1. **填空题（有答案版）**

1、解卷的两种(具体)方法有( 线性预测分析 )和( 同态处理 )。

2、根据人的发音特征，在语音产生的数字模型中清音用( 随机噪声 )模拟,浊音用( 准周期脉冲序列 )模拟。

3、时域波形编码方法主要有( PCM )和( ADPCM )。

4、共振峰模型有:( 级联型 )、( 并联型 )、( 混合型 )三种。

5、语音信号数字分析或处理时为防止混叠干扰和工频干扰预滤波器是带通滤波器(上、下截止频率fH、fL、fx为取样频率)，多数语音编/译码器：fH=( 3.4kHz )、fL=( 60-100Hz )、fx=( 8-10kHz )。

6、语音质量的客观测量主要分两类,共三种方法即( 信噪比 )、( 分段信噪比 )、( 谱系数上的Euclid 距离 )。

7、认识和描述语音和语言的基本特征，即( 语音分析 )是语音信号处理的核心内容。

8、从短时 Fourier 分析的结果 Xn(ejw)恢复出原始语音信号x(n)的方法有:( 滤波器组求和法 )、( 滤波器组求和法 )。

9、参数编码实现语音通信的设备通常称为( 声码器 )。

10、量化可以分为两类，一类是( 标量量化 ), 另一类是矢量量化; 在矢量量化中是把所有M个量化矢量构成的集合称为( 码书或码本 )，将其中每个量化矢量(i=1，2，…,M)称为( 码字或码矢 ).

11、语音信号的预处理包括:( 预加重 )、( 加窗 )、( 分帧 )、( 端点检测 )等。

12、语音质量的主观测量方法有( 绝对等级评定试验(ACR) )、( 降低等级评定试验(DCR) )、( 诊断测量试验(DAM) )。

13、从短时 Fourier 分析的结果 Xn(ejw)恢复出原始语音信号 x(n)的方法有:( 滤波器组求和法 )、( FFT求和法(叠接相加法) )。

14、短时傅立叶变换的总取样率为( 时域 )与( 频域 )取样率之乘积。

15、可以用于浊音基音周期提取的时域参数(函数)有( 短时自相关 )、( 短时修正自相关 )、( 短时平均幅度差 )。

16、STFT线性滤波实现，窄带低通滤波器w(n)，带宽为B，在时域内，以( 2B )速率对 Xn(ejw)取样，不产生混叠失真.

1. **填空题（无答案版）**

1、解卷的两种(具体)方法有( )和( )。

2、根据人的发音特征，在语音产生的数字模型中清音用( )模拟,浊音用( )模拟。

3、时域波形编码方法主要有( )和( )。

4、共振峰模型有:( )、( )、( )三种。

5、、语音信号数字分析或处理时为防止混叠干扰和工频干扰预滤波器是带通滤波器(上、下截止频率fH、fL、fx为取样频率)，多数语音编/译码器：fH=( )、fL=( )、fx=( )。

6、语音质量的客观测量主要分两类,共三种方法即( )、( )、( )。

7、认识和描述语音和语言的基本特征，即( )是语音信号处理的核心内容。

8、从短时 Fourier 分析的结果 Xn(ejw)恢复出原始语音信号x(n)的方法有:( )、( )。

9、参数编码实现语音通信的设备通常称为( )。

10、量化可以分为两类，一类是( ), 另一类是矢量量化; 在矢量量化中是把所有M个量化矢量构成的集合称为( )，将其中每个量化矢量(i=1，2，…,M)称为( ).

11、语音信号的预处理包括:( )、( )、( )、( )等。

12、语音质量的主观测量方法有( )、( )( )。

13、从短时 Fourier 分析的结果 Xn(ejw)恢复出原始语音信号 x(n)的方法有:( )、( )。

14、短时傅立叶变换的总取样率为( )与( )取样率之乘积。

15、可以用于浊音基音周期提取的时域参数(函数)有( )、( )、( )。

16、STFT线性滤波实现，窄带低通滤波器w(n)，带宽为B，在时域内，以( )速率对 Xn(ejw)取样，不产生混叠失真.

