

《多媒体技术》第4次作业

姓名 王泽鹏 学号 2017301510036 完成日期: 2019年11月13日

1 实验要求

1. 录一段声音, 内容为“间隔读出数字0-4”, 转为单声道、8khz、wav格式;
2. 每帧帧长20ms160个样点, 使用汉明窗hamming加窗;
3. 设定能量阈值, 判断语音中无声、有声区间;
4. 计算短时自相关 $R(k)$, k 取20-100, 计算基音周期和基音频率;
5. 信息预测: 数字“0”的预测增益 E_0/E_p (原始信号能量/残差信号能量), 并求预测系数 a_1-a_{10} , 阶数 p 取10。

2 思路

先对声音进行窗函数划分帧, 160个样点一帧, 计算各帧短时能量, 设定能量阈值筛选出5段浊音段, 每段分别进行以下操作: 段中的每一帧计算自相关函数(k 为自相关滞后时间, 取20-100), 通过波峰最大值点计算得到该帧浊音的基音频率, 进而计算该段的基音频率, 进而计算5段浊音段的基音频率。

对于信息预测, 对于数字“0”的语音段的某一帧(如第101帧), 利用lpc函数得到阶数取10下的预测系数, 进而计算预测信号及残差信号, 利用预测增益公式 E_0/E_p (原始信号能量/残差信号能量) 计算得到预测增益。

3 实验代码

```
%读取wav文件
[y,fs,wmode,fixd]=readwav('number.wav','p',-1,-1);
figure(1);stem(y,'.');title('原始音频波形');%显示声音的波形
%l=length(y);

nFrames = floor(length(y)/160) ; % 总帧数

%加窗, 计算短时能量
w=window(@hamming,160);
for k = 1:nFrames
    idy = (k-1) * 160+(1:160);
    y_sub = y(idy) .* w;
    E(k) = sum(y_sub.^2);
end
figure(2);stem(E,'.');title('全区间的短时能量图');

%设定能量阈值为0.001得到有声、无声的区间
for i = 1:nFrames
    if E(i)<0.001
        E(i)=0
    end
end
figure(3);stem(E,'.');title('处理后的帧的短时能量图');%处理后的帧的短时能量图
```

```

%计算短时自相关
n=160;
for m=1:length(y)/n %对每一帧求短时自相关函数，每帧的Rm最大值存在N(m)里
    for k=20:100
        Rm(k)=0;
        for i=1:160
            Rm(k)=Rm(k)+y(i+(m-1)*n)*y(i+k+(m-1)*n);
        end
    end
    p=Rm;
    [Rmax,N(m)]=max(p); %读取第一个自相关函数的最大点位置
end

```

%计算基音周期和基音频率

```

T=(1/8)*N %f=8kHz 对应周期为(1/f)*N 对应单位为ms
figure(4);stem(T, '.');axis([0 length(T) 0 20]);
xlabel('帧数(n)');ylabel('周期(ms)');title('初始未处理的基音周期');

```

```

T1= medfilt1(T,7); %去除野点，中值平滑，同时对于能量为0的无声
区间设置基音周期值为0，因为清音的基音周期为0

```

```

for k = 1:nFrames
    if E(k)==0
        T1(k)=0;
    end
end
figure(5);stem(T1, '.');axis([0 length(T1) 0 20]);
xlabel('帧数(n)');ylabel('周期(ms)');title('中值滤波后的基音周期');

```

```

F1= 1000./T1; %计算基音频率，对于能量为0的无声区间设置基
音频率值为0，因为清音没有基音频率，我们设置为0作为浊音段的分隔

```

```

for k = 1:nFrames
    if E(k)==0
        F1(k)=0;
    end
end
figure(6);stem(F1, '.');
xlabel('帧数(n)');ylabel('频率(Hz)');title('基音频率');

```

%计算数字“0”的差分方程系数和预测增益

```

p=10;
w=window(@hamming,160);

```

```

y1=y(100*160:101*160-1); %取数字“0”的一帧（第101帧），进行汉明加窗
A=lpc(y1.*w,p); %得到系数（11个系数中第一个系数值为1）
est_Frame=filter([0 -A(2:end)],1,y1);%预测帧
FFT_est=fft(est_Frame);
err=y1-est_Frame;%计算残差信号
prodictive_gain=y1/err;
figure(7);
subplot(221);plot(y1);grid;title('原始语音段');
subplot(222);plot(est_Frame);grid;title({'预测语音段及10个预测系
数',[num2str(A(2))',' ',num2str(A(3))',' ',num2str(A(4))',' ',num2str(
A(5))',' ',num2str(A(6))',' ',num2str(A(7))',' ',num2str(A(8))',' ',num
2str(A(9))',' ',num2str(A(10))',' ',num2str(A(11))]});
subplot(223);plot(err);grid;title('误差');
subplot(224);plot(prodictive_gain);grid;title('预测增益');

