

02

第2章 基础知识



人类听觉特点



语音信号处理基础



人类视觉特点



图像信号处理基础



人类听觉特点



人类听觉特点



语音的产生及数字模型



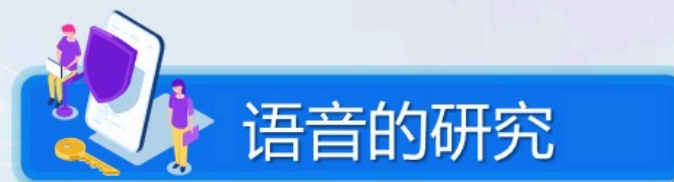
听觉系统和语音感知



语音的质量评价



语音的产生及数字模型



语音的研究



人类对于语音的研究包括两个方面:

从语音的产生和语音
的感知方面来研究

从信号处理的角度来研究



人类对语言的认识

语言的产生

决定想传达给对方的内容是什么
将内容转化成语言的形式（词和语法）

语言学

研究语音中各个音的排列规则及其含义，包括重音、语调、声调等（各国语言，各民族语言，地方方言等）

语音学

研究语音信号是由哪些最基本的单位组成的，发声器官是如何发出声音的，建立语音产生模型，便于人类对语音信号的特性进行研究

单选题 1分

() 研究语音中各个音的排列规则及其含义，包括重音、语调、声调等（各国语言，各民族语言，地方方言等）

A 语言学

B 语音学

语音的产生



声门 (激励源)



产生能量



产生声带的振动



声道 (信道)



口腔



鼻腔



嘴



人能够发出各种不同的声音，取决于声门和声道中各种可变的部分



声带

振动

产生浊音（元音和浊辅音）


不振动

产生清音（清辅音）



声道

是一个具有某种谐振特性的腔体，且其横截面积是可变的
声道的频率特性主要取决于声道截面的最小值出现的位置，主要由舌的位置来控制
嘴是声道的最末端，嘴的形状也影响所发出的声音



用数字模型表示语音信号



→ 用数字模型分别模拟

激励源

声道模型

辐射模型



→ 语音是时变信号，但可以认为在10~20ms的时间范围内是近似不变的



→ 语音的数字模型是一个缓慢时变的线性系统，在10~20ms的时间内是时不变系统

1

激励源模型

激励源

浊音:

激励信号由一个周期脉冲发生器产生，其周期称为基音周期。为了使浊音的激励信号具有声门气流脉冲的实际波形，还需要使这一脉冲序列通过一个声门脉冲模型滤波器，其传输函数为 $G(Z)$

浊音

清音

清音:

激励信号由一个随机噪声发生器产生

浊音

1

激励源模型

声门脉冲滤波器

$$g(n) = \begin{cases} \frac{1}{2} \left[1 - \cos \frac{\pi n}{N_1} \right] & (0 \leq n \leq N_1) \\ \cos \left(\frac{\pi(n - N_1)}{2N_2} \right) & (N_1 \leq n \leq N_1 + N_2) \\ 0 & (\text{其它}) \end{cases}$$

