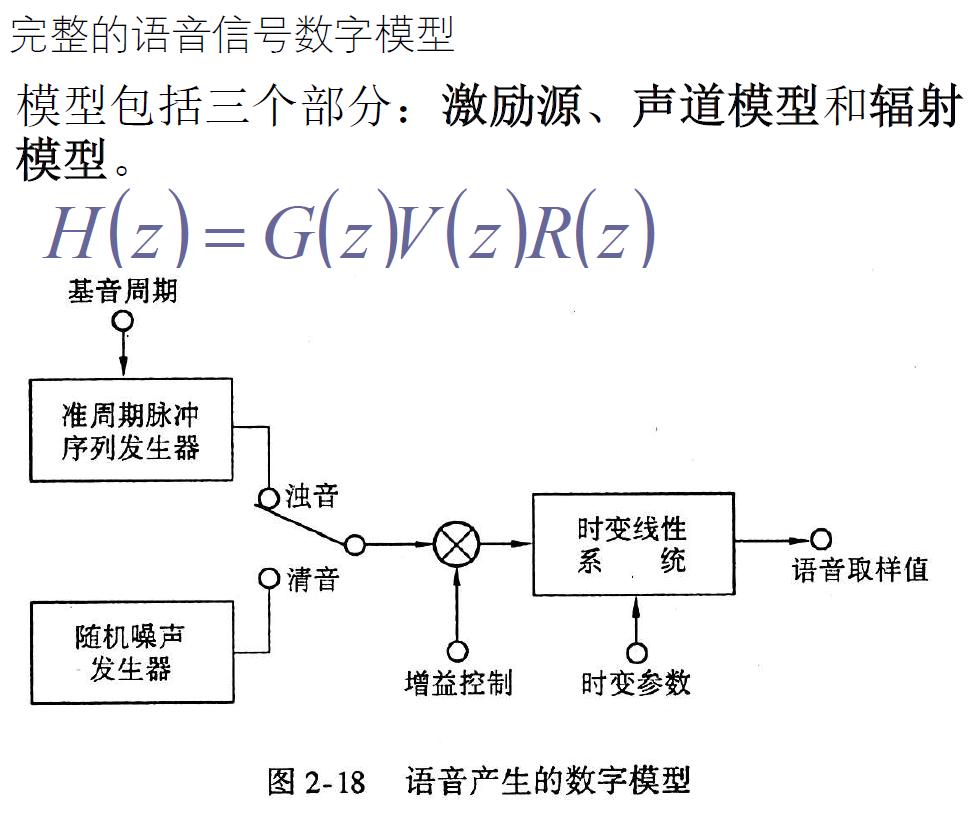
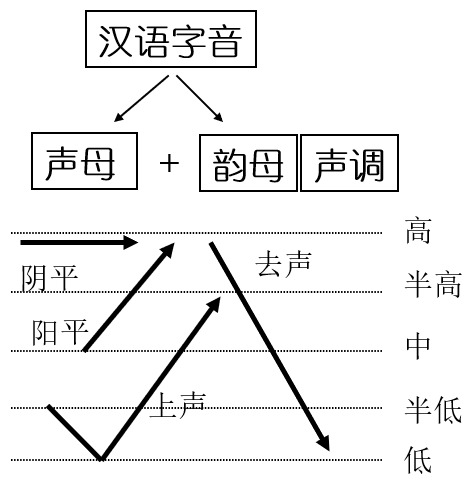
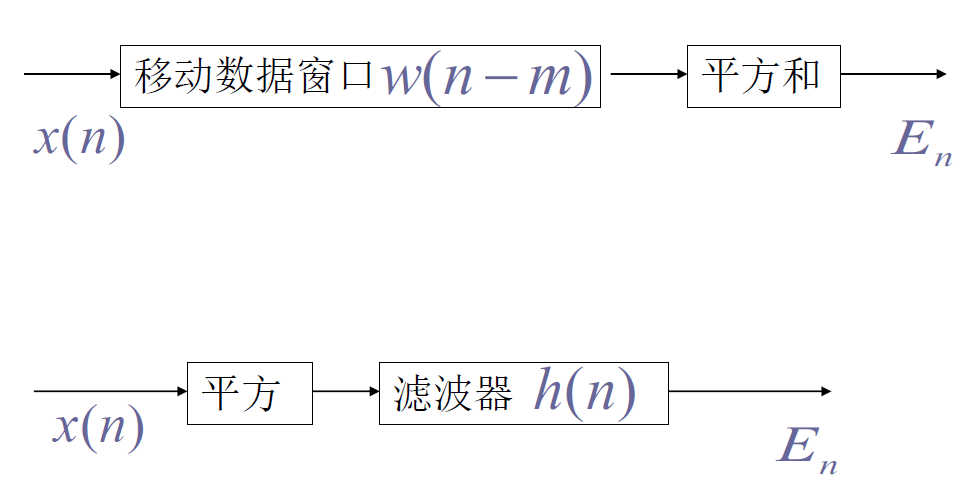
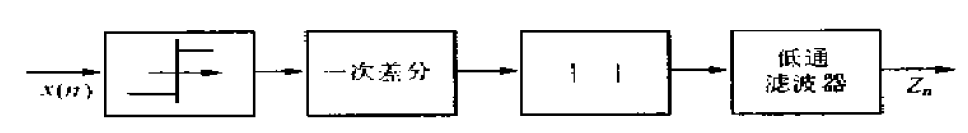
1. 简要叙述人体发出声音的基本原理与过程  
   语音是由空气流激励声道最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而产生的。对于浊音、清音和爆破音来说，激励源是不同的。  
   在说话时，肺部会主动收缩，把空气推向声带，使其振动，并通过咽喉、口腔和鼻腔等声道的变化，调整气流的速度、振动频率和共振峰位，进而塑造出不同的语音音色和音高。
2. 画出语音产生的数字模型，并做简要的说明。

