

MATLAB 语音合成实验报告

吴同 无 93 2019013217

【原创性声明附在最后】

1.

由 z 变换易得其传递函数为： $\frac{1}{1-a_1z^{-1}-a_2z^{-2}}$.

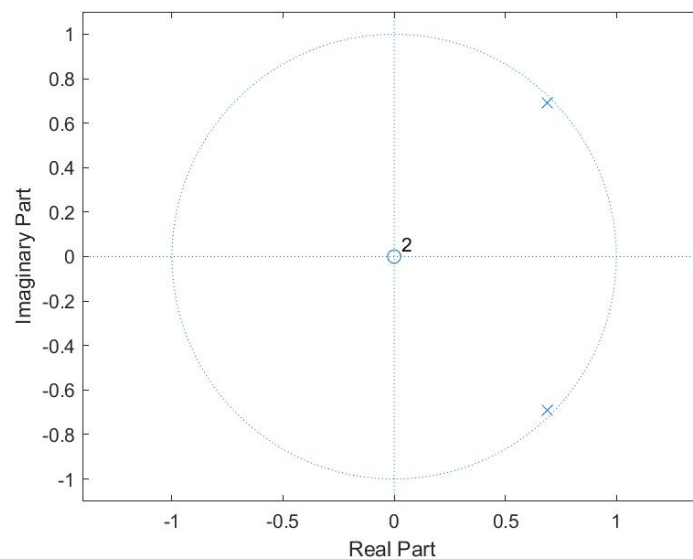
若 $a_1 = 1.3789, a_2 = -0.9506$, 可以得到两个极点如下：

```
>> roots([1, -1.3789, 0.9506])  
  
ans =  
  
0.6895 + 0.6894i  
0.6895 - 0.6894i
```

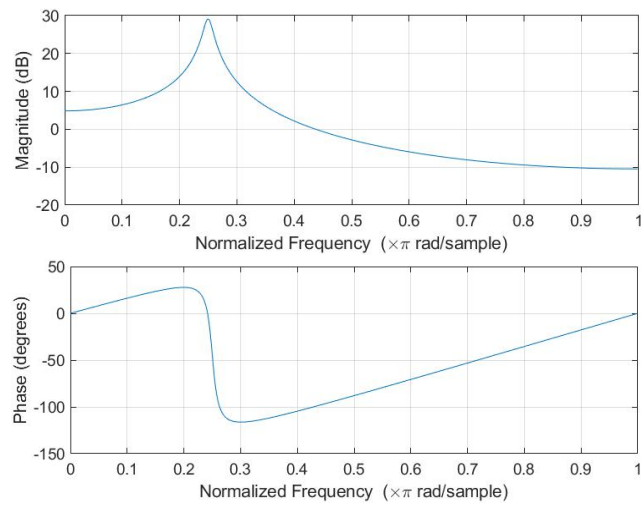
因此共振峰频率为 $\arctan(0.6894/0.6895)/(2\pi T) = \frac{1}{8T} = 1\text{kHz}$.

绘制程序详见附件 VS2_1.m, 结果如下：

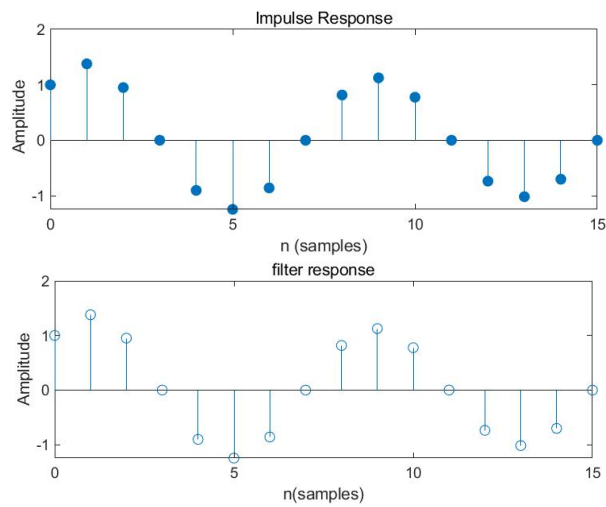
零极点图：



频率响应：



单位样值响应:

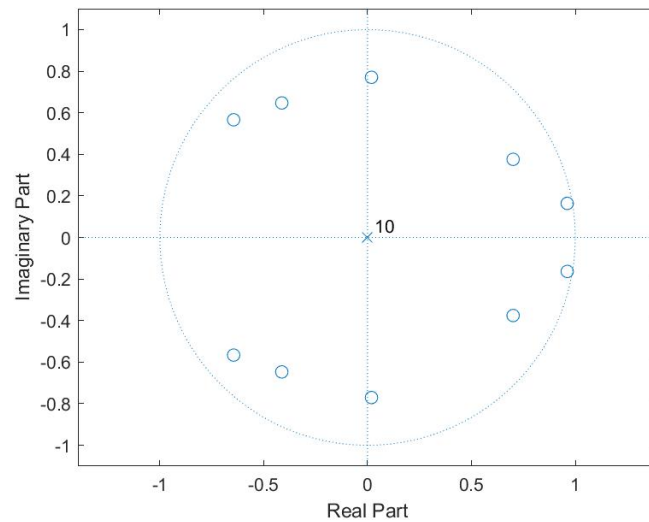


可以看出用 filter 和 impz 绘制的单位样值响应是一致的。

2. 无

3.

程序详见附件 VS2_3.m, 绘制的零极点图如下:



注明：此图为预测系统的零极点图，因此有 10 个零点。

4.

程序详见附件 VS2_4.m.

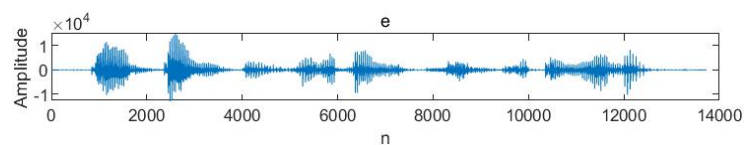
5.

程序详见附件 VS2_5.m.

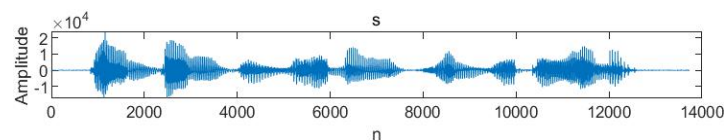
6.

在循环结束后添加了 soundanddraw2_6 函数（程序详见附件 VS2_5.m），用于比较原语音信号、预测激励信号、重建语音信号的听觉效果和实际序列比较。结果如下：（听觉效果请运行程序查看，以下仅为实际序列的比较）

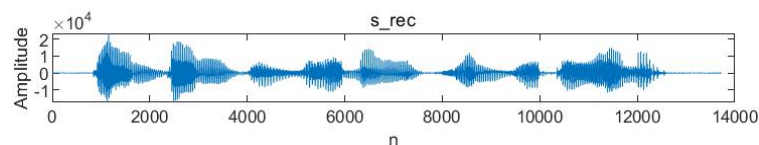
预测激励信号：



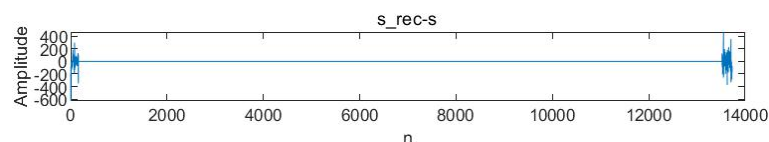
原语音信号：



重建语音信号：



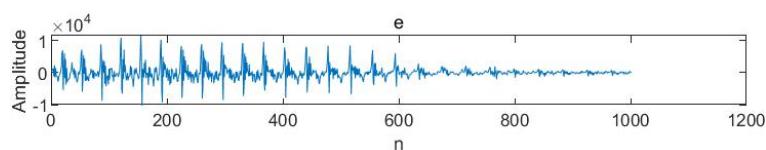
由上可以看出来原语音信号和重建语音信号很相似，而两者和预测激励信号相差较大。因此为了更好比较，另外画出了两者差值的图：



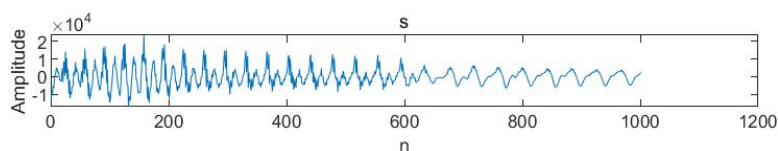
可见原语音信号和重建语音信号主要差别在开始和结束阶段。

此外，从听觉效果上，原语音信号和重建信号完全在我的听觉辨认范围外一致，而预测激励信号则截然不同，相比原语音信号和重建语音信号，它的声调更加尖锐，并且有较大的噪声。为了进一步比较，我截取了 (1000:2000) 的序列片段。结果如下：