答案：1、线性预测分析 同态处理

2、随机噪声 准周期脉冲序列

3、PCM ADPCM

4、级联型 并联型 混合型

5、3.4kHz 60-100Hz 8-10kHz

6、信噪比 分段信噪比 谱系数上的Euclid 距离

7、语音分析

8、滤波器组求和法 FFT求和法(叠接相加法)

9、声码器

10、标量量化 码书或码本 码字或码矢

11、预加重 加窗 分帧 端点检测

12、绝对等级评定试验(ACR) 降低等级评定试验(DCR) 诊断测量试验(DAM)

13、滤波器组求和法 FFT求和法(叠接相加法)

14、时域 频域

15、短时自相关 短时修正自相关 短时平均幅度差

16、2B

1. **简答题**
2. 试画出一种SIFT的线性滤波实现方框图,并给出相应公式？

2、线性预测分析的解法中自相关法和协方差法有何各自的特点?

答：(1)自相关发适用于平稳信号面协方差法适用于非平稳信号

(2)自相关对摩擦音来说可以给出比较好的结果,而协方差法对于周期性语音可以给出比较好的结果结

(3)自相关法引人窗函数来截取,必然会引人误差,其系数的精度不高,但其解很稳定;而协方差法则,精度高,不能保证其稳定性

3、什么是短时平均过零数,其与信号频率和采样率的关系如何;对于语音信号的短时平均过零如何定义，通过它如何区分清音与浊音?

答：**短时平均过零数：**单位时间内的过零次数,称为平均过零数也叫平均过零率2Fo/Fs

**语音信号的短时平均过零数:**对于语音信号x(n)的短时平均过零定义为

**区分清音与浊音:**浊音的低频分量较多所以其过零数也低，清音的高频分量较多齐过零数较高

4、元音信号短时频谱图有两种变化,试说明两种变化及其产生原因?

答：快变化:周期性激励引起,基音频率的各次谐波;

慢变化:声道共振特性引起,各共振峰的频率和带宽。

5、简述什么是语音信号的“短时分析技术”,怎样用 AMDF 判断基音频率?

答：**短时分析技术:**语音信号有时变特性，是一个非平稳的随机过程。但在一个短时间范围内其特性基本保持不变。即语音的“短时平稳性”。“短时分析”，即对语音信号流采用分段处理。将其分为一段一段来分析，其中每一段称为一“帧”,由于语音通常在 10~30ms 之内是保持相对平稳的,因而帧长一般取为10~30 ms。

**AMDF 判断基音频率:**对于浊音信号,在周期倍数点上,幅值相等,Fn=0.

6、短时平均能量的用途是什么?其与短时平均幅度处理信号算法和效果的主要差异?

答：**用途**：可以区分清音段和浊音段,因为浊音时能量比清音时能量大

可区分声母与韵母得分界,无声与有声的分界

作为一种超音段信息用于语音识别中

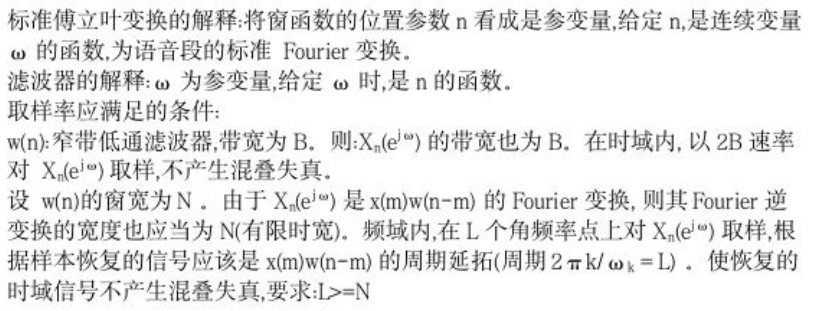
**主要差异:**

A:算法上短时平均能量对于每个序列平方后加和,相对于短时平均幅度差取绝对值后加和计算量大。

B:效果上来说可以达到同样的用途,但短时平均能量的动态范围大。

7、写出短时傅立叶变换的两层含义(解释),并据此说明短时傅立叶变换时域及频域取样率应满足的条件

答:



8、在语音信号分析中,窗口的形状和长度是重要参数，通常选择哪种窗口,请将其与直角窗进行比较,并说明其处理结果的不同?窗口长度N 如何确定,N 很大和太小时对时域分析的影响?