```

4 实验结果

主机配置：CPU: Intel i5-7200U, RAM: 8GB

运行环境：WIN 10

开发环境：MATLAB R2018b

执行时间：2019.11.13

输入文件：number.wav（单声道8kHz）

实验结果图展示：

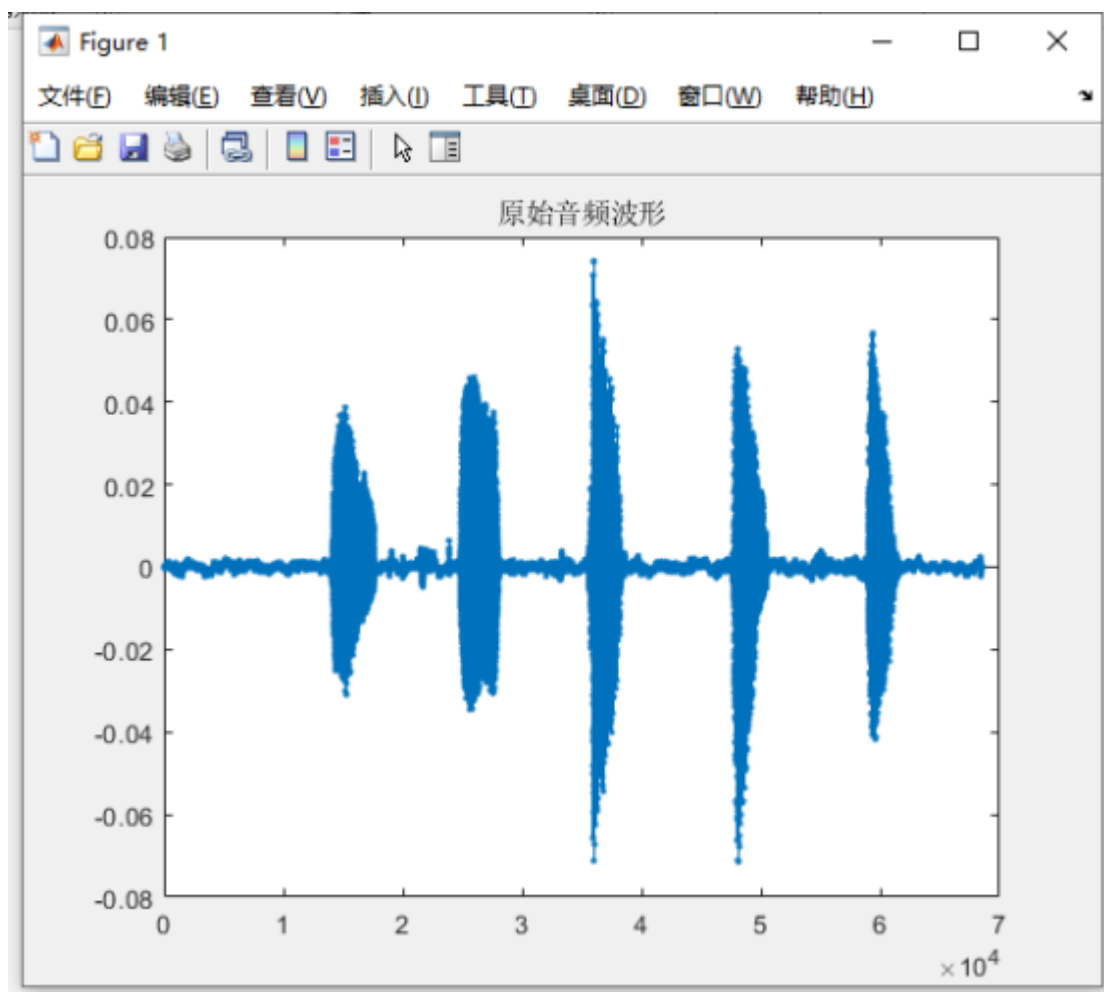


图1 原始信号波形

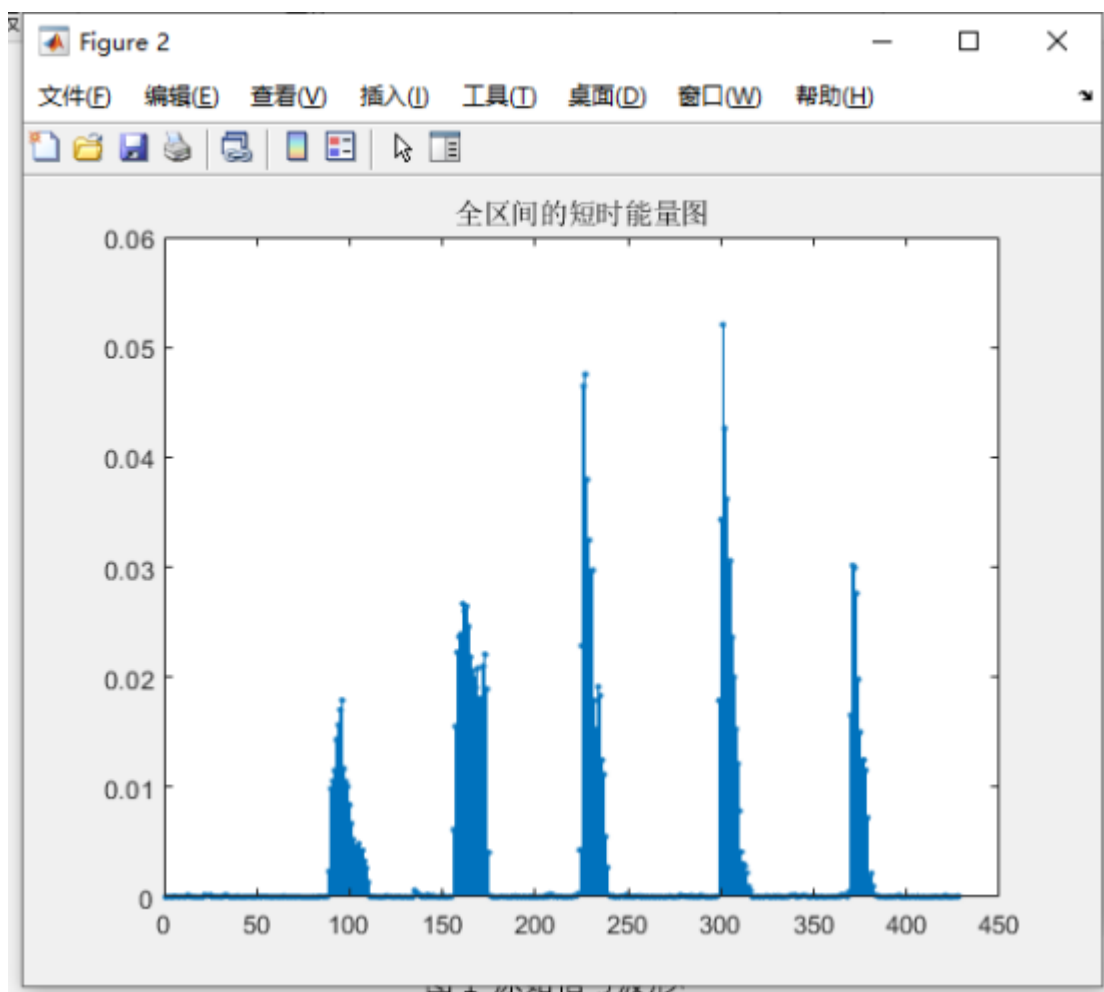


图2 全区间的短时能量图