浊音的产生模型

周期
信号

通过一个声门脉冲滤波器，经过幅度调节，产生浊音信号

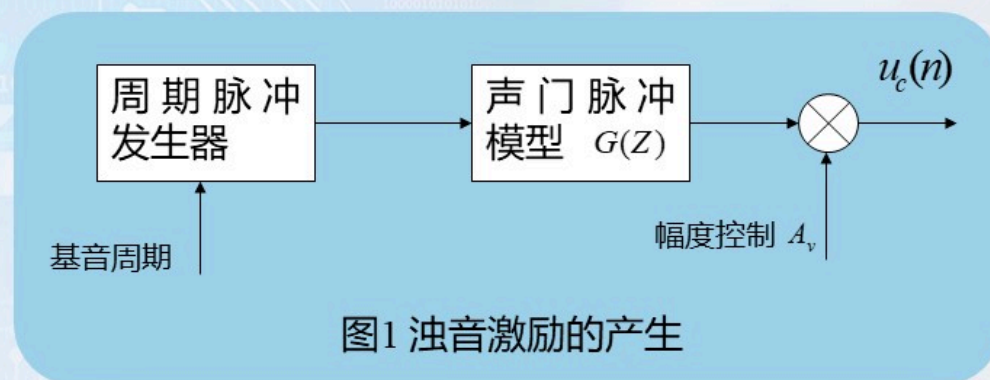


图1 浊音激励的产生

清音

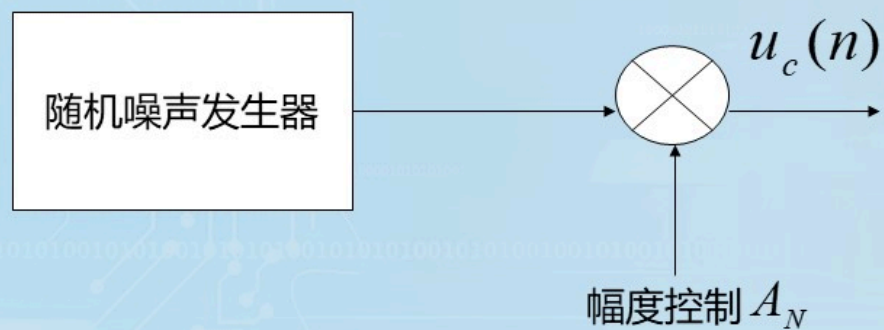


图2 清音激励的产生

2. 声道模型

声道可以用一个全极点模型来模拟，其系统函数的极点对应语音的共振峰

$$V(z) = \frac{G}{1 - \sum_{m=1}^N \alpha_m z^{-m}}$$

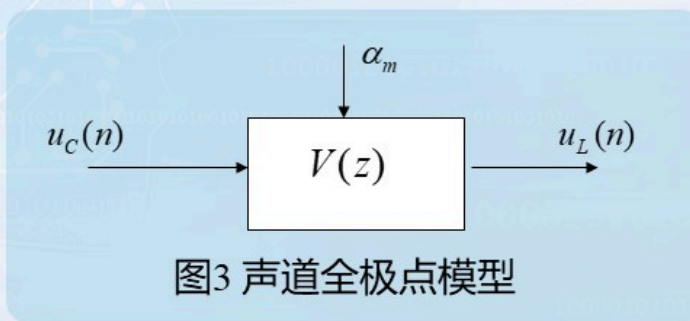


图3 声道全极点模型

3. 嘴唇辐射模型

嘴唇辐射模型与嘴型有关，其传输函数：

$$R(z) = R_0(1 - z^{-1})$$

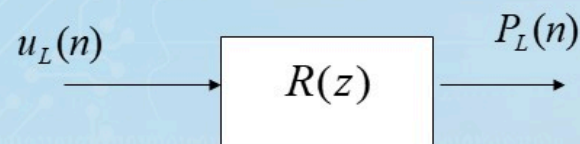
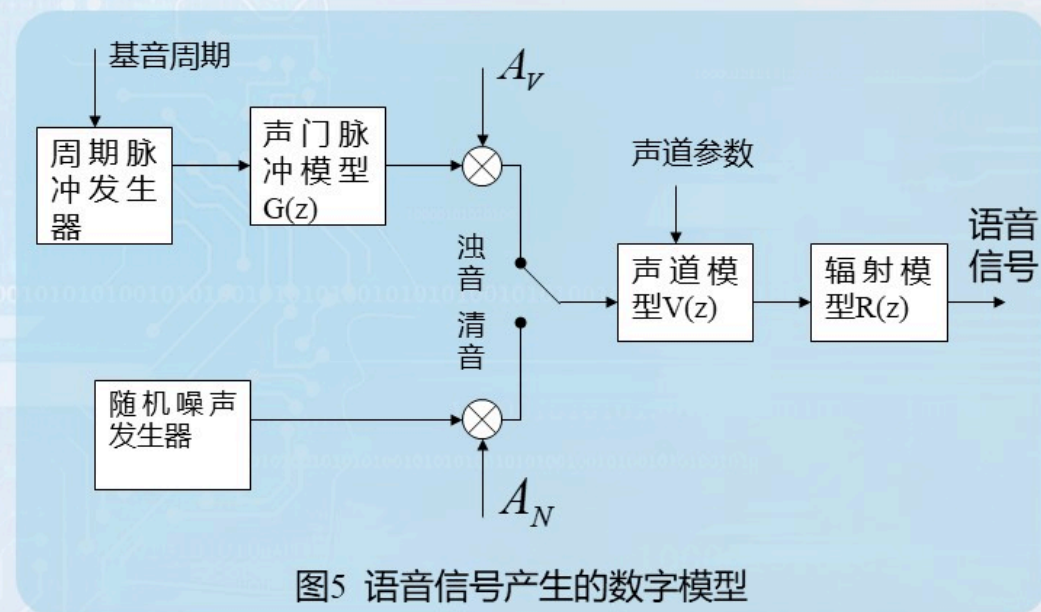


图4 嘴唇辐射模型

完整数学模型



系统函数

$$H(z) = G(z)V(z)R(z)$$

参数

参数：声道参数 α_m ， G ，基音周期， A_v 、 A_N 以及清浊音判别等都是随时间而缓慢变化的

填空题 1分

语音的数字模型是一个 [填空1] 的线性系统，在10~20ms的时间内是时不变系统