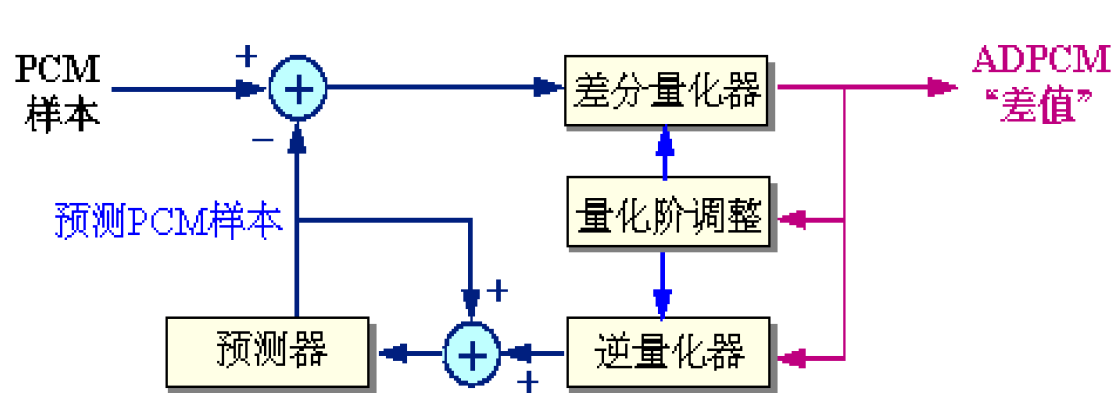
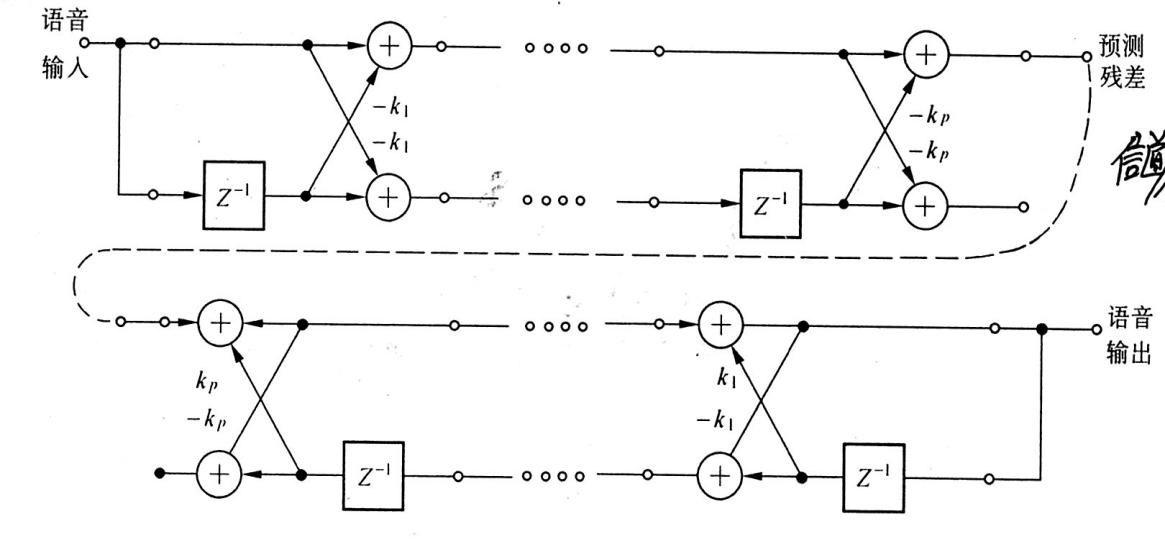
* 
* 该模型由三个子模型：激励模型、声道模型和辐射模型的级联来表示。其转移函数为：
* 这里H(z)是声道传递函数，既可以用声管模型，也可以用共振峰模型来描述。  
  在该图中，清浊音开关模拟了加在声道上的激励的改变情况；当开关接在浊音位置时，激励源是准周期脉冲序列发生器，其重复频率由基因频率来确定；当开关接在清音位置时，激励源是随机噪声发生器。图中线性时变系统主要用来模拟声道的特性。该系统的时变参数反映了语音的时变特性。  
  基本思想是将激励与系统相分离，使语音信号解体来分别进行描述，而不是直接研究信号波形本身的特性。

