

语音信号处理系统设计^{*}

刘惠敏, 王 龙

(吉首大学物理与机电工程学院, 吉首 416000)

【内容提要】 本文简要介绍了语音信号采集与分析以及语音信号的特征、采集与分析方法, 并在采集语音信号后, 在 Matlab 软件平台上进行频谱分析, 然后对所采集的语音信号加入干扰噪声, 对加入噪声的信号进行频谱分析, 设计合适的滤波器滤除噪声, 恢复原信号。利用 Matlab 来读入 (采集) 语音信号, 将它赋值给某一向量, 再将该向量看作一个普通的信号, 对其进行 FFT (Fast Fourier transform) 变换实现频谱分析, 再依据实际情况对它进行滤波, 然后我们还可以通过 sound 命令来对语音信号进行回放, 以便在听觉上来感受声音的变化。

【关键词】 语音信号; 数字信号处理; 傅里叶变换; 滤波器

1 引言

语音是人类之间相互进行信息交流时使用最多、最自然、最有效、最方便的手段, 是人类语言的一种声学表现。在人类社会文明和科学技术迅速发展的今天, 人类已经走进了信息化时代。语音信号处理成为现代通信、办公自动化、多媒体技术、智能系统等新兴科研领域应用的核心技术, 通常这些信号处理的过程要满足实时且快速高效的要求, 随着 DSP 技术的发展, 以 DSP 为内核的设备越来越多, 为语音信号的处理提供了良好的平台, 使人们能更加有效地产生、传输、存储、和获取语音信息, 这对于促进社会的发展具有十分重要的意义。因此, 语音信号处理正越来越受到人们的关注和广泛的研究。

Matlab 是一种数据处理和分析功能十分强大的计算机应用软件, 它可以将声音文件转化成离散的数据文件, 并通过其强大的计算能力来处理各种数据。Matlab 信号处

^{*} 收稿日期: 2016-05-03

作者简介: 刘惠敏, 女, 湖南娄底人, 吉首大学物电学院 2012 级电子信息科学与技术专业学生; E-mail: 397996291@qq.com

王龙, 男, 湖南长沙人, 吉首大学物电学院 2012 级电子信息科学与技术专业学生; E-mail: 362271503@qq.com

理与分析工具箱为语音信号分析提供了丰富的功能，利用这些功能函数可以容易且迅速地
地完成信号的可视化及语音信号的处理和分析，使人机交互更加方便。信号处理是
Matlab 的重要应用领域之一。

2 语音信号处理系统的整体结构

图 2.1 列出了整个语音信号处理系统的工作流程：

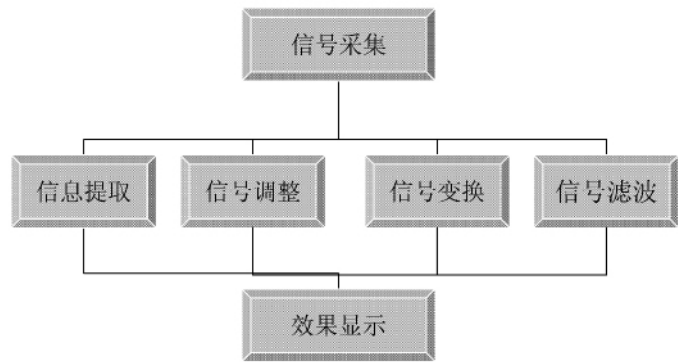


图 2.1 语音信号处理系统的工作流程

其中信号调整包括信号的幅度和频率的任意倍数变化。如下图 2.2

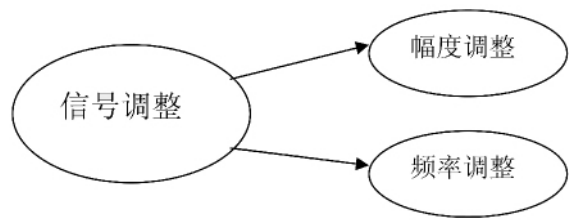


图 2.2 信号调整

信号的滤波采用了四种滤波方式，来观察各种滤波性能的优缺点：

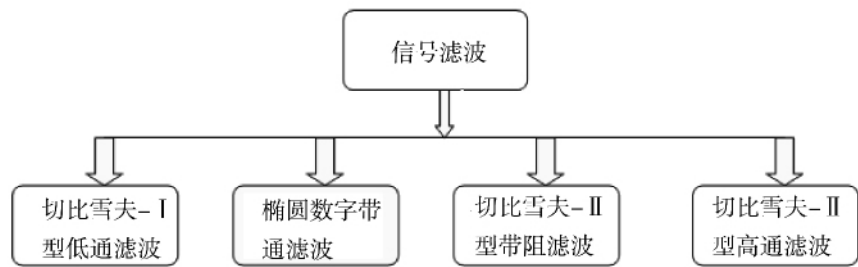


图 2.3 语音信号滤波的方式

在以上三图中，可以看到整个语音信号处理系统的流程大概分为三步，首先要读入待处理的语音信号，然后进行语音信号的处理，包括信息的提取、幅度和频率的变换以及语音信号的傅里叶变换、滤波等；滤波又包括低通滤波、高通滤波、带通滤波和带阻滤波等方式。最后对处理过的语音信号进行处理后的效果显示。以上是本系统的工作流程，本文将从语音信号的采集开始做详细介绍。