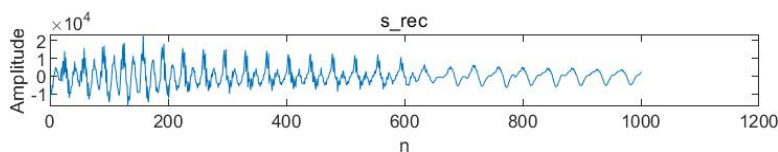
预测激励信号：



原语音信号：



重建语音信号：



可见预测激励信号确实有更多高频分量，出现了很多毛刺。

7.

程序详见附件 VS2_7.m.

在表达式 $x(n) = \sum_{i=0}^{NS-1} \delta(n - iN)$ 中， N 表示数字样值周期，此处应为 40； NS 表示样值数，此处为 200。

运行程序后，可以听到 200Hz 的单位样值串信号的声调比 300Hz 的单位样值串信号的声调低，也即 300Hz 在的单位样值有更强的高频分量。

8.

程序详见 VS2_8.m. (听觉效果运行程序可获得)

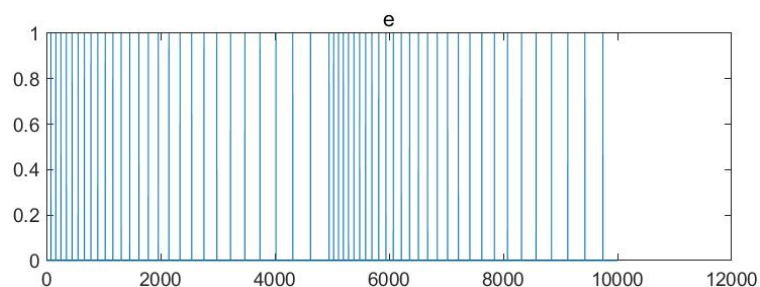
9.

程序详见 VS2_9.m

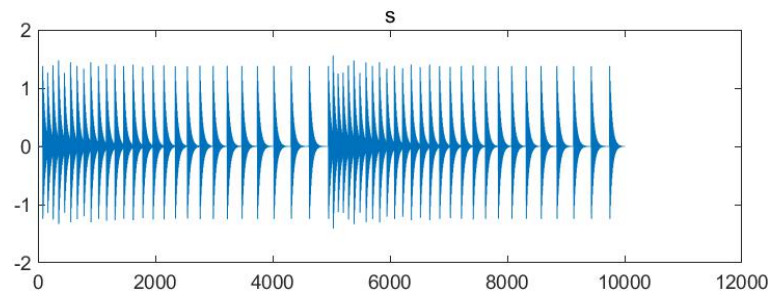
将激励信号和输出信号进行播放对比可得：输出信号相比激励信号声音更加饱满，感觉产生了以下共鸣或者谐振，应该和声道得原理接近。

两者的序列图如下：

激励信号：



输出信号：

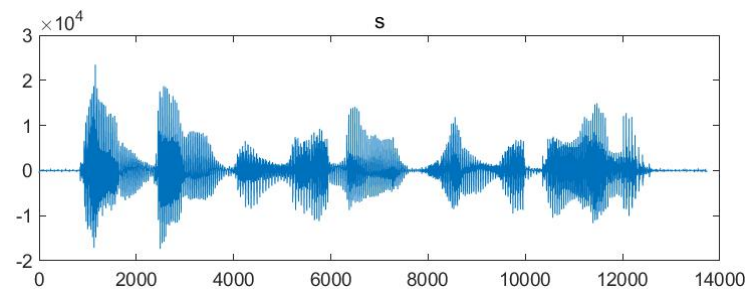


10.

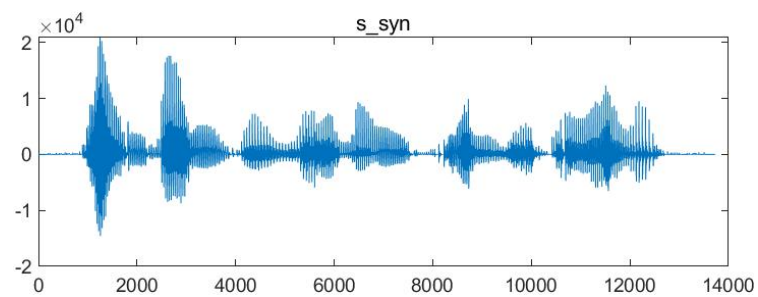
程序详见附件 VS2_10.m.

