

# 基于MATLAB 的语音信号分析和处理

徐靖涛 王金根

(解放军炮兵学院, 合肥 230031)

摘 要: 阐述用 Matlab 进行语音信号处理的具体步骤, 从理论和实践上比较分析不同原型下设计的数字滤波器的滤波效果。

关键词: 语音信号处理; Matlab; 滤波器; 频谱分析

中图分类号: TN912

文献标识码: A

文章编号: 1673-1980(2008)01-0132-05

数字信号处理是一门发展迅速、应用广泛的前沿性学科,其理论性和实践性都很强。为了系统地掌握和理解信号的采集、处理、传输、显示和存储,本文通过具体的语音信号处理实例阐述这一原理和方法。

## 1 语音信号概述

### 1.1 基本组成

语音信号的基本组成单位是音素。音素可分成“浊音”和“清音”两大类。如果将不存在语音而只有背景噪声的情况称为“无声”,那么音素可以分成“无声”、“浊音”、“清音”三类。一个音节由元音和辅音构成。元音在音节中占主要部分。所有元音都是浊音。在汉语普通话中,每个音节都是由“辅音-元音”构成的。

### 1.2 语音信号的“短时谱”

对于非平稳信号,它是非周期的,频谱随时间连续变化,因此由傅里叶变换得到的频谱无法获知其在各个时刻的频谱特性。如果利用加窗的方法从语音流中取出其中一个短断,再进行傅里叶变换,就可以得到该语音的短时谱。

浊音的短时谱有两个特点:第一,有明显的周期性起伏结构,这是因为浊音的激励源为周期脉冲气流;第二,频谱中明显地有凸出点,即“共振峰”,它们的出现频率与声道的谐振频率相对应。清音的短时谱则没有这两个特点,它十分类似于一段随机噪声的频谱。

### 1.3 短时分析技术

语音信号具有时变特性,但在一个短时间范围

内(一般认为在 10~30ms 的短时间内),其特性基本保持不变,即相对稳定,因而可以将其看作是一个准稳态过程,即语音信号具有短时平稳性。任何语音信号的分析必须建立在“短时”的基础上,即进行“短时分析”,将语音信号分段来分析其特征参数,其中每一段称为一“帧”,帧长一般取为 10~30ms。这样,对于整体的语音信号来讲,分析出的是由每一帧特征参数组成的特征参数时间序列。

### 1.4 倒谱分析

倒谱分析是指信号短时振幅谱的对数傅里叶反变换。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。

本文按照以下设计思路进行阐述:录制采集语音信号,对录制的信号进行采样;绘制采样后语音信号的时域波形和频谱图,设计滤波器,并运用所设计的滤波器对采集的信号进行滤波,绘制滤波后信号的时域波形和频谱。

## 2 语音信号分析的实验过程

实验按照录制采集语音信号,对录制的信号进行采样,绘制采样后语音信号的时域波形和频谱图,设计滤波器,运用所设计的滤波器对采集的信号进行滤波,绘制滤波后信号的时域波形和频谱的思路进行。

### 2.1 语音信号采集并进行频谱分析

#### 2.1.1 语音信号采集

该实验以研究者本人的声音为分析样本。在 MATLAB 中使用 Wavread 函数,可得出声音的采样频率  $F_s$  为 22 050Hz,且声音是单通道的。利用 sound

收稿日期: 2007-10-12

函数,可清晰地听到读音为:“炮兵学院”。采集数据并画出波形图(图1),其中 $F_s$ 为采样频率, $Y$ 为采样数据,NBITS表示量化阶数。

```
[Y,FS,NBITS]=wavread('xjt.wav');
```