3 语音信号处理系统涉及的基本知识

3.1 数字滤波器设计原理

S 数字滤波器的作用是利用离散时间系统的特性对输入信号波形（或频谱）进行加工处理，或者说利用数字方法按预定的要求对信号进行变换。

数字滤波器可以理解为一个算法或一个计算程序，在将代表输入信号的数字时间序列转化为代表输出信号的数字时间序列的过程中，按预定的形式改变信号。数字滤波器有多种分类，根据数字滤波器冲激响应的时域特征，可将数字滤波器分为无限长冲激响应（IIR）滤波器和有限长冲激响应（FIR）滤波器两种 [1]。在性能方面，对 IIR 滤波器来说，其传输函数的极点可位于单位圆内的任何地方，因此它可以通过较低的阶数来获得较高的选择性 [2]。但是必须以牺牲相位的非线性为代价来换取这个高效率。相反，FIR 滤波器有严格的线性相位。然而，对于 FIR 滤波器来说，其传输函数的极点固定在原点，所以只能用高阶达到高的选择性。对于同样的滤波器设计指标，FIR 滤波器成本较高，信号延时也较大，因为 FIR 滤波器所要求的阶数可以比 IIR 滤波器高 5—10 倍。对于相同的线性和相同的选择性要求来说，IIR 滤波器就必须加全通网络进行相位校正，同样地大大增加滤波器的复杂性。

整体来看，IIR 滤波器达到同样效果阶数少，延迟小，但是有稳定性问题，属于非线性相位；FIR 滤波器没有稳定性问题，属于线性相位，但阶数多，延迟大。

3.1.1 巴特沃斯滤波器原理

巴特沃斯滤波器具有单调下降的幅频特性：在小于截止频率 Ω_c 的范围内，具有最平幅度的响应，而在 $A^2(\Omega) = |H_a(j\Omega)|^2 = \frac{1}{1 + \epsilon^2 C_N^2(\frac{\Omega}{\Omega_p})}$ 后，幅频响应迅速下降。

巴特沃思低通滤波器幅度平方函数为：

$$H_a(s) \quad (2.1)$$

式中 N 为滤波器阶数， $H_a(s) = \frac{\Omega_p^N}{\epsilon \cdot 2^{N-1} \cdot \prod_{i=1}^N (s - s_i)}$ 为 3dB 截止角频率。将幅度平方

函数写成 s 的函数：

$$H_a(p) = \frac{1}{\epsilon \cdot 2^{N-1} \cdot \prod_{i=1}^N (p - p_i)} \quad (2.2)$$

该幅度平方函数有 $2N$ 个等间隔分布在半径为 $p_i = -ch\xi \cdot \sin[\frac{(2k-1)\pi}{2N}] + jch\xi \cdot$

$\cos[\frac{(2k-1)\pi}{2N}]$ 的圆上的极点为：

$\xi = \frac{1}{N} \text{Arsh}(\frac{1}{\epsilon})$, $k = 0, 1 \dots 2N-1$ 为了形成稳定的滤波器，取左半平面的 N 个极点构成 $H_a(s)$ ，即：

$$H_a(s) = \frac{\Omega_c^N}{\prod_{k=0}^{N-1} (s - s_k)} \quad (2.3)$$

为使设计统一，将频率归一化，得到归一化极点 $p_k = e^{j\pi(\frac{1}{2} + \frac{2k+1}{2N})}$ ，相应的归一化系统函数为：

$$H_a(p) = \frac{1}{\prod_{k=0}^{N-1} (p - p_k)} \quad (2.4)$$

多项式形式为：

$$H_a(p) = \frac{1}{(b_0 + b_1 p + \dots + p^N)} \quad (2.5)$$

3.1.2 切比雪夫滤波器原理

切比雪夫滤波器的幅频特性具有等纹波特性，有两种形式，在通带内等纹波、阻带单调的是 I 型滤波器，在通带内单调、在阻带内等纹波的是 II 滤波器。以 I 型滤波器为例。

切比雪夫滤波器的幅度平方函数为：

$$A^2(\Omega) = |H_a(j\Omega)|^2 = \frac{1}{1 + \epsilon^2 C_N^2(\frac{\Omega}{\Omega_p})} \quad (2.6)$$

式 2.6 中 ϵ 为小于 1 的正数，表示通带内幅度波动的程度。 Ω_p 称为通带截止频率。令 $\lambda = \Omega/\Omega_p$ ，称为对 Ω_p 的归一化频率。 $C_N(x)$ 为 N 阶切比雪夫多项式。幅度平方函数的极点是分布在 $b\Omega_p$ 为长半轴， $a\Omega_p$ 为短半轴的椭圆上的点。同样取 s 平面左半平面的极点构成

$$H_a(s) = \frac{\Omega_p^N}{\epsilon \cdot 2^{N-1} \cdot \prod_{i=1}^N (s - s_i)} \quad (2.7)$$

进行归一化，得到：

$$H_a(p) = \frac{1}{\epsilon \cdot 2^{N-1} \cdot \prod_{i=1}^N (p - p_i)} \quad (2.8)$$

3.1.3 模拟滤波器数字化原理

将模拟滤波器转化为数字滤波器在工程上常用的有双线性变换法和脉冲响应不变法。

脉冲响应不变法是一种时域上的转换方法，它是数字滤波器的单位取样响应在抽样

点上等于模拟滤波器的单位冲激响应，即：

$$h(n) = h_a(nT) \quad (2.9)$$

设模拟滤波器只有单阶极点，其系统函数为：

$$(2.10)$$

对 $H_a(s)$ 进行拉氏反变换得到 $h_a(t)$ ，对 $h_a(t)$ 进行等间隔抽样，得到 $h(n) = h_a(nT)$ ，对 $h(n)$ 进行 Z 变换，得到数字滤波器系统函数：

$$H(z) = \sum_{i=1}^N \frac{A_i}{1 - e^{s_i T} z^{-1}} \quad (2.11)$$

这种方法 s 和 z 的关系是： $z = e^{sT}$ 。该方法的优点是频率坐标变换时线性的数字滤波器的单位脉冲响应完全模仿模拟滤波器的单位冲激响应，时域特性逼近好；缺点是会产生频谱混叠现象，适合低通、带通滤波器的设计，不适合高通、带阻滤波器的设计。