播放两段音频序列后，重构语音相比合成语音较模糊，原来的噪声变得更加尖锐。两者的序列图如下；

原始语音：



合成语音：

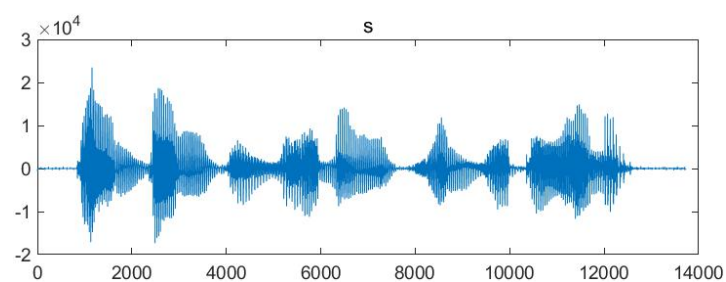


11.

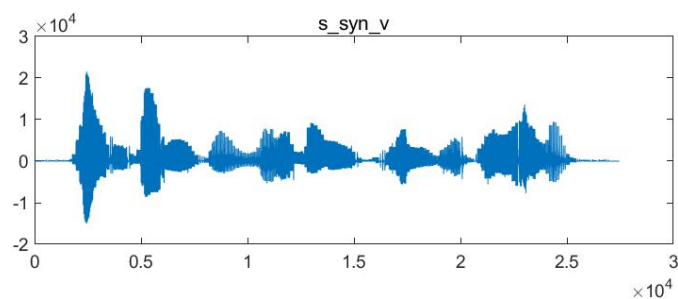
程序详见附件 VS2_11.m.

播放后，可以很容易发现合成的语音信号是原信号的变调不变速。两者的序列图如下：

原始语音：



合成语音：



12.

若将频率提高到 1150Hz,由 $\omega = \frac{\Omega}{T}$ 可知 $\Delta\omega = \frac{\Delta\Omega}{T}$, 因此 $\Delta\Omega = \Delta\omega T = 0.1178$.

因此处于 y 正半平面的极点要乘上 $e^{j\Delta\Omega}$, 处于 y 负半平面的极点要除以 $e^{j\Delta\Omega}$, 可以得到新的极点为：

$p =$

$$0.6036 + 0.7656i$$

$$0.6036 - 0.7656i$$

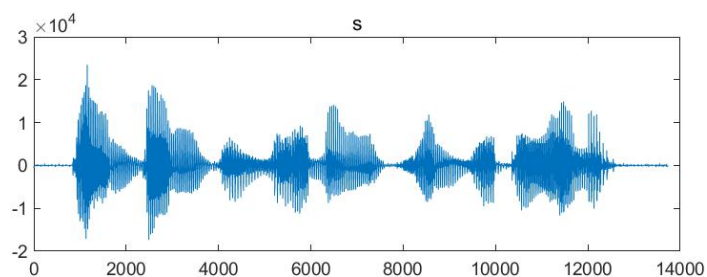
因此新系统的参数 $a_1 = 1.2073, a_2 = -0.9506$

13.

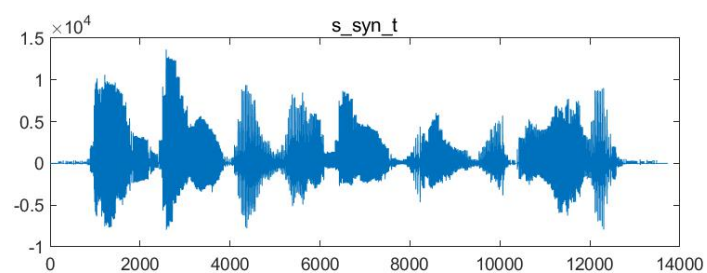
程序详见附件 VS2_13.m.

播放两个音频可以清晰地辨认出合成语音是原始语音地变调不变速。两者地序列图如下：

原始语音：



合成语音：



原创性声明：

以上所有内容均由自己玩成，未参考网络资料，未与同学讨论。