接下来对采样数据作快速傅里叶(fft)变换并画出频谱图和以dB为单位的信号频谱。

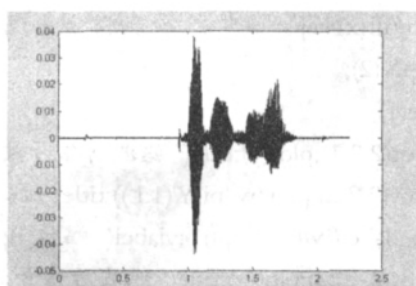


图1 声音样本波形图

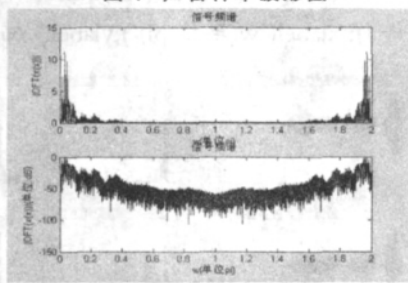


图2 频谱图

由频谱图2可清楚地看到样本声音主要以低频为主。人的语音信号频率一般集中在200 kHz到4.5 kHz之间,从声音频谱的包络来看,样本声音的能量集中在0.1pi(即1 102.5Hz)以内,0.4pi以外的高频部分很少。所以信号宽度近似取为1.1kHz,由采样定理可得 $F_s \geq 2F_0 = 2 \times 1102.5 = 2205\text{Hz}$ ,重放语音后仍可较清晰的听出原声,不存在声音混叠现象。

### 2.1.2 对语音信号进行倒谱分析

```
%获取一个*.wav文件
fn=input('Enter WAV filename: ','s');
[x,fs,nb]=wavread(fn);
ms2=floor(fs*0.002);
ms10=floor(fs*.01);
ms20=floor(fs*0.02);
ms30=floor(fs*0.03);
%画波形图
t=(0:length(x)-1)/fs;
subplot(2,1,1);
plot(t,x);
legend('Waveform');
xlabel('Time (s)');
```

```
ylabel('Amplitude');
%加窗
w=hamming(ms30);
pos=1;
fx=[];
while(pos+ms30)<=length(x)
y=x(pos:pos+ms30-1);y=y-mean(y);
Y=fft(y.*w); %作加窗信号的傅利叶变换
C=fft(log(abs(Y)+eps)); %倒谱分析
[c,fxval]=max(abs(C(ms2:ms20)));
%找出 2ms(=500Hz)和 20m(=50Hz)之间的最大值
fx=[fx fs/(ms2+fxval-1)];
pos=pos+ms10;
end;
%画出倒谱图
t=(0:length(fx)-1)*0.01;
subplot(2,1,2);
plot(t,fx);
legend('FX Trace');
xlabel('Time(s)');
ylabel('Frequency(Hz)');
分析后得出倒谱图(图3)。
```

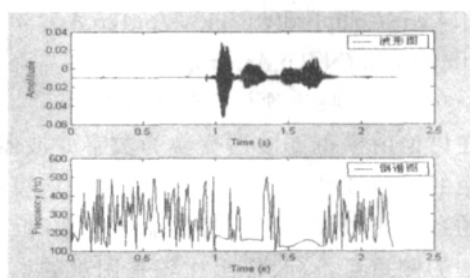


图3 倒谱图

从上面的倒谱图可以看出,当读“炮兵学院”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 kHz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(200~500Hz)应该属于背景噪声。

### 2.2 采用原型低通滤波器设计滤波器

通过频谱分析,要设计各型滤波器,可分别取近似指标如下

(过渡带宽均近似取0.05pi)

低通滤波器性能指标:  $w_p=0.075\pi$ ,  $w_s=0.125\pi$ ,  $R_p=0.25$ ;  $A_s=50\text{dB}$ ;

高通滤波器性能指标:  $w_p=0.375\pi$ ,  $w_s=0.425\pi$ ,  $R_p=0.25$ ;  $A_s=50\text{dB}$ ;

带通滤波器性能指标:  $w_{p1}=0.125\pi, w_{p2}=0.375\pi$ ,  
 $w_{s1}=0.075\pi, w_{s2}=0.425\pi$   $R_p=0.25; A_s=50\text{dB}$ ;

带阻滤波器性能指标:  $w_{p1}=0.075\pi, w_{p2}=0.425\pi$ ,  
 $w_{s1}=0.125\pi, w_{s2}=0.375\pi$   $R_p=0.25; A_s=50\text{dB}$ ;

### 2.2.1 为滤波器设计函数

%数字巴特沃兹低通滤波器函数

```
function [b,a]=digital_lowpass_butter(wp,ws,Rp,As,Fs)
```

```
[N,wn]=buttord(wp/pi,ws/pi,Rp,As)
```

```
[b,a]=butter(N,wn);
```

%数字切比雪夫-1型高通滤波器函数

```
function [b,a]=digital_highpass_chby1(wp,ws,Rp,As,Fs)
```

```
[b,a]=cheb1hpf(wp,ws,Rp,As);
```