双线性变换法为了克服频谱混叠现象，采用非线性频率压缩方法，将整个频轴上的频率范围压缩到 $\pm \pi/T$ 之间，再用 $z = e^{sT}$ 转换到 Z 平面上。这种方法 s 和 z 的关系是： $s = (2/T)(1 - z^{-1}/1 + z^{-1})$ 。该方法克服了频谱混叠现象，但带来了频率坐标变换的非线性： $\Omega = (2/T)\tan(\omega/2)$ ，由模拟滤波器系统函数转换为数字滤波器系统函数公式为：

$$H(z) = H_a(s) \Big|_{s=\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}} \quad (2.12)$$

3.2 FFT 分析

FFT 即为快速傅氏变换，是离散傅氏变换的快速算法，它是根据离散傅氏变换的奇、偶、虚、实等特性，对离散傅立叶变换的算法进行改进获得的。在 Matlab 的信号处理工具箱中函数 FFT 和 IFFT 用于快速傅立叶变换和逆变换 [3]。函数 FFT 用于序列快速傅立叶变换，其调用格式为 $y = \text{fft}(x)$ ，其中 x 是序列， y 是 x 的 FFT， x 可以为一向量或矩阵。若 x 为一向量， y 是 x 的 FFT 且和 x 长度相同；若 x 为一矩阵，则 y 是对矩阵的每一列向量进行 FFT。如果 x 长度是 2 的幂次方，函数 fft 执行高速基——2FFT 算法，否则 fft 执行一种混合基的离散傅立叶变换算法，计算速度较慢。函数 FFT 的另一种调用格式为 $y = \text{fft}(x, N)$ ，式中， x ， y 意义同前， N 为正整数。函数执行 N 点的 FFT，若 x 为向量且长度小于 N ，则函数将 x 补零至长度 N ；若向量 x 的长度大于 N ，则函数截短 x 使之长度为 N ；若 x 为矩阵，按相同方法对 x 进行处理。

3.3 倒谱的概念

定义：倒谱定义为信号短时振幅谱的对数傅里叶反变换。

特点：具有可近似地分离并能提取出频谱包络信息和细微结构信息的特点

用途：①提取声道特征信息：提取频谱包络特征，以此作为描述音韵的特征参数而应用于语音识别；②提取音源信息：提取基音特征，以此作为描述音韵特征的辅助参数而应用于语音识别；

4 语音信号处理系统涉及的程序及结果

4.1 语音信号的提取

在 Matlab 中使用 Wavread 函数，可得出信号的采样频率为 22500，并且声音是单

声道的。利用 Sound 函数可以清晰的听到“主人，信息收到了”的语音。采集数据并画出波形图。

其中声音的采样频率 $F_s=22050\text{Hz}$ ， y 为采样数据，NBITS 表示量化阶数。

部分程序如下：

```
fn=input('Enter WAV filename:', 's'); %获取一个*.wav 的文件
[x, fs, nb]=wavread(fn);
ms2=floor(fs*0.002);
ms10=floor(fs*0.01);
ms20=floor(fs*0.02);
ms30=floor(fs*0.03);
t=(0:length(x)-1)/fs; %计算样本时刻
subplot(2,1,1); %确定显示位置
plot(t,x); %画波形图
legend('Waveform');
xlabel('Time (s)');
ylabel('Amplitude');
```