1. 为什么语音信号要采用“短时”方法进行分析处理，说明原因  
   语音信号具有很强的“时变特性”，有些波形段具有很强的周期性 ，有些波形段具有很强的噪声特性 ，且周期性语音和噪声性语音的特征也在不断变化之中。但在较短的时间内，语音信号的特征可以认为基本保持不变，这就是语音信号的“短时平稳性”。语音信号的短时平稳性是语音信号数字处理的基础，通常我们截取具有短时平稳性的一段语音进行分析处理，这一段语音称为一“帧”语音。  
   在较短的时间间隔内可以认为语音信号的特征基本保持不变，这是按帧进行语音处理的基础 。
2. 画图说明汉语中四种声调和基音频率的关系  
   声带开启和闭合使气流形成一系列脉冲每开启和闭合一次的时间称为音调周期或基音周期，其倒数称为基音频率。  
   无论是单音节语音还是连续语音，其中浊音段的基音频率是随时间而变化的，基音频率的不同轨迹成为声调。  
   
3. 窗函数有什么作用？简要说明语音信号窗函数宽度的选取原则。  
   窗函数的作用：用波形乘以窗函数，使边缘平滑；对信号谱和窗函数的傅立叶变换进行卷积。  
   窗函数如果选择太大，则语音信号随时间变化就很小，不能充分反映语音信号的幅度变化；而选择得太小，即选择等于或小于一个基音周期时，语音信号将按照信号波形的细微变化而起伏不定，以致短时能量不够匀化和平滑 。 因此，折衷考虑的值，在通常情况下，当取样频率为10 KHz 时，N= 100~200 被认为是合适的 。对应于短时语音段的长度为 10~20 ms 。
4. 矩形窗与汉明窗相比，各有什么特点  
   根据幅频特性曲线可以看出海明窗的带宽大约是同等宽度矩形窗带宽的两倍。 此外，海明窗在通带外的衰减要比矩形窗大得多 而且通带与阻带的起伏比较小。直角窗的谱平滑较好但波形细节丢失，海明窗则相反。
5. 为什么语音信号要采用非均匀量化，常用的非均匀量化有哪两种形式  
   均匀量化器特性曲线不能同输入信号的概率密度函数相匹配，量化噪声比较大。非均匀量化器考虑到量化特性曲线同输入信号的概率密度函数相匹配，使量化噪声得到进一步降低。从概念上讲，可以在 输入信号概率密度分布相对比较集中的区域，选择较小的量化间隔，而在其它区域选择较大的量化间隔，使过载噪声和颗粒噪声同时得到减少。  
   实现非均匀量化通常有两种途径。  
   一种是采用最佳非均匀量化器的迭代计算方法来设计量化器，这时量化噪声达到最小。  
   另一种是先将输入信号按照非线性曲线压缩然后进行均匀量化。 在接收端将译码后的样值再经过扩张恢复出应用的量化值。
6. 均匀量化时，量化比特数与信噪比有什么关系，定量说明  
   设是输入信号的均方差值，R是量化比特数，则量化后的信号噪声比为：

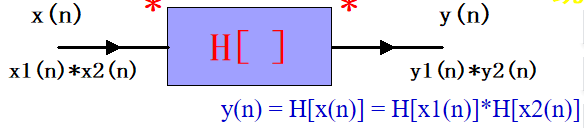
* 表明：每bit字长对信噪比的贡献为6dB

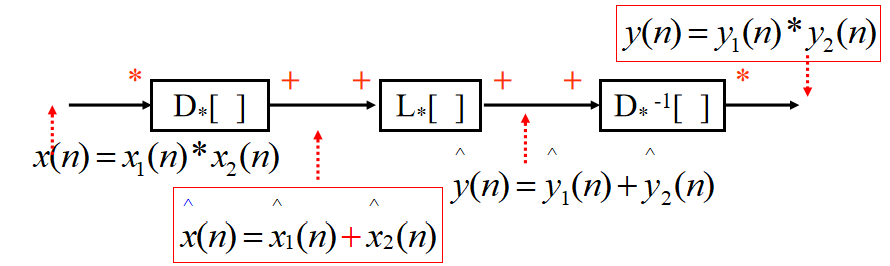
1. 人耳听觉感知的声音三要素  
   响度、音调、音色
2. 短时能量与短时平均过零率的含义，实现框图  
   短时平均能量：

* 上式变形得：
* 上述两种短时能量计算方法分别对应如下两个框图：  
    
  平均过零率：  
  对于离散时间信号的相邻两个取样值具有不同符号时便出现过零现象。单位时间过零的次数叫做过零率，也叫做平均过零数 。 对于窄带信号，过零率可以比较准确地反映该信号的频率。  
  然而，语音信号序列是宽带信号，所以不能简单地使用平均过零数计算频率。但是，仍然可以用短时平均过零数来得到其频谱的粗略估计。  
  短时平均过零数定义为：
* 实现框图为：  
  