%数字切比雪夫-1型带阻滤波器函数

```
function [b,a]=digital_bandstop_chby2(wp,ws,Rp,As,Fs)
```

```
[N,wn]=cheb2ord(wp/pi,ws/pi,Rp,As);
```

```
[b,a]=cheby2(N,Aswn,'stop');
```

%数字椭圆带通滤波器函数

```
function [b,a]=digital_bandpass_ellip(wp,ws,Rp,As,Fs)
```

```
[N,wn]=ellipord(wp/pi,ws/pi,Rp,As);
```

```
[b,a]=ellip(N,Rp,Aswn);
```

### 2.2.2 双线性变换法设计各种滤波器

在讨论由模拟滤波器变换为数字滤波器时, 普遍采用的是双线性变换法, 它保留的是从模拟到数字域的系统函数表示。在滤波器的选取时, 由于设计方法的侧重点不同, 作出比较是困难的。如果在IIR滤波器情况下, 最优的设计将是椭圆滤波器。

我们以双线性变换法设计数字椭圆滤波器为例, 分别给出低通、高通、带通、带阻四种情况下滤波前后信号的波形和频谱进行比较, 见图4、图5、图6、图7。

给出用双线性变换法设计低通滤波器的源程序:

```
clear;clc;
```

```
[y,fs,bits]=wavread('xjt.wav',[22050 44100]); %
```

取第2秒中的数据

```
ws=0.125*pi;wp=0.075*pi;Rp=0.25;As=50;
```

```
[b,a]=digital_lowpass_ellip(wp,ws,Rp,As,fs)
```

```
[Rows,Col]=size(y);
```

```
y(:,1)=y(:,1)-mean(y(:,1));
```

```
Y=abs(fft(y)+eps);
```

%为防止出现0值

```
Y(1)=Y(2);
```

```
L=Rows
```

```
w=0:(L-1);
```

```
w=2*pi/Rows*w;
```

```
x=filter(b,a,y);
```

```
X=abs(fft(x)+eps);
```

%为防止出现0值

```
X(1)=X(2);
```

```
figure(1);
```

```
subplot(2,2,1);plot(y);title('滤波前信号波形');
```

```
subplot(2,2,2);plot(w/pi,Y(1:L));title('滤波前信号频谱');xlabel('w(单位 pi)');ylabel('幅值');
```

```
subplot(2,2,3);plot(x);title('滤波后信号波形');
```

```
subplot(2,2,4);plot(w/pi,X(1:L));title('低通滤波后信号频谱');xlabel('w(单位 pi)');ylabel('幅值');
```

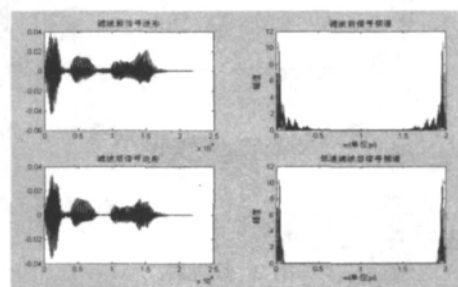


图4 用双线性变换法设计低通滤波器

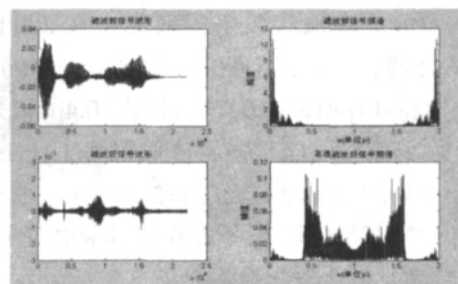


图5 用双线性变换法设计高通滤波器

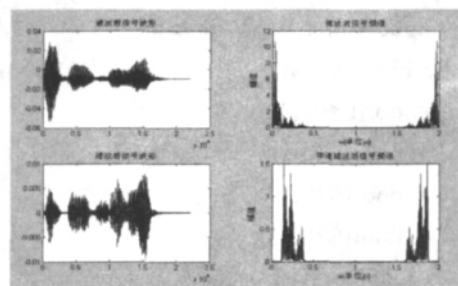


图6 用双线性变换法设计带通滤波器

### 2.3 回放语音信号

在matlab中, 函数sound可以对声音进行回放, 其调用格式为: sound(x,fs,bits)。通过回放可以发现滤波前后的声音有变化。高通滤波后只能听到少量