运行后弹出语音信号处理系统的操作界面如图 4.1：

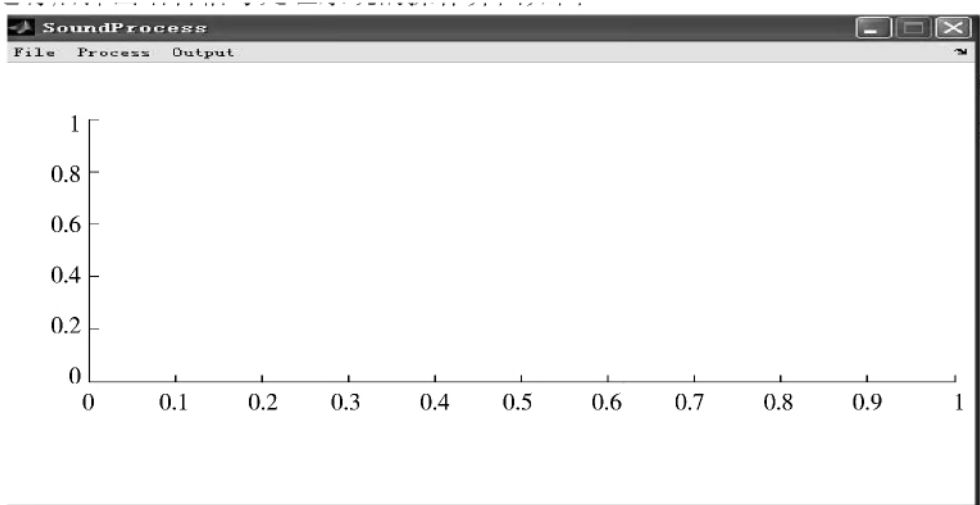


图 4.1 语音信号处理系统的操作界面

然后点击 File 菜单中的子菜单 Input，回到 Matlab 软件的输入界面如图 4.2：

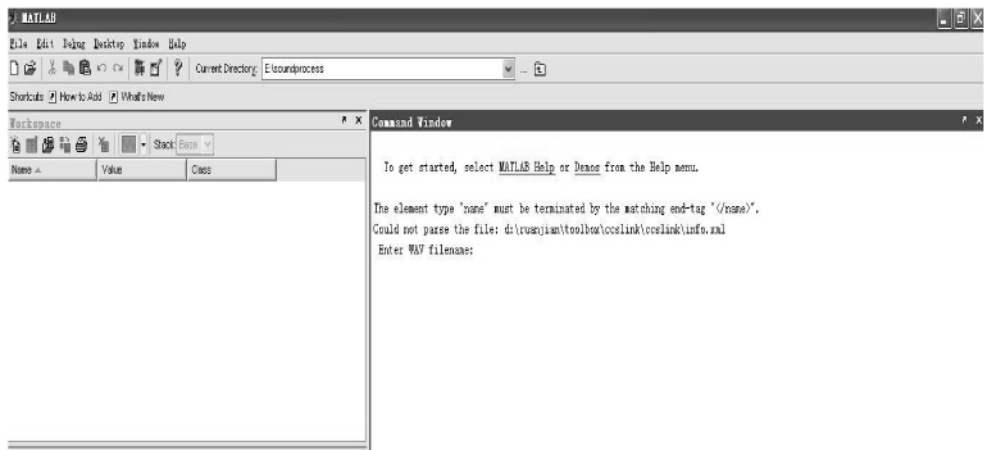


图 4.2 输入界面

输入要处理的语音信号的名称，便可得到语音语音的波形图如图 4.3：

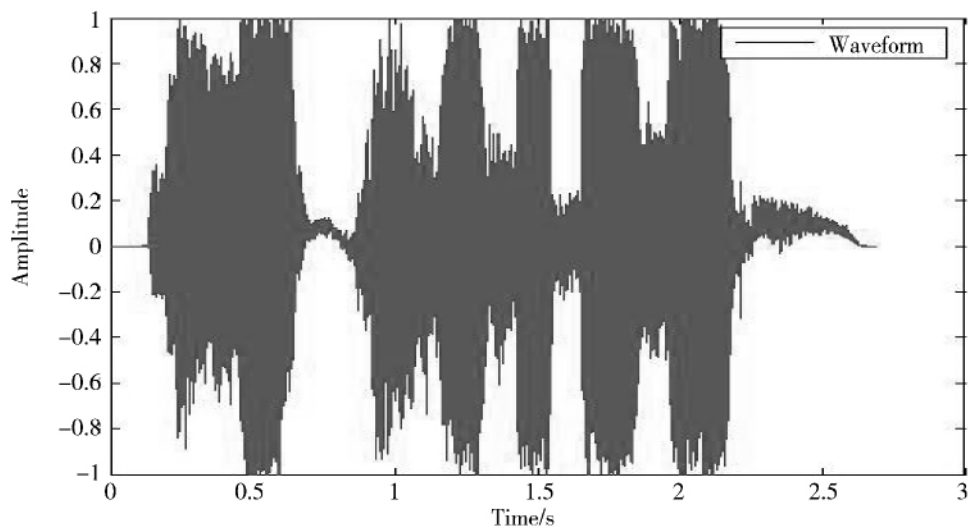


图 4.3 语音语音的波形图

如图中提取的语音的波形图所示，整段音频数据中得声音高低起伏与录入的声音信号基本一致，并且可以观察到其中包含部分高频噪声。

4.2 语音信号的调整

在语音信号的研究中，经常会对语音信号进行进行多倍频率以及多倍幅度变换调整，日常应用中，这种变换调整也经常要用到。所以在设计中也添加了这种功能，并能够观察调整后的信号的波形图得变化，而且能通过语音处理界面的输出功能试听处理后的语音信号。

4.2.1 语音信号的频率调整

在设计中，可以将语音信号的采样频率提高或降低，来实现语音信号的调整，得到理想的语音信号。例如将采样频率提高一倍，即可得到语音信号频率为原频率 2 倍新的语音信号。运行 Process (Adjust (Frequency，得到如图 4.4 的信号波形图，并试听调整后的效果。

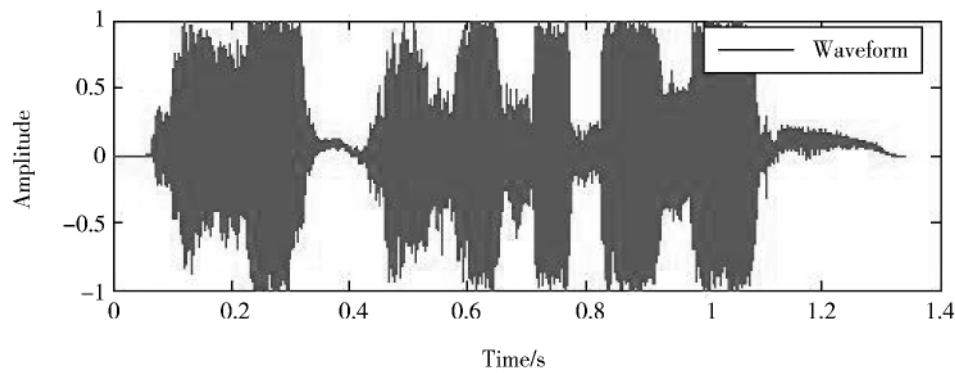


图 4.4 频率调整后波形图

与原语音信号相比，经过调整后的信号周期变为原来的 $1/2$ ，此时的语速明显变快，即实现了信号的 2 倍频功能。

4.2.2 语音信号的振幅调整

在设计中，可以将语音信号的幅度进行提高或降低操作，来实现语音信号的调整，得到声音音量大小不同的语音信号，例如将原语音信号的幅度提高一倍，得到如下图 4.5 的信号波形图，可以通过 GUI 操作界面的输出功能试听调整后的效果。

