**《语音编码》知识点汇编**

1、语音编码是语音相关各研究方向中最早获得成功应用的研究方向，它是支撑各种语音通讯方式发展（市话、长途电话、移动通讯、VoIP、即时通讯等）的最关键技术之一。也普遍存在于各种音视频多媒体、流媒体数据中。

2、语音数据不论是磁盘上存储，还是在网络信道上实时传输，都需要考虑数据规模的大小。每秒钟语音需要占据的磁盘空间可以计算为：采样频率◊量化bit数◊声道数。要达到实时通讯的效果，语音的传输速率也应该这样计算，单位为比特率，bps 或bit/s。

3、根据采样定理，采样频率要大于信号最大频率的两倍，否则就会发生混叠失真。所以采样前应该用低通滤波器对信号进行滤波，只保留必要的频率成分。对语音而言，采样频率不应该低于8KHz, 即4kHz以内的频率成分如果损失，会影响语音的可懂度。采样频率越大，音质越好，但语音数据的规模也就越大。同样、量化bit数越高，量化误差越小，声道数越多，声音越立体，声音质感越好，但数据规模也会越大。为了节省磁盘空间，或降低对传输速率（网络带宽）的要求，需要对语音进行编码压缩。

4、语音编码方法可以粗略地分为“波形编码”、“参数编码” 和“混合编码”，其概念见课件。

5、信噪比的计算公式（见课件），其反映了信号的失真程度，单位为dB，信噪比越大，失真程度越小。

6、脉冲编码调制（Pulse Code Modulation，PCM）是最简单的编码方式，就是声卡采样量化后的形式，其采样点按顺序存放。根据采样值中大幅值比较少的特性，可以对采样值进行对数变化，然后可以降低量化比特数，进行8bit量化，这种编码方式被称为压扩PCM。要求会计算PCM和压扩PCM的比特率。

7、自适应差分脉冲编码ADPCM是一种应用广泛的波形编码，利用了相邻采样点间的强相关性了。

8、掌握DPCM的编码原理，内容见课件。掌握需要在编码器中包含一个解码器的道理。

9 各种波形编码算法的输入一般是PCM或者压扩PCM。由于波形编码以采样点为单位来进行编码压缩，为了达到压缩的效果，一般需要减少压缩后信号的量化bit数，这通过量化器来完成，通常会带来量化误差。可以理解为新量化bit数没有足够的分辨率来存储压缩后的采样值。

10、掌握ADPCM的编码原理，内容见课件。要求能写出ADPCM编码的时域计算公式。边信息是以帧为单位的信息，有别于波形编码中其它以采样点为单位的信息。掌握ADPCM中的边信息有哪些。

11、掌握LPC-10的编码原理，内容见课件。其编码参数有哪些？其中线性预测系数采用了反射系数的形式，二者可以互相转换。反射系数插值特性好。解码端需要生成特定形状的声门波（以基音周期为周期）。

12、混合编码不生成特定形状的声门波，而是用“合成分析”和“闭环搜索”的方法来确定能最好生成语音信号的声门波，遍历所有的候选声门波，逐一合成语音，选择其中误差最小的候选声门波。为了使候选声门波更简洁，更容易编码表示，采用长时相关滤波器去除声门波中周期相关性，得到更稀疏的没有周期性的激励信号，因而语音生成时，也需要多一个长时综合滤波器。并在合成分析中，计算合成语音与真实语音间的误差时，采用了感觉加权滤波器。其效果是计算误差时降低误差信号中被掩蔽频段的贡献，增强未被掩蔽误差频段的贡献。详细的计算公式见课件。

13 在频谱上0dB的等响度曲线也被称为闻域曲线，低于此闻域曲线的信号无法被人耳听见。语音信号的存在会改变此闻域曲线而形成掩蔽曲线，某个频段信号强，掩蔽曲线也会被相应拉高。从而使之前可被人耳听见的信号，无法再被听见。具体图示见课件。

14 CELP的候选激励信号从一个码本中挑选，最佳码本（合成分析后感觉加权误差最小）的编号被传输到解码端，从同样的码本中获得该激励，用于语音生成。其过程见课件。

15 解码语音的质量常用MOS评分来评估。

16 深度自编码器等深度学习技术可用于实现更高效的语音编码。