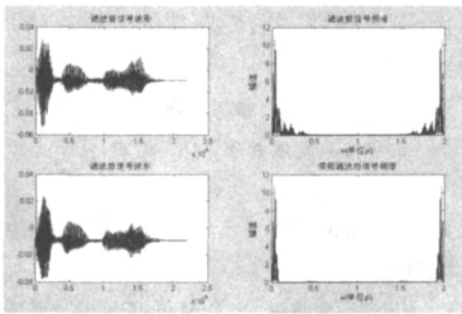


图7 用双线性变换法设计带阻滤波器

的杂音, 低通、带通和带阻均能听到“炮兵学院”四个字。低通滤波后, 声音稍微有些发闷, 但是很接近原来的声音。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( $0.1\pi$  即  $1102.5\text{Hz}$  以内) 部分。

#### 2.4 利用各型滤波器函数滤波并进行比较

```
[Y,FS,NBITS]=wavread('xjt.wav');
[Rows,Cols]=size(Y); %得到数据
Y的维数,Rows为数据量,Cols为通道数
wp=0.075*pi; ws=0.125*pi; %滤波器指标参数
Rp=0.25;As=50;
%在同一张图中画出原始语音信号通过各型滤波器后的频谱图进行比较。
```

(重复以下代码段, 分别改变滤波器函数即可得图8、图9、图10、图11)

Figure(1)

```
[b,a]=digital_lowpass_butter(wp,ws,Rp,As,FS);
%根据指标, 分别利用各型滤波器函数
Y_new=filter(b,a,Y(:,1)); %利用filter
函数对原始信号滤波
sound(Y_new',FS) %再次播放滤波后的声音
Y_new=Y_new';
X1=abs(fft(Y_new(:,1))+floor); %避免
对数值为零,取一很小的值 floor=0.1
L1=Rows
w=0:(L1-1);
w=2*pi/Rows*w;
subplot(2,2,1);
plot(w/pi,X1(1:L)); title(' Butterworth 滤波后语音信号频谱 ');
xlabel(' w(单位 pi)'); ylabel(' 幅值 ');
利用 sound 函数重放滤波后的声音, 图8中低通滤波后的声音跟原始声音差别不大, 通过观察频
```

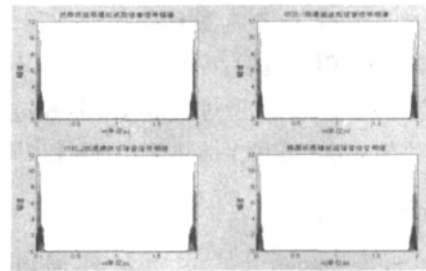


图8 各型低通滤波器滤波后的频谱图比较

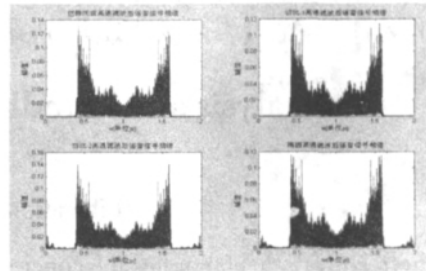


图9 各型高通滤波器滤波后的频谱图比较

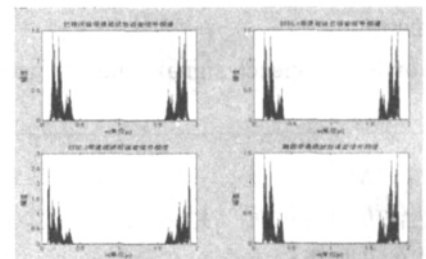


图10 各型带通滤波器滤波后的频谱图比较

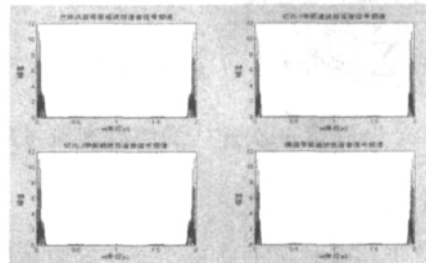


图11 各型带阻滤波器滤波后的频谱图比较

谱以及过渡带的衰减特性, 结合表1中的滤波器阶次可知此时最优的为椭圆型数字低通滤波器。

图9中高通滤波后, 此时只有少许杂音。我们可以看出切比-2型和椭圆型存在频谱泄漏情况, 而巴特沃兹和切比-1型不存在。又因为切比-1型所需要的阶次相对较小, 可知此时最优的为切比-1型数字高通滤波器。

