具市场价值和挑战性的工作。因其实用性使之能够长期地、深深地与当代信息科学中最活跃的前沿学科保持密切的联系,并且相互促进发展。然而许多信息处理的新方法的提出,首先是在语音信号处理中获得成功,然后才能得以推广到其他领域,由此可见,对语音信号进行处理的重要性在现如今的科学领域是举足轻重的,而对语音信号的处理其第一步便是对语音信号进行采集与分析。

二、语音的录制与播放

由于语音信号是一种连续变化的模拟信号,而计算机只能处理和记录二进制的数字信号,因此,由自然音而得的音频信号必须经过采样、量化和编码,变成二进制数据后才能送到计算机进行再编辑和存储。语音信号输出时,则与上述过程相反。用计算机的声音编辑工具进行语音信号的录制时,已经利用了计算机上的 A/D 转换器,将模拟的声音信号变成了离散的量化了的数字信号。量化了的数字信号又通过 D/A 转换器,把保存起来的数字数据恢复成原来的模拟的语音信号。采用 WINDOWS 系统自带的录音机进行录制语音文件,基于 PC 机的语音信号原理和采集过程如下图所示。



使用 MATLAB 编程中的声音处理功能语句“`y=wavrecord(m, Fs, ch)`”,在此时刻以后的 m/Fs 秒时段内的声音信号将以 y 为文件名,以数字声音信号 .wav 格式存储在 MATLAB 的工作空间里。然后使用 MATLAB 语句“`sound(y, Fs)`”,将存在工作空间的矢量 y 以 Fs 为采样频率还原为声音以用于继续进行对语音信号的采集与分析。

三、语音信号的采集与分析原理

对语音信号进行频谱分析,引入傅里叶分析即分析信号的频谱(频率构成)、频带宽度等。由于语音信号也是连续的,则对语音信号的分析也需采用傅里叶变换。对于连续时间信号 $f(t)$,其傅里叶变换 $F(\omega)$ 为: $F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt$ 连续时间傅里叶变换适用于连续时间信号的分析,但 $f(t)$ 和 $F(\omega)$ 都是连续函数,不适于计算机处理。工程应用中经常需要对抽样数据进行傅里叶分析,这种情况下往往无法得到信号的解析表达式,因而必须采用傅里叶变换的数值计算方法。即,如果 $f(t)$ 的主要取值区间为 $[t_1, t_2]$,定义 $T = t_2 - t_1$ 为区间长度。在该区间内抽样 N 个点,抽样间隔为 $\Delta t = \frac{T}{N}$,则有:

$$F(\omega) = \sum_{n=0}^{N-1} f(t_1 + n\Delta t)e^{-j\omega(t_1 + n\Delta t)} \Delta t = \Delta t \cdot \sum_{n=0}^{N-1} f(t_1 + n\Delta t)e^{-j\omega(t_1 + n\Delta t)}$$

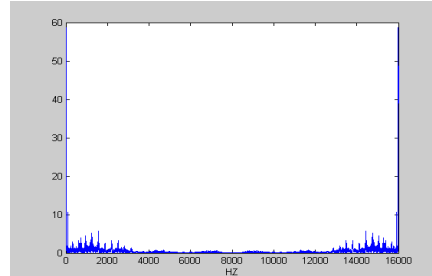
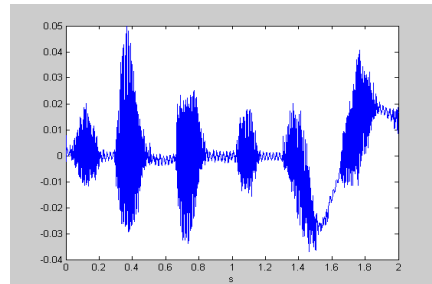
可以计算出任意频点的傅里叶变换值,假设 $F(\omega)$ 的主要取值区间位于 $[\omega_1, \omega_2]$,要计算其间均匀抽样的 k 个值,则有:

$$F(\omega_1 + k\Delta\omega) = \Delta t \cdot \sum_{n=0}^{N-1} f(t_1 + n\Delta t)e^{-j(\omega_1 + k\Delta\omega)(t_1 + n\Delta t)}$$

$\Delta\omega = \frac{\omega_2 - \omega_1}{k}$ 为频率抽样间隔。傅里叶变换后,需对其根据采样定理进行适当的采样,即利用抽样脉冲序列从连续信号 $f(t)$ 中抽取一系列离散样值构成抽样信号 $f_s(t)$ 。再利用时域抽样定理无失真地恢复原连续信号。时域抽样定理如下:一个频谱受限的信号 $f(t)$,如果频谱只占据 $-\omega_m \sim +\omega_m$ 的范围,则信号 $f(t)$ 可以用唯一的等间隔抽样值表示。抽样间隔最大为 $\frac{1}{2f_m}$ 。把最低抽样频率 $f_s = 2f_m$ 称为奈奎斯特频率,最大抽样间隔 $T_s = \frac{1}{\omega_m} = \frac{1}{2f_m}$ 称为奈奎斯特间隔。

四、仿真结果与分析

录制一段简短的语音信号,进行采样,然后用 MATLAB 软件进行仿真,得出采样后的时域波形(如上图)与频谱图(如下图):



观察频域波形,由图可知该信号的频率主要集中在 1000Hz,上限频率大约为 3000Hz。现令二倍上限频率为 6000Hz,然后分别选择等于、大于和小于两倍上限频率三种不同的采样频率重新录制该语音信号,试听回放效果,进行比较。通过观察比较可得:

等于二倍上限频率时:

`Y=wavrecord(6000, 3000, 2)`,

`Sound(y, 6000)`,此时语音信号出现失真;

大于二倍上限频率时:

`Y=wavrecord(20000, 8000, 2)`,

`Sound(y, 8000)`,此时语音信号不出现失真;

小于二倍上限频率时:

`Y=wavrecord(4000, 2000, 2)`,

`Sound(y, 2000)`,此时语音信号出现失真。

由此可得,适当的参数设置有利于对语音信号进行较好的分析,因此在进行对语音信号的采集与分析时,需要使用适当的采样间隔和各频率的选择。

五、结束语

对于语音信号的采集与分析是对语音信号进行处理关键第一步,只有在对其进行较好地操作之后才能为后续的工作打下良好的基础,本文简要介绍了对自行录制的一段简短的语音信号进行采集与分析,并验证了参数在各种设定的情况下所分析出的结果。

【参考文献】

- [1] 郑君里、应启珩、杨为理,信号与系统,高等教育出版社,2011
- [2] 李敏、陈兴文,信号分析与处理的软硬件实现,大连海事出版社,2009
- [3] 赵淑清、郑薇,随机信号分析,哈尔滨工业大学出版社,2009
- [4] 胡航,语音信号处理,哈尔滨工业大学出版社,2002
- [5] 樊昌信、曹丽娜,通信原理,国防工业出版社,2010