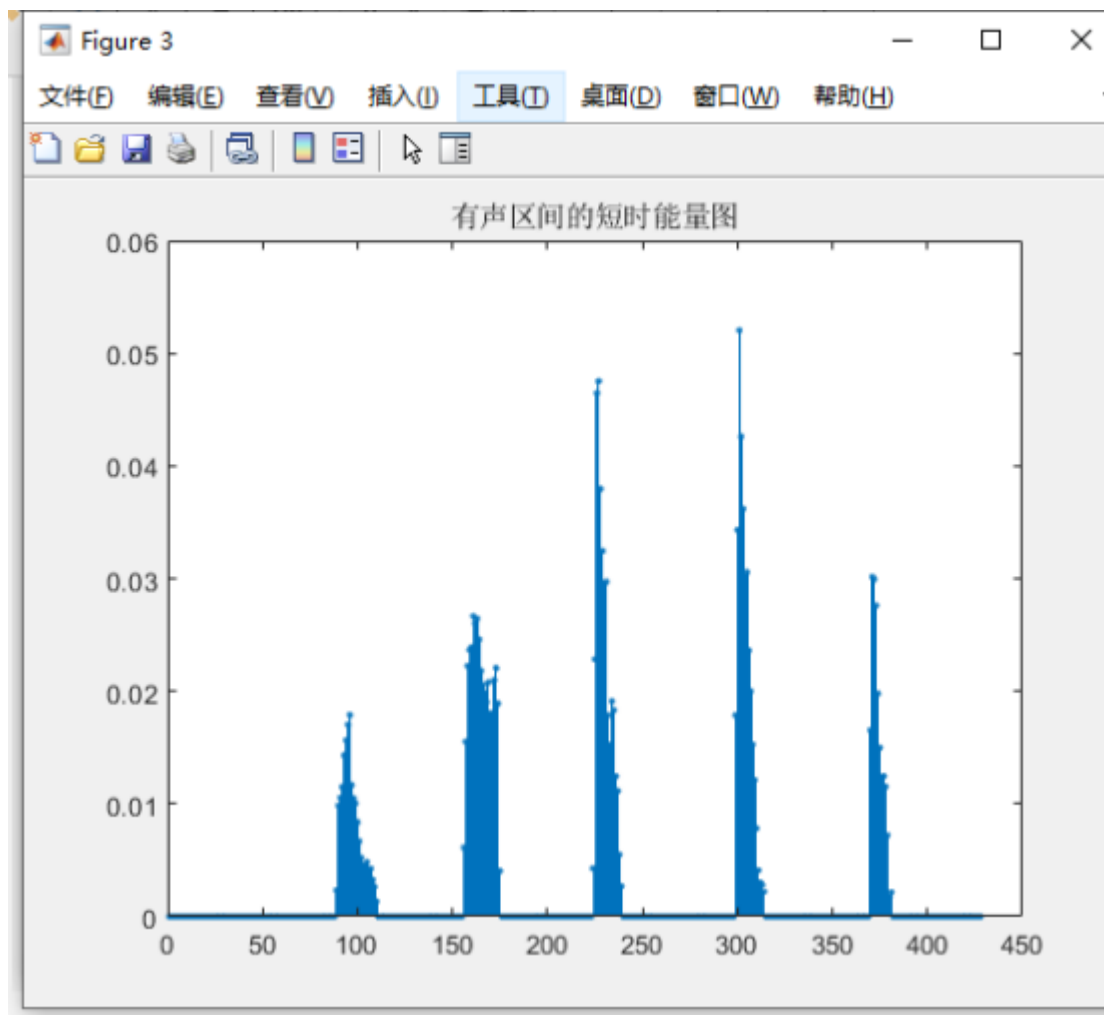


图3 限定阈值处理后的有声区间的短时能量图

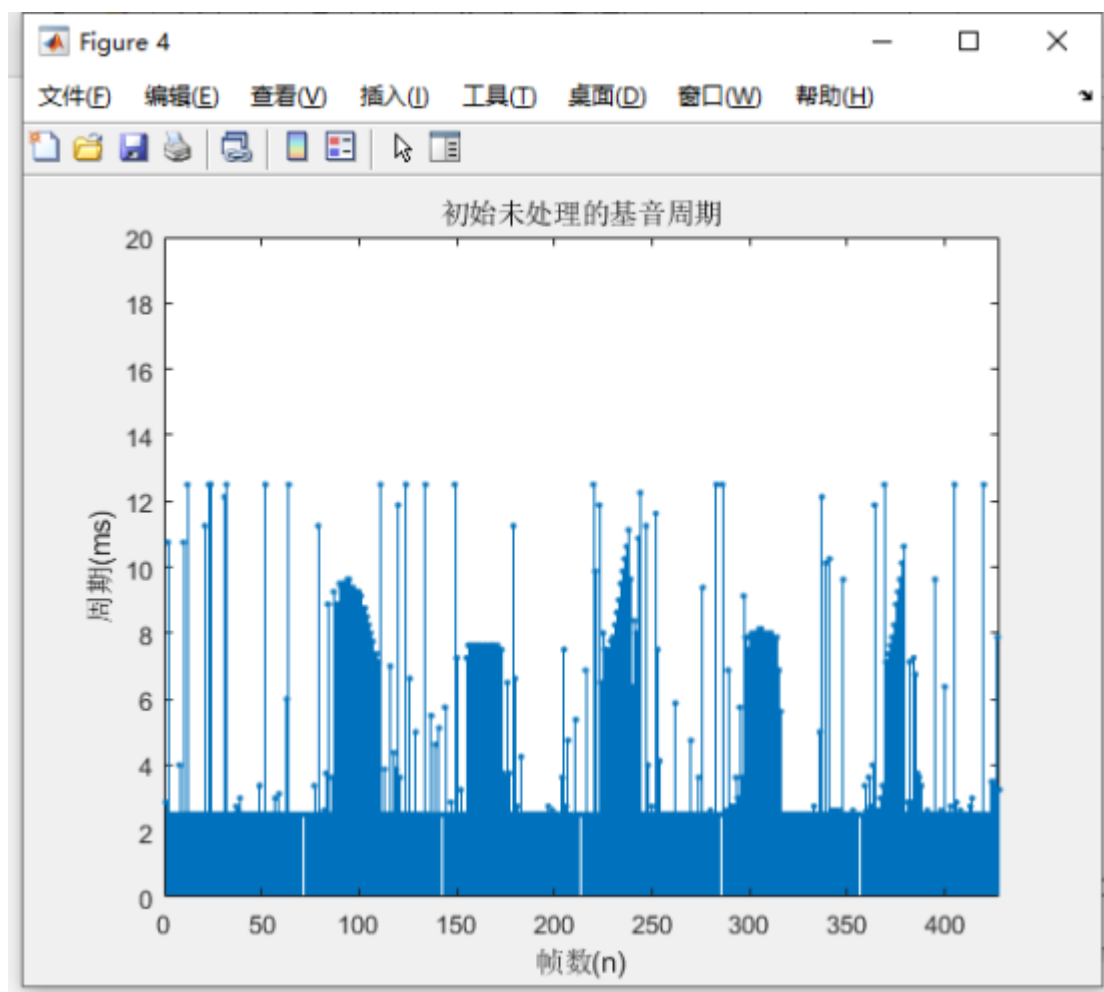


图4 初始未处理的基音周期