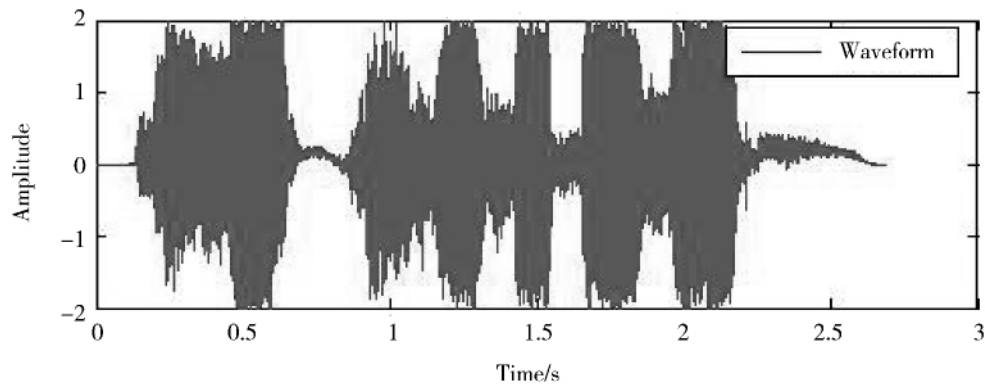


图 4.5 幅度调整后波形图

此时听到的调整后声音声调变高，但不是很明显，可以将幅度的变化值设置的比较大，那样的话就可以得到效果相当明显的语音信号了。

4.3 语音信号的傅里叶变换

倒谱分析是指信号短时振幅谱的对数进行傅里叶反变换。它具有可近似地分离并提取出频谱包络信息和细微结构信息的特点。

对语音信号进行频谱分析，在 Matlab 中可以利用函数 `fft` 对信号行快速傅里叶变换，得到信号的频谱图，并进行倒谱分析，得到倒谱图。

傅里叶变换的部分程序如下：

```
x=y (44101: 55050, 1);           %提取原语音信号的一部分
t= (0: length (x) -1) /fs;       %计算样本时刻
subplot (3, 1, 1);               %确定显示位置
plot (t, x);                     %画波形图
legend (' 波形图');
xlabel (' Time (s)');
ylabel (' Amplitude');
Y=fft (x, hamming (length (x))); %做加窗傅里叶变换
fm=5000 * length (Y) /fs;        %限定频率范围
f= (0: fm) * fs/length (Y);      %确定频率刻度
subplot (3, 1, 2);
plot (f, 20 * log10 (abs (Y (1: length (f))) + eps));
legend (' 频谱图');              %画频谱图
ylabel (' 幅度 (db)');
xlabel (' 频率 (Hz)');
c=fft (log (abs (x) + eps));       %倒频谱计算
ms1=fs/1000;
ms20=fs/50
q= (ms1: ms20) /fs;              %确定倒频刻度
subplot (3, 1, 3);
plot (q, abs (c (ms1: ms20)));    %画倒谱图
legend (' 倒谱图');
xlabel (' 倒频 (s)');
ylabel (' 倒频谱幅度 (Hz)');
```

运行 `Process (Transform`，对语音信号的一部分进行傅里叶变换，并进行倒谱分析，得到如图 4.6

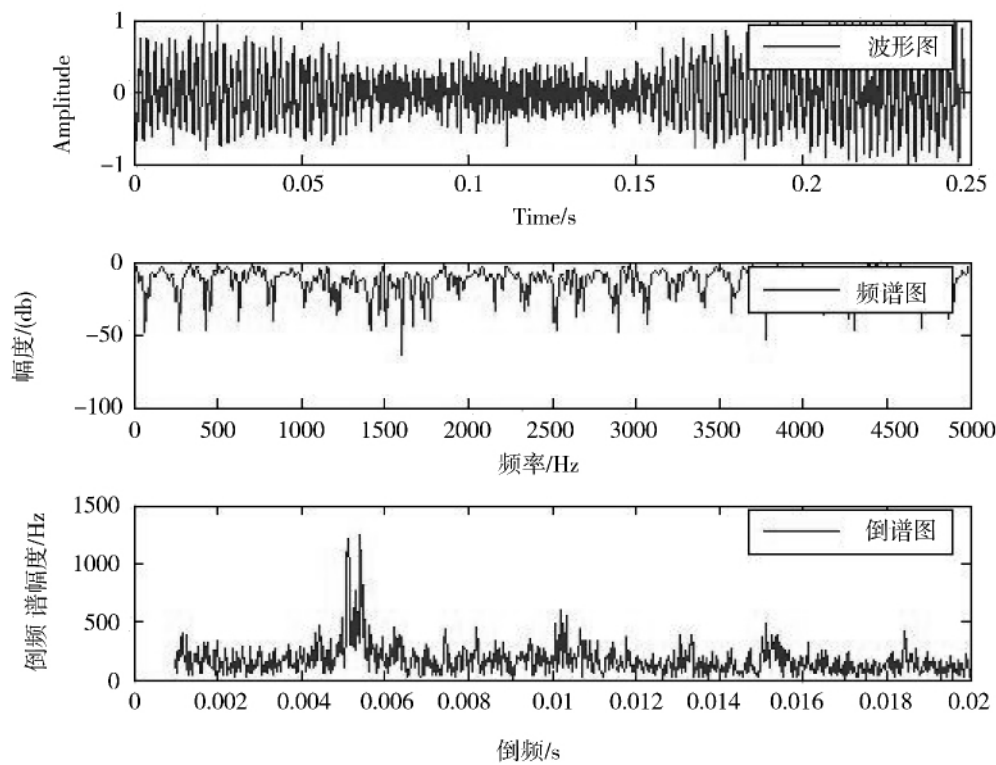


图 4.6 声音样本波形图、频谱图和倒谱图

从上面的倒谱图可以看出，当读“主人，信息收到了”时，所对应的频率大概在 200Hz 左右。这与人的语音信号频率集中在 200 Hz 到 4.5 kHz 之间是相一致的。而在未发声的时间段内，相对的小高频部分（200500Hz）应该属于背景噪声。