答：

**窗口形状：**在短时傅立叶变换中通常选择海明窗，海明窗的第一零点频率为之比直角窗大,其带外衰减比直角窗大。通过卷积,直角窗的谱平滑较好,但波形细节丢失;而海明窗则相反。

**窗口大小：**窗口长度对于能否反映语音信号的幅度变化将起决定作用。

N很大,它等效于带宽很窄的低通滤波器,此时时域参量随时间变化很小,不能反映语音信号的特性变化,波形的变化细节就看不出来了;

反之,N 太小时,滤波器的通带变宽,时域参量随时间有急剧变化,不能得到平滑的函数。

9、解释什么是人耳的掩蔽效应,其掩蔽类型有那几种和各自的特点?

答：**人耳的掩蔽效应:**一个较弱的声音(被掩蔽音)的听觉感受被另一个较强的声音(掩蔽音)影响的现象称为人耳的“掩蔽效应”

**掩蔽类型:**

**频域掩蔽:**所谓频域掩蔽是指掩蔽声与被掩蔽声同时作用时发生掩蔽效应,又称同时掩蔽，这时,掩蔽声在掩蔽效应发生期间一直起作用,是一种较强的掩蔽效应，通常,频域中的一个强音会掩蔽与之同时发声的附近的弱音,弱音离强音越近,一般越容易被掩蔽;反之,离强音较远的弱音不容易被掩蔽，一般来说,低频的音容易掩蔽高频的音;在距离强音较远处,绝对闻阈比该强音所引起的掩蔽阈值高,这时,噪声的掩蔽阈值应取绝对闻阈.

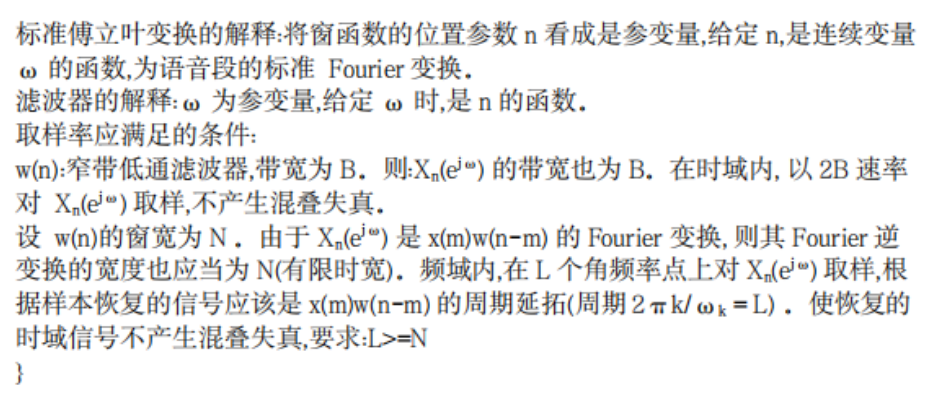
**时域掩蔽:**所谓时域掩蔽是指掩蔽效应发生在掩蔽声与被掩蔽声不同时出现时,又称异时掩蔽。异时掩蔽又分为导前掩蔽和滞后掩蔽，若掩蔽声音出现之前的一段时间内发生掩蔽效应,则称为导前掩蔽;否则称为滞后掩蔽。产生时域掩蔽的主要原因是人的大脑处理信息需要花费一定的时间,异时掩蔽也随着时间的推移很快会衰减,是一种弱掩蔽效应。一般情况下,导前掩蔽只有 3ms-20ms,而滞后掩蔽却可以持

续 50ms-100ms.