1. 短时自相关函数与短时幅度差函数的主要用途  
   短时自相关函数是语音分析的重要参数，它有两个用途：  
   1、判断清/浊音，并估计基音周期  
   2、它的傅里叶变换就是短时谱（利用FFT减少运算量）  
   计算短时自相关函数的计算量很大，为了避免乘法运算，可利用差值（短时平均幅度差 AMDF）  
   平均幅度差函数能够代替自相关函数进行计算量更少的语音分析。
2. 清音和浊音的短时能量、短时自相关函数、短时过零率各具有什么样的特点？  
   在短时能量上：语音信号清音的能量值明显小于浊音段的能量 。  
   在短时过零率上：浊音语音的能量集中在大约 3 kHz 以下而对于清音语音 其能量大多出现在较高频率上 。 由于高的频率对应有高的过零率 低的频率对应有低的过零率。若测得的过零率高 则对应为清音 反之为浊音。  
   在短时自相关函数上：  
   短时自相关函数可以很明显地反映出浊音信号的周期性，而清音的短时自相关函数没有周期性，也不具有明显突起的峰值，其性质类似于噪声。
3. APCM编码中，改变量化阶大小的办法有哪些？  
   APCM是一种根据输入信号幅度大小来改变量化阶大小的一种波形编码技术。这种自适应可以是瞬时自适应，即量化阶的大小每隔几个样本就改变，也可以是音节自适应，即量化阶的大小在较长时间周期里发生变化。  
   改变量化阶大小的方法有两种：  
   一种称为前向自适应：是根据未量化的样本值的均方根值来估算输入信号的电平，以此来确定量化阶的大小，并对其电平进行编码作为边信息传送到接收端。  
   另一种称为后向自适应：是从量化器刚输出的过去样本中来提取量化阶信息。由于后向自适应能在发收两端自动生成量化阶，所以它不需要传送边信息。
4. 线性预测分析的基本原理  
   线性预测分析所包含的基本概念是，一个语音的抽样能够用过去若干个语音抽样的线性组合来逼近。通过使实际语音抽样和线性预测抽样之间差值的平方和 在一个有限间隔上达到最小值，即进行最小均方误差的逼近，能够决定唯一的一组预测系数。 这里的预测系数就是线性组合中所用的加权系数。
5. 为什么线性预测滤波器最常用的是全极点模型，说明理由
   1. 全极点模型最容易计算，对全极点模型作参数估计是对线性方程组的求解过程，相对来说比较容易。而若模型中含有有限个零点，则是解非线性方程组，实现起来非常困难。
   2. 有时无法知道输人序列，比如对一些地震应用、脑电图及解卷积等问题。
   3. 如果不考虑鼻音和摩擦音,那么语音的声道传递函数就是一个全极点模型。
   4. 人的听觉对于那种只能用零点来表现的频谱陡峭谷点是迟钝的。
6. 评价语音编码质量有哪两种方法，各有什么特点  
   客观标准：  
   SNR ( 对声码器无效）  
   主观标准：听后打分。  
   MOS( 平均意见得分）  
   DRT( 可懂度评价）  
   DAM( 判断满意度） 。
7. ADPCM的核心想法？画出其基本实现框图，并简要说明  
   核心想法是：①利用自适应的思想改变量化阶的大小，即使用小的量化阶 (step size) 去编码小的差值，使用大的量化阶去编码大的差值，②使用过去的样本值估算下一个输入样本的预测值，使实际样本值和预测值之间的差值总是最小。  
   