4.4 语音信号的滤波

从图 4.4 中发现，语音信号中包含背景噪声，这些噪声的频率一般较高。所以可以利用 MATLAB 软件中的滤波器进行滤波处理，得到较为理想的语音信号。

4.4.1 语音信号的低通滤波

系统中设计了一个截止频率为 200Hz 切比雪夫—I 型低通滤波器，它的幅频特性如下图 4.7：

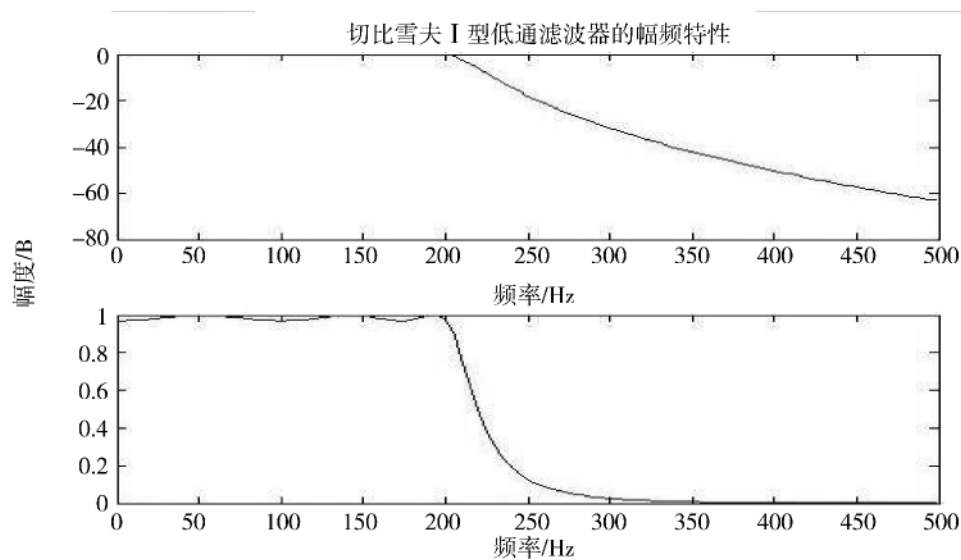


图 4.7 低通滤波器的幅频特性

低通滤波器性能指标： $w_p=0.075\pi$ ， $w_s=0.125\pi$ ， $R_p=0.25$ ； $A_s=50\text{dB}$ ；
经过低通滤波器处理后，比较处理前后的波形图的变化，如下图 4.8：

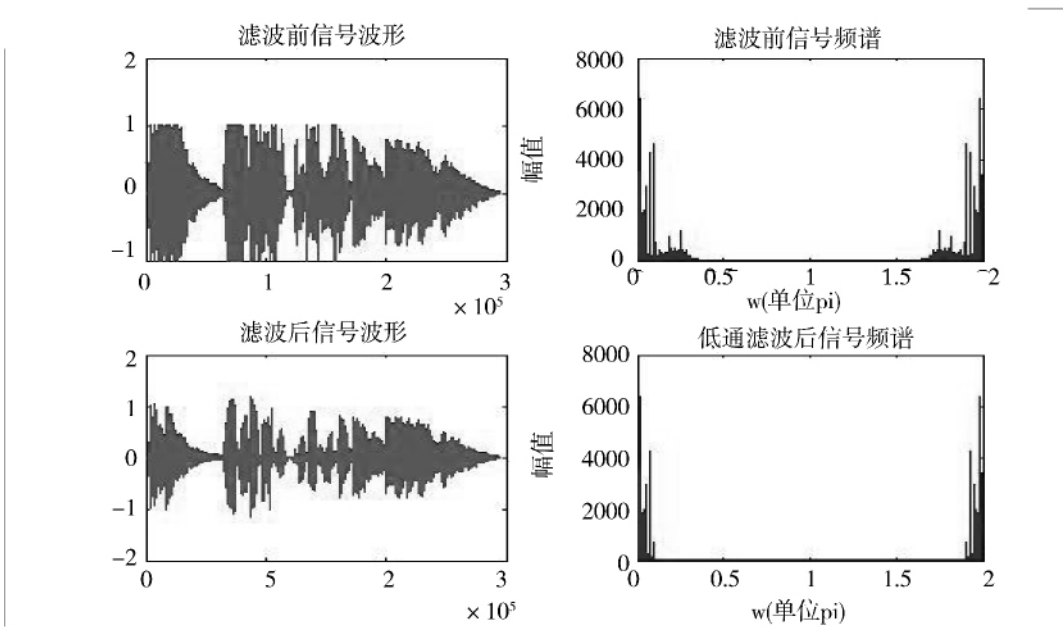


图 4.8 低通滤波后波形和频谱的变化

低通滤波后，声音稍微有些发闷、低沉，原因是高频分量被低通滤波器衰减，但是很接近原来的声音。

4.4.2 语音信号的高通滤波

运用切比雪夫—Ⅱ型数字高通滤波器，对语音信号进行滤波处理。高通滤波器性能指标： $\omega_p=0.375\pi$ ， $\omega_s=0.425\pi$ ， $R_p=0.25$ ； $A_s=50\text{dB}$ ；然后将其与原信号的比较图如下图 4.9：