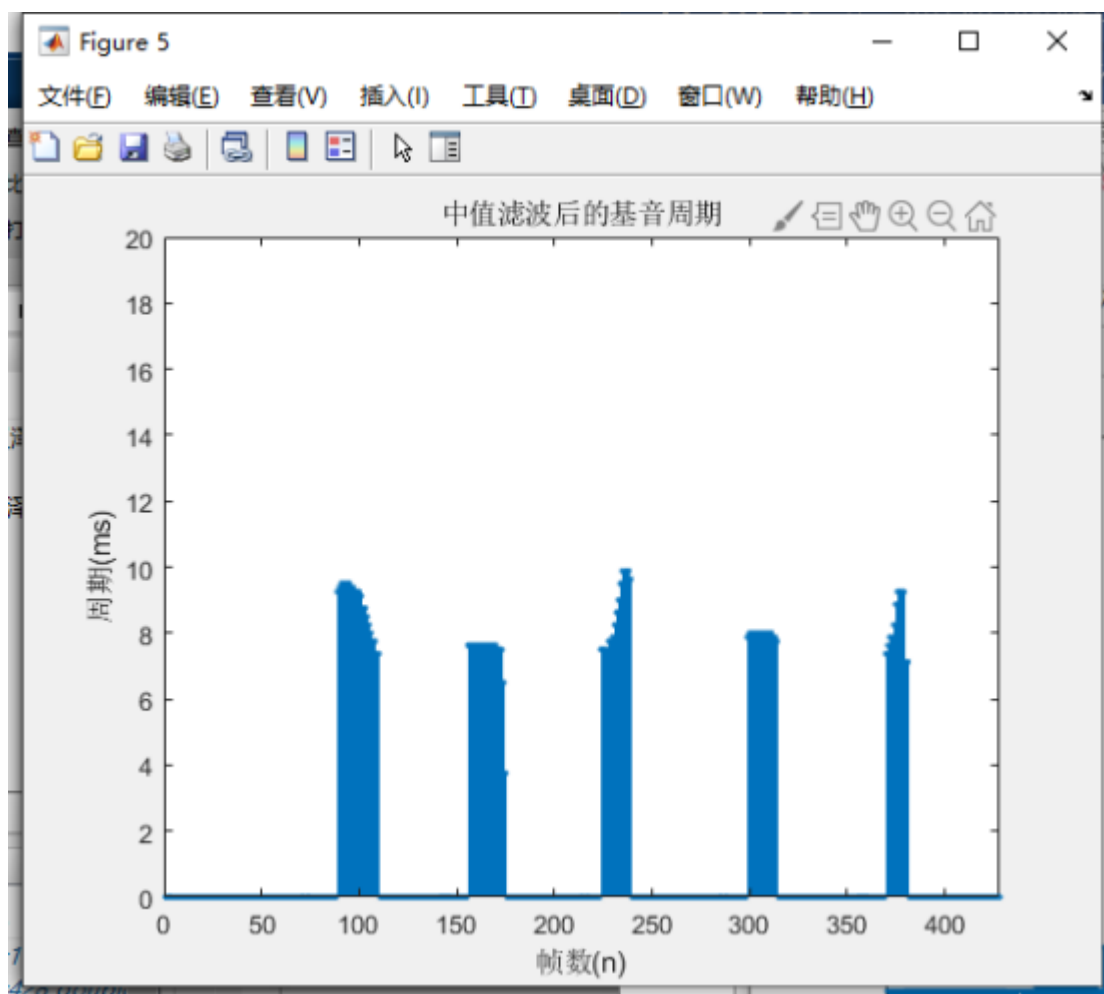


图5 中值滤波后的基音周期

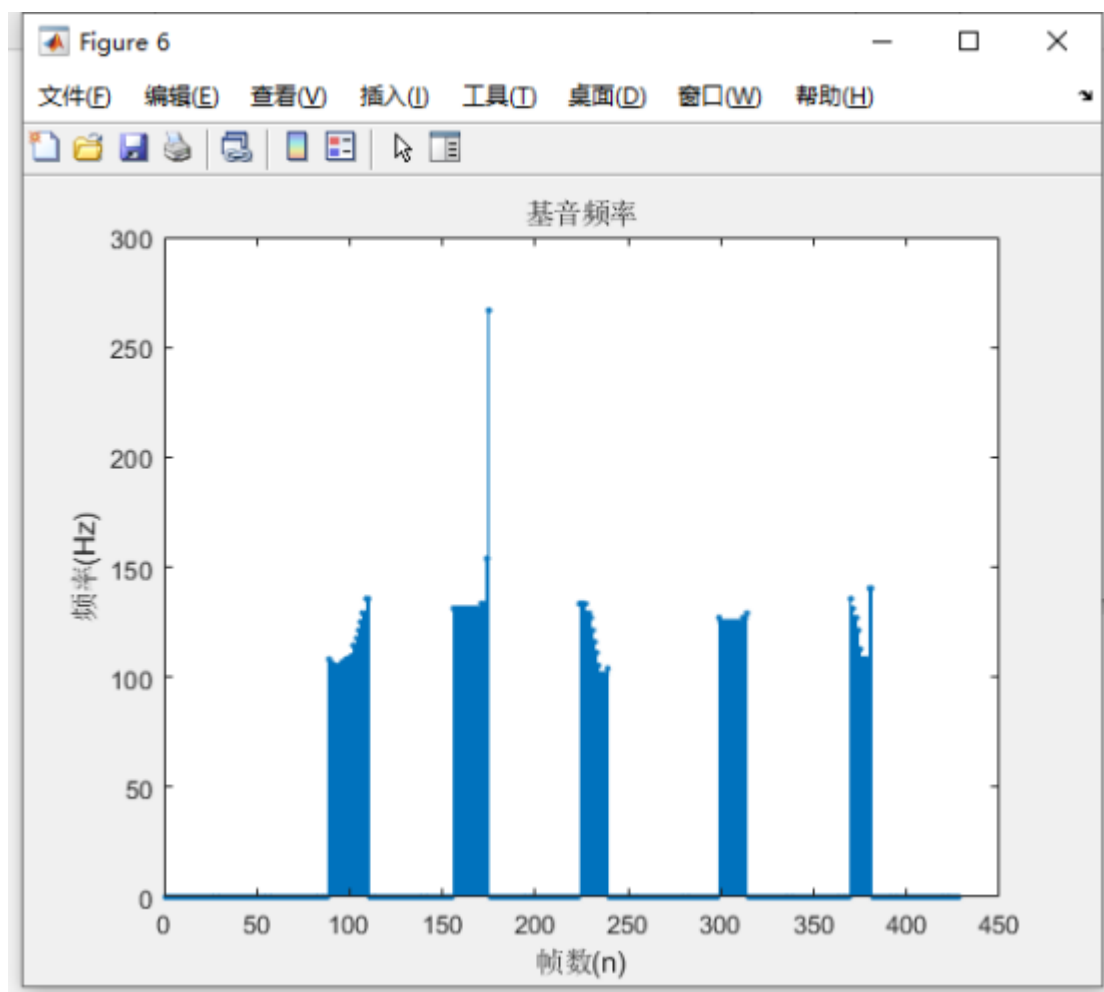


图6 基音频率