图10中带通滤波后滤波后声音与原声符合最好的是切比-2型带通滤波。经过放大观察频谱, 可明显看到切比-2和椭圆型均有频谱泄漏到低频 $0.1\pi$ 以内, 这也是为什么这两者与原声最接近的原因, 巴特沃兹型在过渡带内波动较大, 而切比-1型过渡带最窄且在过渡带内衰减最快, 且无频谱泄漏



发生, 此时切比-1 型为最优带通滤波器。

图 11 带阻滤波后, 声音与原声差别不大, 放大观察频谱可知巴特沃兹型过渡带较宽且在过渡带内衰减较慢, 比较阶次可知此时椭圆型为最优带阻滤波器。

### 3 结语

通过设计的数字低通、高通、带通和带阻等滤波器对语音信号分别进行滤波处理, 比较不同原型下设计出来的数字滤波器的滤波效果, 从理论上进行

分析比较, 对掌握数字信号处理的分析方法具有良好的指导作用。

#### 参考文献

- [1] 程佩青. 数字信号处理教程[M]. 北京: 清华大学出版社, 2002.
- [2] 刘敏, 魏玲. Matlab 通信仿真与应用[M]. 北京: 国防工业出版社, 2001.
- [3] 楼顺天, 刘小东, 李博菡. 基于 Matlab7.x 的系统分析与设计[J]. 信号处理. 西安: 西安电子科技大学, 2005.

## Analysis and Handle Based on the MATLAB Speech Signal

XU Jing- tao WANG Jin- gen

(Liberation Arm Artillery College, Hefei 230031)

Abstract: The article elaborates the concrete step of the processing which carry on speech signal with the matlab. It analyzes the effect that the different numerical filter descend different result from the view of theory and fulfillment.

Key words: the speech signal handles; matlab; filter; frequency chart analysis

(上接第 131 页)

随机迭代方向变化而产生波动, 使得收敛准则能够满足迭代的总体下降方向, 在最后 m 次迭代的路段客流量的基础上确定收敛准则。收敛准则满足下式:

$$\sqrt{\sum_a (\bar{x}_a^{m+1} - \bar{x}_a^n)^2} / \sum_a \bar{x}_a^n \quad (21)$$

式中,  $\bar{x}_a^n$  为最后 m 次迭代的路段客流量的平均值。

### 3 结语

将拥挤条件下的轨道交通乘客阻抗函数看作随机变量, 建立了随机用户均衡配流(SUE)的理论基础。而 Fisk 轨道交通随机均衡配流模型具有形如 Logit 形式的路径流量解, 可以给出以随机加载模型

方法为基础的用于求解轨道交通 SUE 模型的相继平均算法。

#### 参考文献

- [1] 吴祥云, 刘灿齐. 轨道交通客流量均衡分配模型与算法[J]. 同济大学学报(自然科学版), 2004, 32(9): 159-163.
- [2] Heinz Spiess, Michael Florian. Optimal strategies: A New Assignment Model for Transit Networks [J]. Transportation Research Part B, 1989, 23(2): 83-102.
- [3] Fisk, C. Some developments in equilibrium traffic assignment [J]. Transportation Research B, 1980, 14(3): 243-255.
- [4] Dial, R.B. A probabilistic multipath assignment model which obviates path enumeration [J]. Transportation Research, 1971(5): 83-111.

## Study on Stochastic User Equilibrium Model and Algorithm for Rail Transition

DENG Jun- sheng<sup>1</sup> KONG Fan- yu<sup>2</sup>

(1. China Airport Construction Second Squad, Xuzhou 220115; 2. Tongji University, Shanghai 200331)

Abstract: The article introduces the passenger impedance functions based on congested conditions. It presents a stochastic user equilibrium condition to construct logit-based fisk stochastic user equilibrium model. Stochastic network load method as the base of MSA, an algorithm is given to solve rail transit stochastic user equilibrium model.

Key words: stochastic user equilibrium; rail transit; congested conditions; MSA