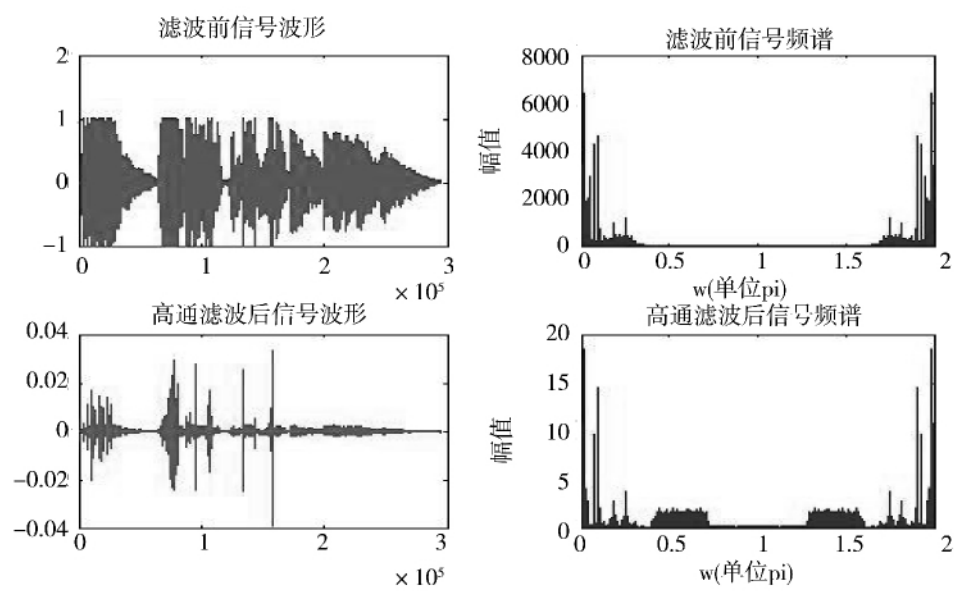


图 4.9 高通滤波后波形和频谱的变化

高通滤波后，此时只有少许杂音，原因是低频分量被高通滤波器衰减，而人声部分正好是低频部分，所以只剩下杂音，或者发出高频杂音但人的耳朵听不到。

4.4.3 语音信号的带通滤波

运用椭圆数字带通滤波器函数，对语音信号进行滤波处理后其与原信号的比较图如下图 4.10：

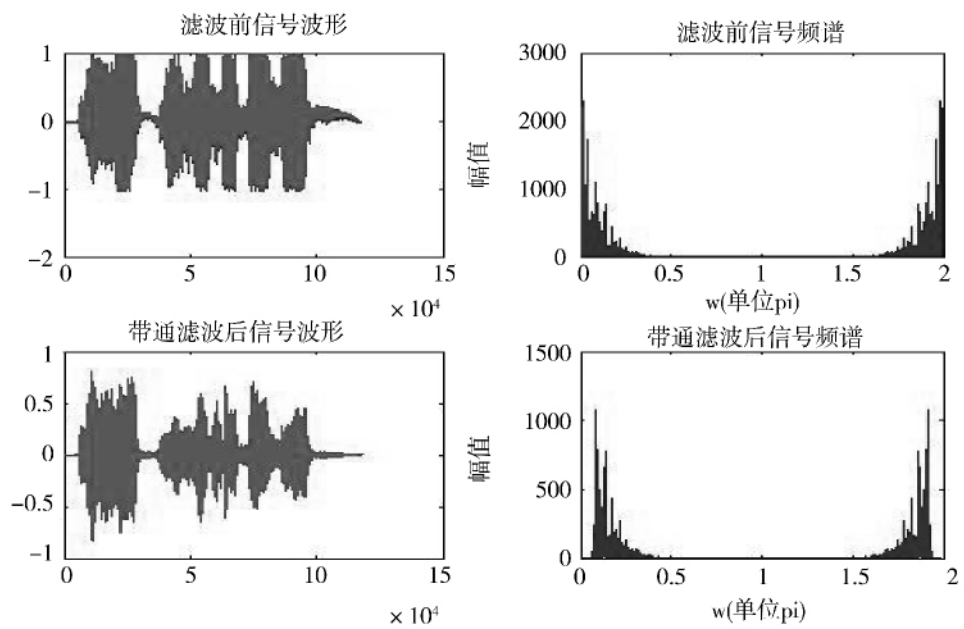


图 4.10 带通滤波后波形和频谱的变化

4.4.4 语音信号的带阻滤波

运用切比雪夫—Ⅱ型数字带阻滤波器，对语音信号进行滤波处理后其与原信号的比较图如下图 3.11：

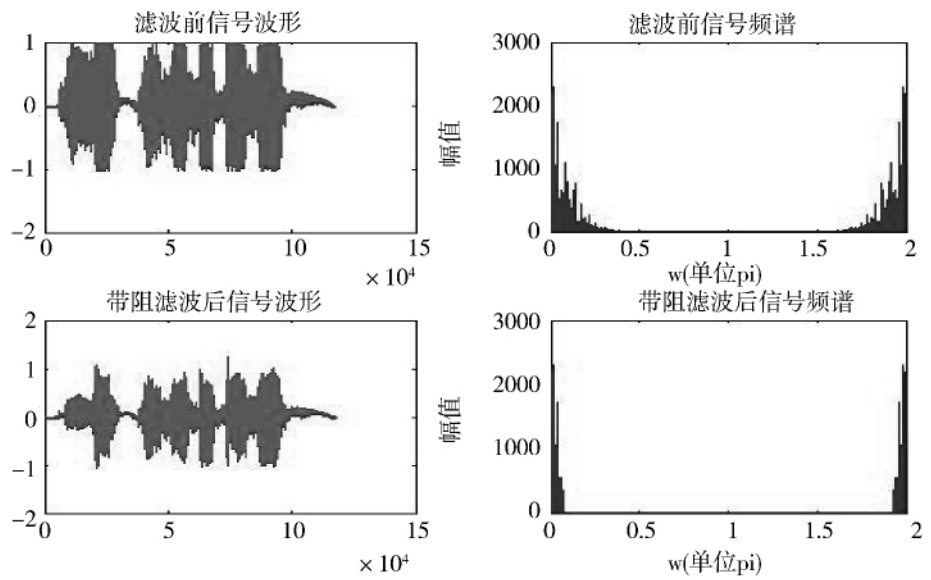


图 4.11 带阻滤波后波形和频谱的变化

从以上各种数字滤波器经过滤波后得出的语音信号相比较，低通滤波后，声音稍微有些发闷，但是很接近原来的声音；高通滤波后听不到人的声音；带通滤波后声音有点像机器人小叮当发出的声音。带阻滤波后，声音比较接近原来的声音。从频谱图中我们可以看出声音的能量主要集中在低频（ 0.2π 即 2204.5Hz 以内）部分。