10、写出短时傅立叶变换的两层含义(解释),并据此说明短时傅立叶变换时域及频域

取样率应满足的条件。

答：



11、线性预测模型中一般采用什么模型?为什么?

答：选用全极点模型

全极点模型最容易计算，有时无法知道输人序列，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，人的听觉对那种只能用零点来表现的频谱陡峭谷点是迟钝的

12、简述什么是均匀量化PCM,什么是对数PCM?为了满足高质量的电话通信要求均匀量化PCM和μ律PCM本用需要多少bit?

答：语音信号带宽WHz.2W速率采样,每样本用Bbit表示,总比特率:2WBbit/s.

特点:均匀量化 PCM 与普通 AD完全相同,未利用语音信号的任何性质,未压缩。

非线性量化的工程方法:先进行非线性压缩后,再进行线性量化,对数PCM:先对数

幅度压缩,解码用指数函数扩张，

位μ律PCM的性能与11位均匀量化PCM 相同,可以满足高质量的电话通信要求

13、解释语音预处理中预加重的原理和如何进行预加重和去加重?

答;

**原理:**语音的一个特征:约8kHz高频端按-6dB/倍频程跌落.

**预加重:**提升语音的高频部分,使信号的频谱变得平坦,可在反混叠滤波前,可压缩动态范围,提高信噪比，也可在数字化后、参数分析之前，

**预加重和去加重:**预加重用6B/倍频程的提升高频特性的预加重数字滤波器:式中,为常系数,值接近于1.通常取μ=0.92~0.94.H(z)=1-uz-1,恢复原信号,对测量值进行去加重处理,即加上-6 dB/倍频程的下降的频率特性来还原成原来的特性。

14、请说明什么是窄带语谱图、宽带语谱图,其各自特点和主要用途?

答：

**语谱图:**显示大量与语音特性有关的信息,它综合了频谱图和时域波形的特点,明显地显示出语音频谱随时间的变化情况,或者说是一种动态的频谱。用语谱图可确定语音参数,例如共振峰频率及基音频率，语谱图的纹路,称为“声纹”;因人而异,可用于讲话人识别

**宽带语谱图:**带宽约为300z,具有良好的时间分辨率,但是频率分辨率较差;

**窄带语谱图:**带宽约为45Hz,具有良好的频率分辨率,但是时间分辨率较差。在不同的语谱图上所表现的语音特征也不同。

**宽带语谱图的典型谱型:宽横杠(Bar):**代表元音的共振峰位置,表现为图中与水平时间轴平行的较宽的黑杠,不同元音的共振峰位置不同,根据宽带语谱图上各横杠的位置可以区分不同的 元音,不同人发音的第一共振峰位置会不同,但其分布结构是相似的。冲直条(spike)代表塞音(b,d,g,p,t,k)或塞擦音(z,zh,i.c,ch,g),表现为图中与垂直频率轴平行的较宽的 黑条,在时间上持续时间很短,在频率轴上集中区位置随不同的辅音而不同 摩擦乱纹(61):代表摩擦音(s,sh,x,f,h)或者送气音的送气部分,表现为图中无规则的乱纹。

**窄带语谱图的典型谱型:窄横条:**代表元音的基音频率及各次谐波,表现为图中与水平轴平行的线条,窄横条在频率轴的位置对应了音高频率值,随时间轴的曲折、升降变化代表了音高变化的模式无声隙段:对应于语音的停顿间隙,在图中表现为空区,在两种语谱图中都存在。

