

9.2.1) 语言预测模型

1. 给定:

$$e(n) = s(n) - a_1s(n-1) - a_2s(n-2)$$

传递函数:

$$H(z) = 1 - a_1z^{-1} - a_2z^{-2}$$

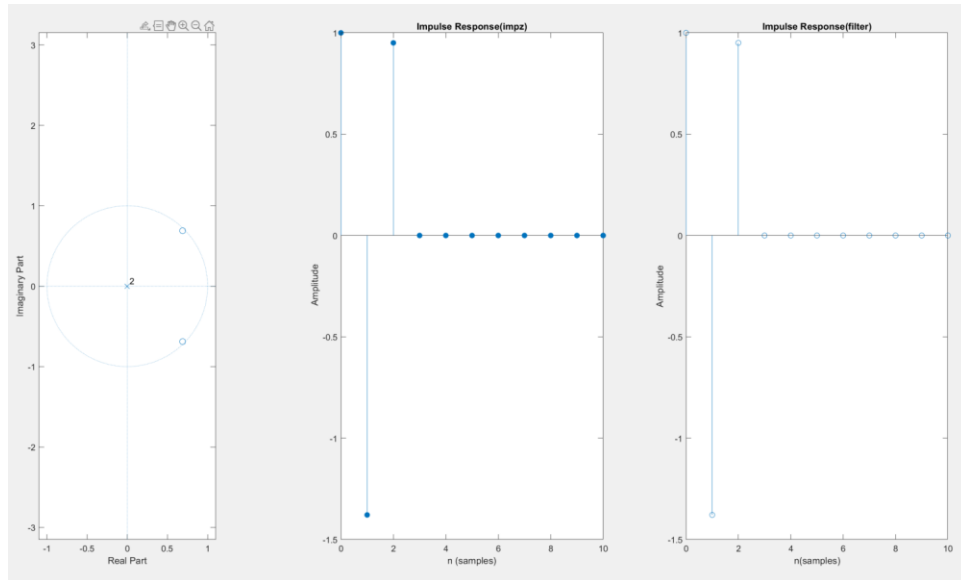
共振峰频率:

$$f = \frac{\Omega}{2\pi T}$$

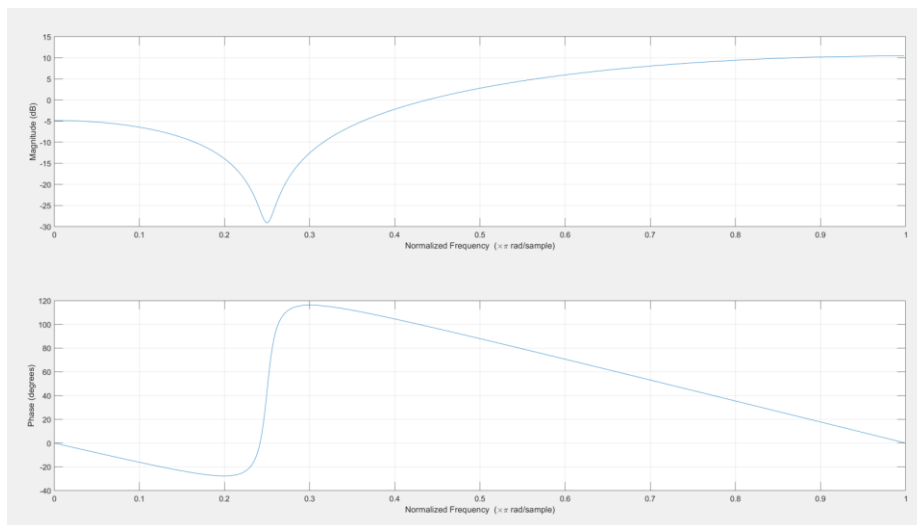
其中, $T = \frac{1}{8000\text{Hz}}$ 为抽样间隔, Ω 为数字频率, 且也是系统极点的幅角。

结果:

- 合成模型共振峰频率: 999.9447Hz (约 1000Hz)



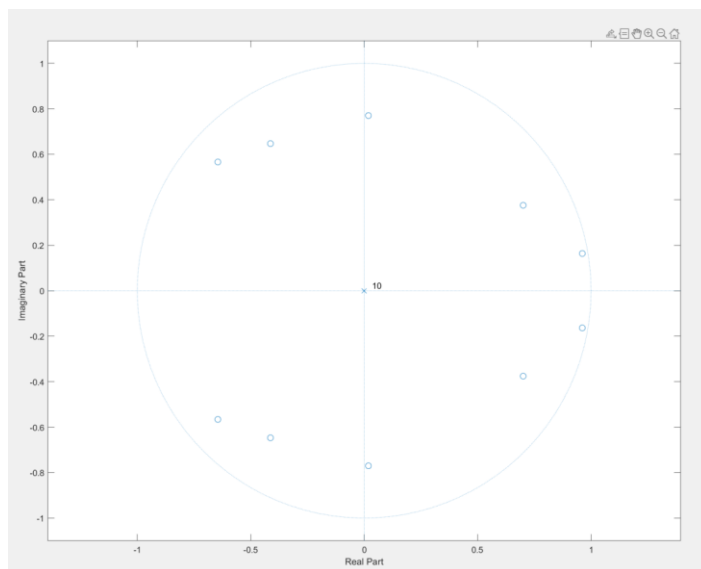
图【1】零极点分布图, 单位样值响应



图【2】频率响应

3.