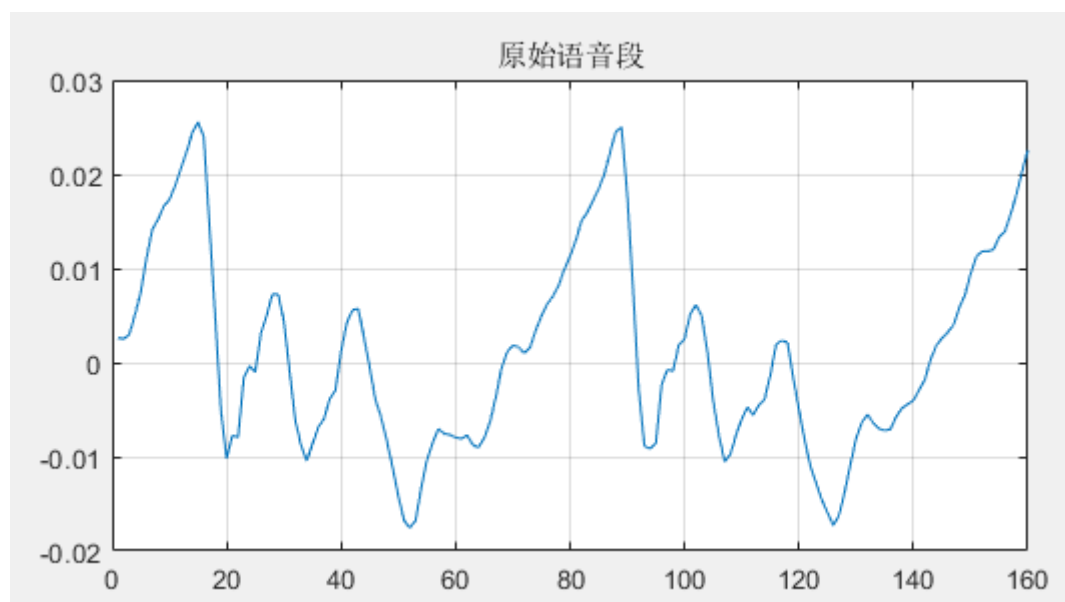


图7 原始语音帧（第101帧）

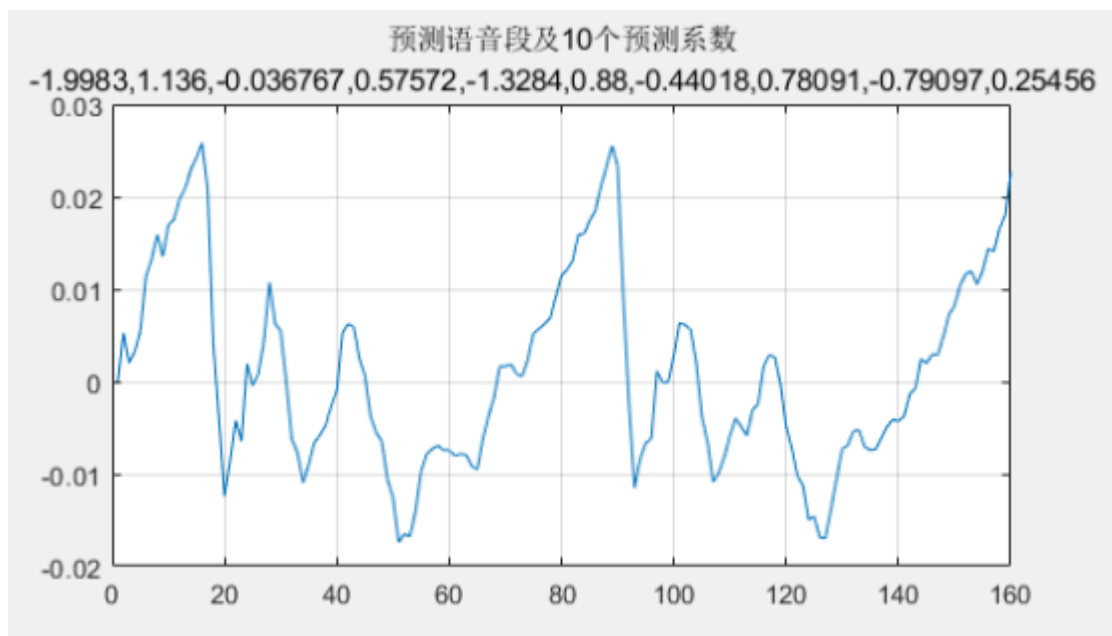


图8 预测语音帧及10个预测系数