1. **论述题**

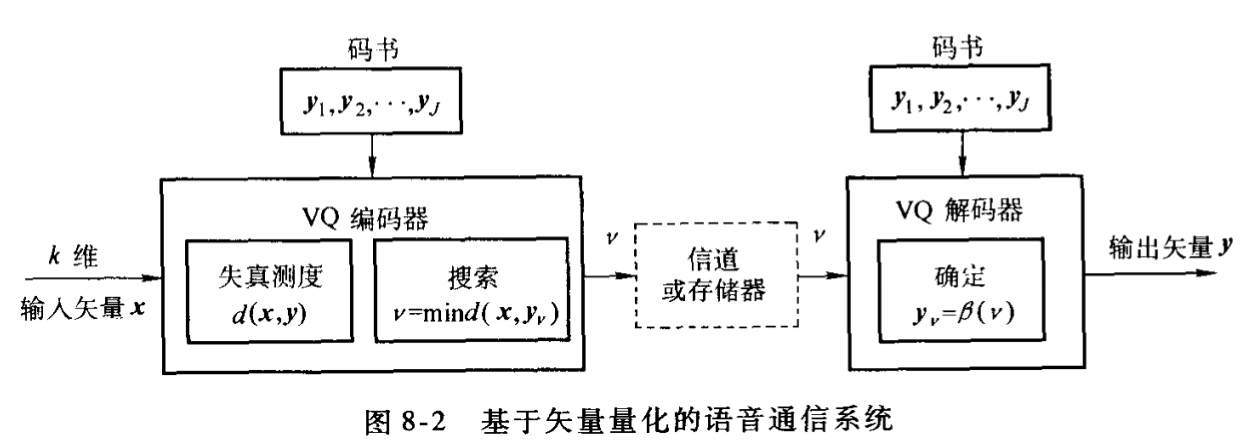
1、画图说明语音信号中卷积同态系统及各子系统模型，标注其运算符号的改变，用英文和公式简要说明其处理过程。

答：

2、画出基于矢量量化的语音通信系统图，分析工作原理，及性能。

答：

**矢量量化的语音通信系统图：**



**工作原理：**

(1)每输入一帧语音(帧长N)，形成与之相应的k维特征矢量(k<<N)，并送人 VQ 编码器.

(2)根据输入特征矢量从编码器码书中选择一与之失真误差最小的码矢，取的编码(标号)v，即:

(3)传输v，若不产生误差，则收端的信号仍是v

(4)解码器按照v从解码器码书中选出具有相同下标的码字,作为输出，只即为万的重构矢量(恢复矢量)，即:

**性能：**

收发两端即编码器译码器两端都没有反馈回路，因此很稳定

由于传输或存储的并不是矢量本身而是其序号，因此具有高度的保密性将语音信号波形的K个样点的每一帧,或有K个参数的每一参数帧,构成K维空间中的一个矢量,然后对这个矢量进行量化。(5分)

3、详细画出语音信号的产生的模型,并简要分析其原理。

答：

**语音信号的产生的模型：**

**原理**：1、发浊音的激励源为声带(声门)，以准周期脉冲序列来模拟声带的震动根据测结果脉冲波类似于斜三角脉冲(2分)

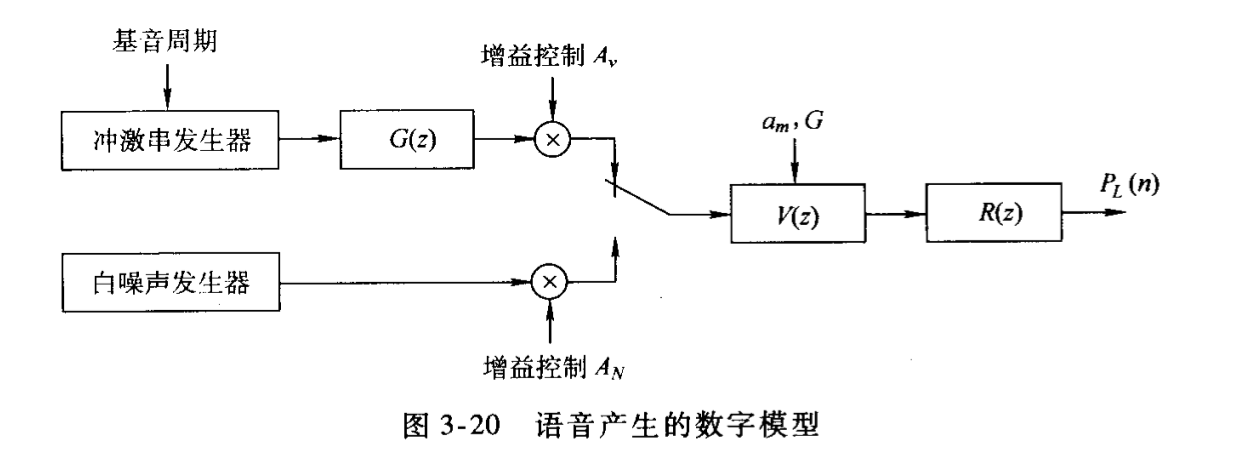
2、以随机噪声发生器来发清音时声带中空气湍流的震动;(1 分)