8. 语音合成的目的是什么，有哪几种合成方法  
   语音合成的目的是让机器能讲话。具体地说，语音合成是要根据输入的语音符号产生出具有一定音质和可懂度的语音。  
   语音合成技术可细分为四类：波形编码合成、 参数式分析合成 、 规则合成和文语转换 。
9. 简要叙述语音信号子带编码的基本工作原理  
   子带编码(Sub- Band Coding,简称SBC)也称为频带分割编码。它与10.7中将要分绍的自适应变换编码一样,也是在频域上寻求语音压缩的途径但与后者不同的是，它不对信号直接变换,而是首先使用带通滤波器组(也称为分析子带滤波器组)将语音信号分割成若干个频段,也称为子带。然后用调制的方法对滤波后的信号进行频谱平移变成低通信号(即基带信号)，以利于降低取样率进行抽取(即进行欠采样);再利用奈奎斯特速率对其进行取样，最后再进行编码处理。在接收端，信号通过内插恢复原始的取样率,通过调制恢复到原来的颊带,这样各个频带的分量相合成以得到重构的语音信号。
10. 子带编码的优点  
    SBC 的优点是对应于人的听觉特性，可以比较容易地考虑噪声的抑制;即各子带可以选用不同的量化参数以分别控制其信噪比，满足主观听觉的要求。例如，由于语音能量的不平衡，对于含有基音频率和第一共振峰的低频部分，对语音清晰度等主观品质影响较大,应分配比较多的信息，量化细些；反之，高频部分的量化就可粗些。这样，可以减少量化噪声对听觉的妨害程度，整体上也能降低比特数。另外，量化噪声只能出现在各被分割的频带内，对其他频带没有任何影响，所以可以较容易地控制噪声谱。
11. 比较语音的波形编码和参数编码的优缺点， 混合编码方法是如何利用两者的优点和克服两者的缺点的。  
    波形编码将语音信号作为一般的波形信号处理,具有适应能力强、语音质量好等优点,但所需要的编码速率高。它们在16~64 kbit/s的数码率上能给出高的编码质量,当数码率进一步降低时,其性能下降较快。  
    参数编码的优点是编码速率低。主要问题是重建语音质量差，特别是自然度较低。另外，这类编码器对讲话环境噪声十分敏感，需要安静的环境才能给出较高的可懂度。  
    混合编码结合了波形编码和声码器的优点:既利用了语音产生模型,通过对模型中的参数(主要是声道参数)进行编码,减少了波形编码中被编码对象的动态范围或数目;又使编码的过程产生接近原始语音波形的合成语音,以保留说话入的各种自然特征,提高了合成语音质量。
12. 语音信号的参数编码包含哪些参数：激励源参数及声道响应参数：每种声码器的特征参数，基音周期、清浊音信息、增益参数、声道参数，幅度等等（声码器种类不同参数也不同）。
13. 语音信号的参数编码中，为什么用反射系数代替滤波器参数：因为{ai}很小的变化都将导致合成滤波器极点位置的较大变化，甚至造成不稳定的现象，这表明需要用较多的bit数去量化每个{ai}。而反射系数在系数幅度值和谱灵敏度都是分布不相同不均匀的，采用非线性量化会更高效，降低码率。
14. 增量调制的基本原理：DM是对实际的采样信号与预测的采样信号之差的极性进行编码，将极性变成“0”和“1”这两种可能的取值之一。如果实际的采样信号与预测的采样信号之差的极性为“正”，则用“1”表示；相反则用“0”表示，或者相反。
15. 宽带语谱图和窄带语谱图各具有什么样的特点：宽带语谱图可以获得较高的时间分辨率，反映频谱的快速时变过程（较差的频率分辨率），给出语音的共振峰频率及清辅音的能量汇集区；窄带语谱图可以获得较高的频率分辨率，反映频谱的精细结构（较差的时间分辨率），时间坐标方向表示的是基音及其各次谐波。
16. 语音合成技术包括哪四类，简要说明：波形编码合成、参数式分析合成、规则合成和文-语转换。**波形编码：**以语句，短语为合成单元录音后编码压缩成库，播放时串接重组解码还原；**参数式：**以音节音素为合成单元提取参数编码成库，输出时取参连接编辑合成；**规则合成：**以音节因素的参数及其组合成句子的规则存储为库，输出利用规则转换为语音；**文语转换：**以文字串为输入的规则合成系统，分解文字串为词和读音符号，根据各种规则转换为代码串，再进行合成
17. 画出格型滤波器合成语音的框图，并简要说明：