结果:



图【3】预测模型零极点分布图

4. & 5.

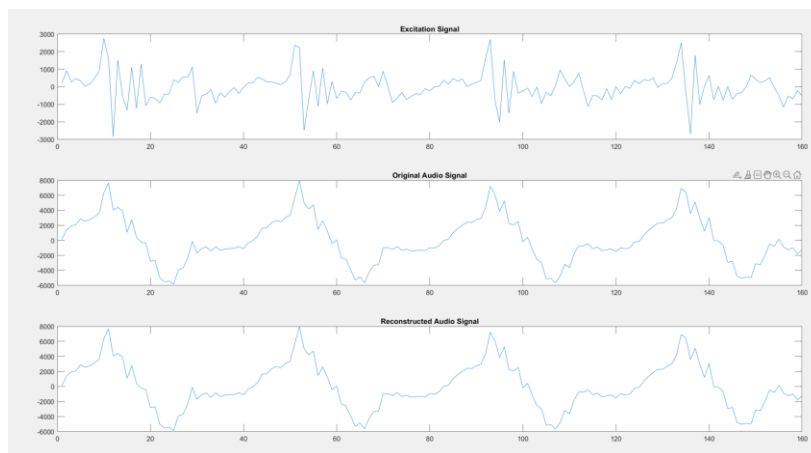
结果:

- 已在 speechproc.m 程序中加上相应代码

6.

结果:

- 激励信号 $e(n)$ 与原来的语音 $s(n)$ 相比，多出了不少噪音，因此较不清晰
- 观察图 4 可以发现，原语音和重建语音与激励信号相比少了高频的分量，且其幅度也增加了许多。另外，前两者之间的差别从图中无法清除看出，试听中两者也似乎没有差别。



图【4】三个信号的比较

9.2.2) 语音合成模型

1. 生成频率为 200Hz 的单位样值串:

$$x(n) = \sum_{i=0}^{NS-1} \delta(n - iN)$$

易知 N 为数字域周期，故:

$$N = \frac{1}{200\text{Hz} \times T} = \frac{8000}{200} = 40$$

在抽样频率为 8kHz 的情况下，1s 的数字信号含有 8000 个样点，因此可以计算出单位样值的数量，表示为 NS :

$$NS = \frac{8000}{40} = 200$$

结果:

- 通过试听，可发现 200Hz 的信号比较“低音”，反之 300Hz 的信号较高音，与预想的结果相符。

2. 给定每个段的基音周期，产生一个激励信号：

$$PT = 80 + 5\text{mod}(m, 50)$$

根据提示，可知 PT 为数字域的周期，即两个单位样值在数字域之间的距离。为满足要求，在程序中使用一个变量记录前一段的最后一个单位脉冲位置，以作为下一段的起始位置。

结果：

- 通过试听，可以听到语音可以分成几段，每一段的幅度都是由小变大。（不太会描述这个语音）

3. 将上一题中制造的激励信号通过 9.2.1.1 的合成模型内。

结果：

- 与激励信号对比，产生的信号似乎与空气穿过一个通道产生的声音类似。（不太会描述这个语音）

4.

结果：

- 合成语音听起来比原始语音清晰，噪音也较少了。

9.2.3) 变速不变调

结果：

- 通过试听，可以发现语音速度慢了，可是音调仍然保持不变。

9.2.4) 变调不变速

1. 提高共振峰频率会更改系统极点。下面推导计算新极点的方法。

$$f_{new} = \frac{\Omega_{new}}{2\pi T}$$

$$2\pi T(f + \Delta f) = \Omega_{new}$$

$$\Omega_{new} = \Omega + 2\pi T \Delta f$$

已知极点的幅度幅角形式，提高共振峰频率不会影响幅度，但需要改变幅角（逆时针转），因此可利用指数性质：

$$\therefore p_{new} = p_{old} \times e^{i2\pi T \Delta f}$$

需要注意的是，针对幅角为负的极点，需要将 Δf 变号（顺时针转）

结果：

- 共振峰频率提高后，新的 **a1** 与 **a2** 分别为 **1.2073** 和 **0.9506**

2.

结果：

- 通过试听，可以发现语音音调明显变高了，语音速度仍然保持不变。

总结与感想

这次的大作业让我对离散系统的实际应用有一定的了解，也让我充分感受到数字信号处理的强大之处，更重要的是其也激发了我对信号处理的兴趣。犹记得在学习信号与系统时，由于涉及较多理论知识和数学计算，所以我觉得有些枯燥。学习了 **Matlab** 后，毫无疑问可以将复杂的计算进行简化，让我能花更多心思在系统特性上。虽然此次作业中，我对语音相关知识的了解依旧有限，但是其让我明白了对语音进行各种有趣处理的原理，这点大大地加深了我对后续课程如数字信号处理和视听信息系统导论等专业课程的兴趣。

综上所述，我认为学习使用 **Matlab** 对后续的课程学习（尤其是需要自主设计系统等工程问题的时候）有很大的帮助，也非常感谢老师和助教们在暑假期间的付出。