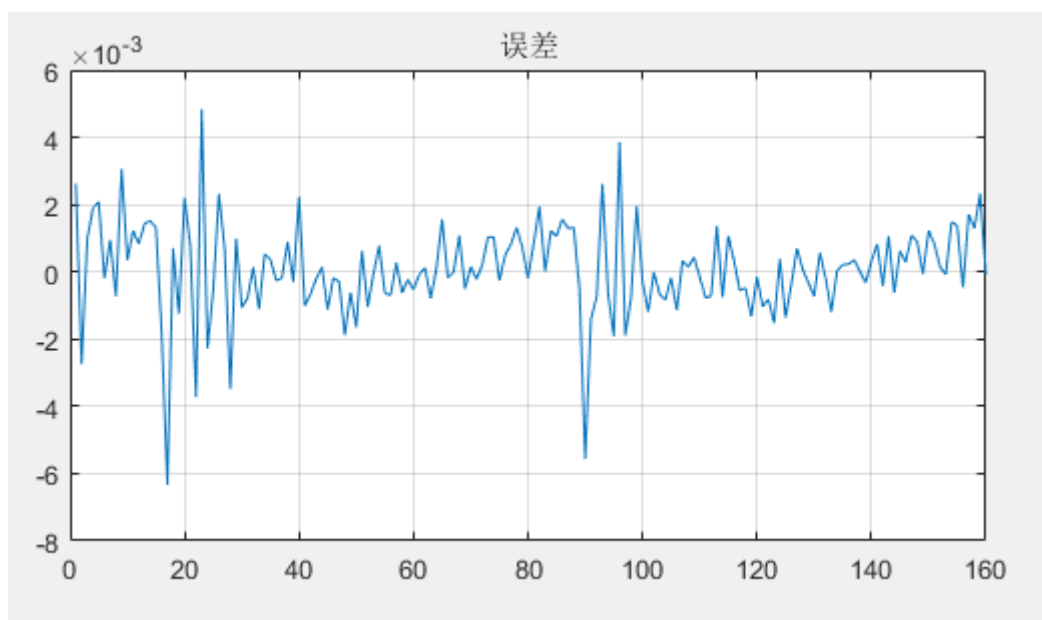


图9 预测语音帧与实际语音帧的误差值

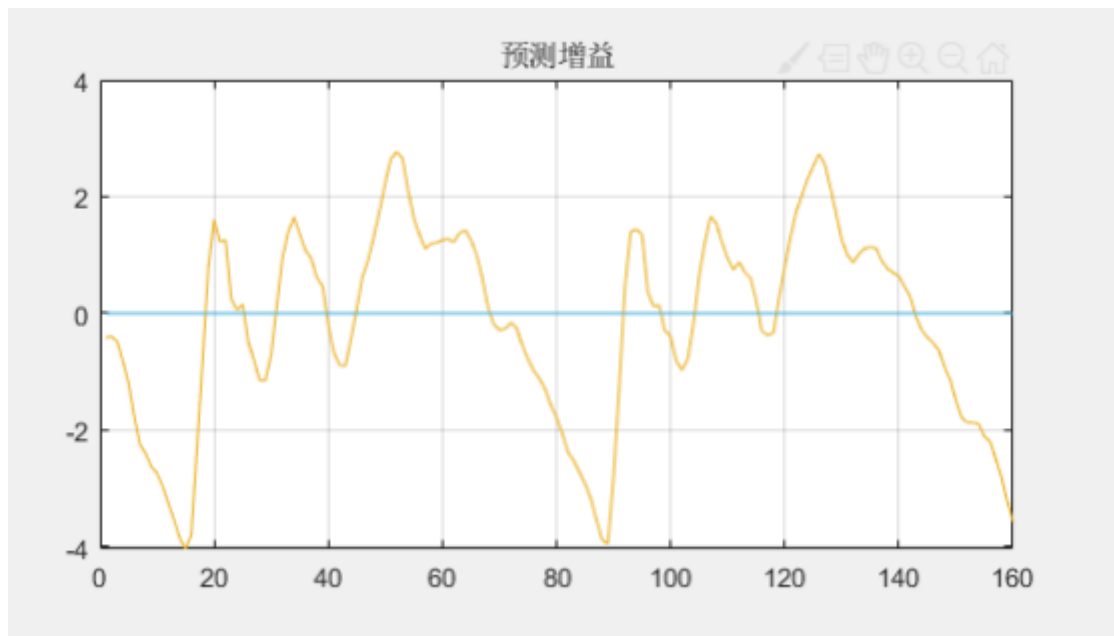


图10 第101帧的预测增益

5 参考文献

[1] 付青青, 吴爱平. 基于Matlab的语音信号自相关基音检测[J]. 长江大学学报(自科版), 2006(10):8+108-110.

[\[2\] 什么是汉明窗? 加Hamming窗的作用?](#)

[\[3\] 语音信号线性预测\(LPC\)分析](#)