3、时变线性系统(声道、辐射模型)的参数是时变的，包括增益及线性系统的滤波器参数，反映了语音的时变特性，(2分)

4、画出语音信号处理系统框图，说明图中滤波器的作用;说明预加重的位置及其作用.

答：

**语音信号处理系统框图：**

****

**滤波器的作用：**

**反混叠滤波器**:低通，在信号带宽不明确时，使其带宽限制在某个范围内(1分)

**平滑滤波器:**对重构的语音波形的高次谐波起平滑作用，以去除高次谐波失真(1分)

**预加重:**提升语音的高频部分，使信号的频谱变得平坦，可在反混叠滤波前，可压缩动态范围，提高信噪比，也可在数字化后、参数分析之前，(3分)

1. **计算题**

1、下图是元音段[]的LPC谱和语音信号谱的比较，其中信号谱由T分得到;LPC谱:Hamming窗，14阶，自相关法;请说明1、分别从两种谱中能得到有关于语音的什么参数?2、两种谱的区别、联系和产生的原因?

答：FFT 谱中可以看到包络和谐波，可以得到共振峰和基音频率(2分)

LPC谱中主要是包络仅能得到共振峰(2分)

1、该信号的谱有很好的谐波结构;

在信号谱的峰值处，IPC谱和信号谱匹配良好;(1分)

在信号谱的谷底处，LPC谱和信号谱匹配较差;(2分)

2、浊音语音谱，在谐波成分处匹配效果要远比谐波之间好得多。(1分)

原因:源于方均误差最小的准则，谱值大时误差要小。(2分)

2、试根据下面元音[i]的波形和短时频谱图分别说明短时谱图特点和窗函数及窗宽对短时谱的影响(元音田的基音周期大约是13ms，此题不用画图)

答：

1)短时频谱图有两种变化

快变化:周期性激励引起，基音频率的各次谐波

慢变化:声道共振特性引起，各共振峰的频率和带宽.

2)两个频谱图间的差别:

矩形窗时:谐波各峰较尖锐，谱图较破碎(类似于噪声)，主较窄(较高频率分辨率);旁瓣较高，“泄漏”严重;

Hamming 窗时:短时频谱平滑些.

3)窗宽对短时频谱的影响

窗宽6.4ms，元音的基音周期大约是13ms;窗选语音段长不到一个基音周期，丢失了基音周期的信息;频的快变化(谐波频率)消失。频的慢变化(较宽的峰)保留，是声道的共振特性，矩形窗比Hamming时，呈现较多的细致结构，由于矩形窗比Hammming窗具有更高的频率分辨率。

3、设 Matlab的工作空间中有一段语音数据x,其长度为2000点，抽样率为10k;试完成下面的脚本程序来进行短时平均能量的计算(写出2个程序空即可)。并写出算式，计算该算法本身所需进行的乘法和加法的次数。

WL=100;%窗长

ML=20:%帧间隔

M= ;%所计算的帧数

E-zems(1,M);

For i=1:M

E(1 i)= ;%短时平均能量

End

**答：**