4.5 语音信号的输出

可以将处理后的语音信号在 Matlab 软件先播放，体验处理后的语音信号的效果。还可以将处理后的语音信号保存在电脑上。

运行 File (Save, 保存处理后的语音信号。如果没有语音信号被处理，则系统会出现提示如下图 4.12:

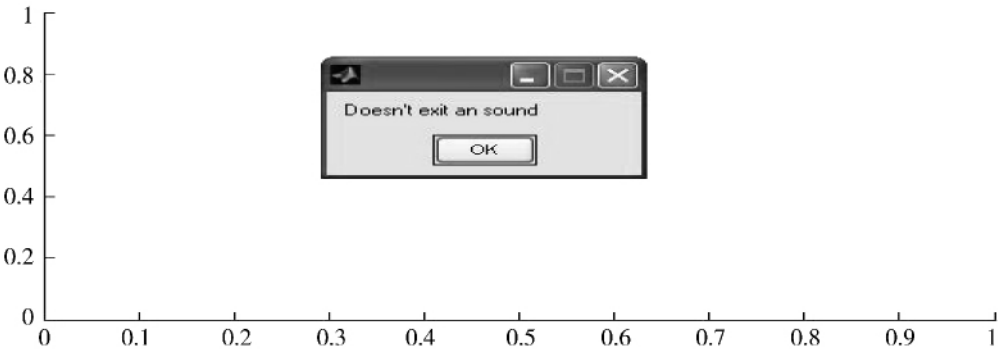


图 4.12 保存提示界面

如果有语音信号被处理，运行 File (Save, 系统会出现提示如下图 4.13:

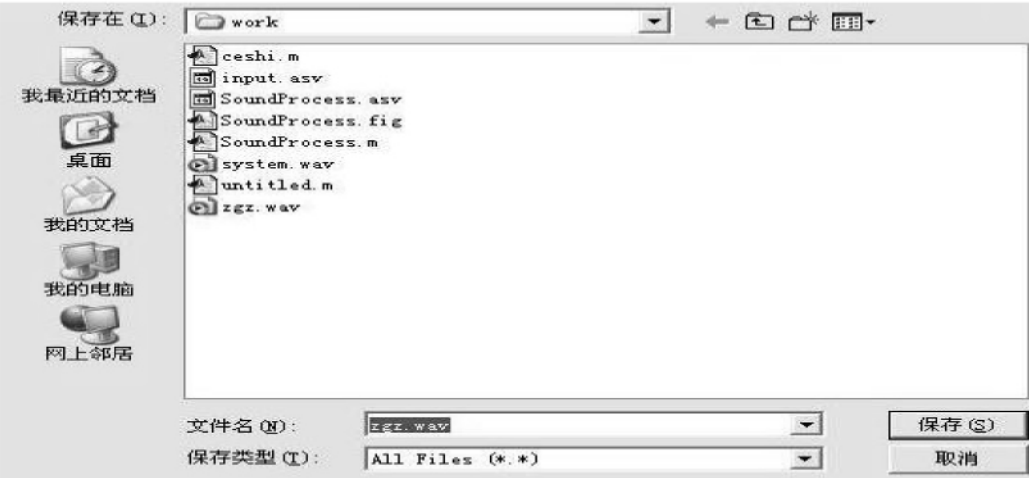


图 4.13 保存界面

保存后，整个操作过程就完成了。

5 结束语

本文对语音信号处理系统的设计作了详细的介绍，采用一系列图像分析和处理技术，实现了语音信号的基本处理的功能，经过测试运行，本设计圆满的完成了对语音信号的读取与打开；较好的完成了对语音信号的频谱分析，通过 fft 变换，得出了语音信号的频谱图；在滤波这一块，课题主要是从数字滤波器入手来设计滤波器，基本实现了滤波，完成了各种滤波器的滤波效果比较，与课题的要求十分相符。

参考文献：

- [1] 王卓交. 基于频率法的缆索承重结构索力分析方法的改进 [D]. 河北：河北工业大学，2009.
- [2] 叶 懋. 电视台前端音频处理器的研究与设计 [D]. 广西：桂林电子科技大学，2008.
- [3] 罗晓琳. 基于 DSP 的电参量数字信号测量算法 [D]. 四川：电子科技大学，2008.

The Design of Processing Voice Signal

Liu Huimin, Wang Long

(College of Physics and Mechanical Engineering, Jishou University, Jishou 416000)

【Abstract】 This article briefly introduced in this paper voice signal acquisition and analysis and the characteristics of speech signal, collection and analysis method of, and after the voice signal acquisition, on the platform of Matlab software for spectral analysis and on the acquisition of the speech signal with the interference noise, to join the noise signal spectrum analysis, design a suitable filter to filter out the noise and restore the original signal. Using Matlab to read (acquisition) of speech signal, will assign it to a vector, the vector as a common signal, the of FFT spectrum analysis, again according to the actual situation of filtering it. Then we can through the sound commands to the speech signal playback, in order to auditory up feeling sound changes.

【Keywords】 Matlab; Speech Signal; Digital Signal Processing; Fourier Transform; Filt

【责任编辑：金 龙】