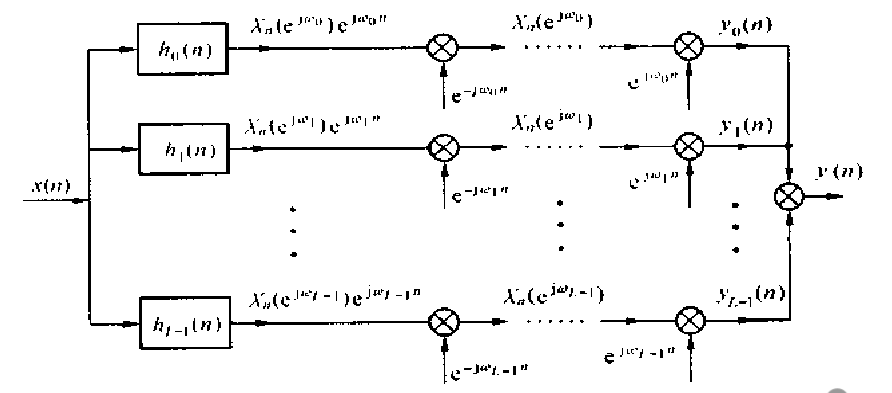
分析格型链上通道的末端与合成格型链上通道的始端相连接。下通道不相连，但分析链下的通道的输出信号等于合成链下通道的输入信号。具有相同的反射系数k1,k2,…，kp。但排列次序倒过来了，从中间向两边看，分析级和合成级成对出现。

1. 简要说明增量调制 (DM) 编码中颗粒噪声和斜率过载产生的原因及解决办法：1.产生原因：输入信号的变化速度超过反馈回路输出信号的最大变化速度时，就会出现斜率过载，后者受到量化阶大小的限制，然而量化阶大小为固定的；输入信号与预测信号的差值接近零的区域，增量调制器的输出出现随机交变的“0”和“1”。这就是粒状噪声。2.解决办法：为了尽可能避免出现斜率过载，就要加大量化阶Δ，如果要减小粒状噪声，就要减小量化阶Δ，可以用自适应增量调制来实现这两种办法。
2. 在子带编码中，子带分解为什么要采用正交镜像滤波器，并画出正交镜像滤波器的幅频特性曲线：因为实际中滤波器阻带和通带存在波动，易出现频带重叠，混叠，回声现象，需要更大的数码率，正交镜像滤波器可以消除混叠。
3. 画出同态滤波实现方法的框图，并对各个部分进行说明，并说明同态滤波的适用领域：



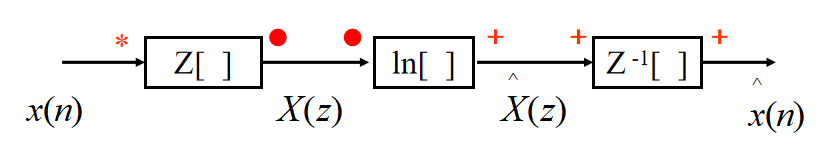
第一个是特征系统：把卷积转换为和，把非线性变为线性，第二个是满足叠加原理来进行线性处理的线性系统，第三个是逆特征系统：把和转换为卷积，把线性变为非线性。

适用领域：分离非加性组合信号，地震信号，图像信号，语音信号，通信中的衰落信号，调制信号，图像语音视频通信等等。

1. 简要说明人耳的听觉掩蔽效应：当两个响度不同的声音作用于人耳时，响度较高的频率成分的存在会影响到人对响度较低的频率成分的感受，使其变得不易察觉，且当一个声音突然停止，人耳约在150ms内对其他弱音听不清楚。
2. 简要说明语音编码质量的主观评价法和客观评价法含义：客观评价法：通过性能指标来评判如snr信噪比，谱失真测度，编码速率，顽健性等等。这些在最初的评价上很有用，但最终判决还是采用主观评价方法。主观评价法：如MOS(平均意见得分)，DRT(可懂度评价)，DAM(判断满意度)，自然度等，主观评价需要耗费大量人力和时间，结果存在非重复性，评价结果也很难对改进提供有益方向，因此需要客观评价，既能反映主观评价初步结果也可以为设计通信系统提供准则。
3. 人耳的主观感觉有哪几个指标可以衡量：响度（对于声音强弱的主观反应），音调（对声音频率高低的感受），音色（每个声音独有的属性），音长，音质。
4. 语音压缩编码要在哪几个方面折衷考虑：压缩编码需要在保持可懂度与音质，降低数码率和降低编码的复杂度三方面折衷考虑。
5. 语音编码要解决的两类问题是什么？简要说明：**第一类**是解决语音的存储问题，用尽可能少的存储单元存放尽量多的语音信息，在需要时将存储的语音进行回放，实现语音编码—存储—回放。对编码器的实时性要求不高，但要有较高的压缩比，以降低存储量。对解码要求算法尽量简单、成本要低，并基本上实时解码。**第二类**应用主要是用于通信，实现编码—传输—解码，具有抗干扰性强、保密性好，易于集成化以及适应性强的优点，用不同方法有不同的编码速率、编码质量。
6. 画出滤波器组求和法进行语音信号短时综合的实现原理框图，并做简要说明：

K从0到L-1。hk（n）为一个带通滤波器，中心频率为wk，yk（n）是第k个滤波器的输出，共有L个滤波器，都用相同的窗，有相同输入，将其输出相加，在恢复时这些通带信号被移回到原来的中心频率上，即得恢复信号y（n）

1. 同态信号处理具有什么样的特点？语音信号的声门激励复倒谱具有什么特点？：同态信号：非参数解卷，实现了将卷积关系变换为求和关系的分离处理，可以分离非加性组合。是一种非线性滤波，但服从广义叠加原理，可表示为三个同态系统级联；复倒谱：声门激励信号的复倒谱是无限冲激序列，幅度变，周期不变。且振幅随k值增大而衰减，除原点外，可以采用 “高复倒谱窗”从语音信号的倒谱中提取浊音激励信号的倒谱。从而可使用复倒谱提取基音。
2. 画出由原始语音求复倒谱的实现原理图，并用公式描述具体的实现过程：



（图应改成DFT）

先用窗w（n）选择一个语音段作为输入信号x（n）通过下式计算复倒谱。

算得频谱算得对数频谱算得复倒谱（此处用Z变换或者傅里叶都可以）

1. 求解语音信号的短时谱包络有哪几种方法？分别说明各自的特点。LPC复倒谱：运算量小，时间是FFT的一半，LPC系数直接估计：相比另外两种频谱包络不平滑，在谐波成分处比谐波之间效果好。LPC很好的表示共振峰结构但不出现额外的起伏和峰起。对元音发声是极佳的。FFT倒谱：计算便利，但易有相位卷绕，与LPC相比更好重现谱的峰值。
2. 在矢量量化器的设计中，所选择的失真测度必须具有哪几个特性，简要说明：1.必须在主观评价上有意义，小的失真应对应于好的主观语音质量。2.必须是易于处理的，在数学上易于实现，用于实际设计。3.平均失真存在且可以计算。4.易于硬件实现。
3. 在矢量量化过程中，初始码书的生成有哪几种方法，简要说明：1.随机选取法：从训练序列X中随机地选取J个矢量作为初始码字。2.分裂法1：(1)由所有训练矢量X，求形心Y1(0)

(2)利用较小的矢量ε将Y1(0)一分为二，以这两个矢量为初始码本，用LBG算法，求出Y1(1)，Y2(1)。(3)将Y1(1)，Y2(1)再二分为四，重复(2)直到有J个码字。3.分裂法2：求出Y1(0)后，对所有X∈S找max[ d(Xk,Y1(0)) ]对应的Xk作为Y2(1)，依次类推。4.乘积码书法：用若干个低维数的码书作为乘积码，求得所需的高维数的码书。