# 语音编码概述

在通信系统中，语音编码是相当重要的。因为在很大程度上，语音编码决定了接收到的语音质量和系统容量。在移动通信系统中，宽带是十分宝贵的。低比特率语音编码提供了解决该问题的一种方法。在编码器能够传送高质量语音的前提下，如果比特率越低，可在一定的宽带内能传更多的高质量语音。

　　语音编码为信源编码，是将模拟语音信号转变为数字信号以便在信道中传输。语音编码的目的是在保持一定的算法复杂程度和通信时延的前提下，占用尽可能少的通信容量，传送尽可能高质量的语音。语音编码技术又可分为波形编码、参量编码和混合编码三大类。

　　波形编码是对模拟语音波形信号经过取样、量化、编码而形成的数字语音技术。为了保证数字语音技术解码后的高保真度，波形编码需要较高的编码速率，一般在16~64kbps.可对各种各样的模拟语音波形信号进行编码均可达到很好的效果。它的优点是适用于很宽范围的语音特性，以及在噪音环境下都能保持稳定。实现所需的技术复杂度很低，而费用中等程度，但其所占用的频带较宽，多用于有线通信中。波形编码包括脉冲编码调制（PCM）、差分脉冲编码调制（DPCM）、自适应差分脉冲编码调制（ADPCM）、增量调制（DM）、连续可变斜率增量调制（CVSDM）、自适应变换编码（ATC）、子带编码（SBC）、和自适应预测编码（APC）等。参量编码是基于人类语言的发声机理，找出表征语音的特征参量，对特征参量进行编码的一种方法。在接收端，根据所收的语音特征参量信息，恢复出原来的语音。由于参量编码只需传送语音特征参数，可实现低速率的语音编码，一般在1.2~4.8kbps。线性预测编码（LPC）及其变形均属于参量编码。参量编码的缺点在于语音质量只能达到中等水平，不能满足商用语音通信的要求。对此，综合参量编码和波形编码各自的长点，即保持参量编码的低速率和波形编码的高质量的优点，又提出了混合编码方法。

　　混合编码是基于参量编码和波形编码发展的一类新的编码技术。在混合编码的信号中，既含有若干语音特征参量又含有部分波形编码信息。其编码速率一般在4~16kbps。当编码速率在8~16kbps范围时，其语音质量可达商用语音通信标准的要求。因此混合编码技术在数字移动通信中得到了广泛应用。混合编码包括规则脉冲激励长期预测编解码器（RPE-LTP）、矢量和激励线性预测编码（VSELP）和码激励线性预测编码（CELP）等。

## G.711

G.711也称为PCM（脉冲编码调制），是国际电信联盟订定出来的一套语音压缩标准，主要用于电话。它主要用脉冲编码调制对音频采样，采样率为8k每秒。它利用一个 64Kbps 未压缩通道传输语音讯号。起压缩率为1：2，即把16位数据压缩成8位。G.711是主流的波形声音编解码器。

　　G.711 标准下主要有两种压缩算法。一种是µ-law algorithm （又称often u-law, ulaw, mu-law），主要运用于北美和日本；另一种是A-law algorithm，主要运用于欧洲和世界其他地区。其中，后者是特别设计用来方便计算机处理的。

## G.723

　　G.723是ITU-T在1996年制订成型的一种多媒体语音编解码标准。其典型应用包括VoIP服务、H.324视频电话、无线电话、数字卫星系统、数电倍增设备(DCME)、公共交换电话网（PSTN）、ISDN及各种多媒体语音信息产品。G.723标准传输码率有5.3kb/s和6.3kb/s两种，在编程过程中可随时切换。该标准主要包含了编码算法和解码算法。原理是：从采集的语音信号中解析出声道模型参数，构造一个合成滤波器，采用合适的激励源激励,编码传输的参数主要是激励源与合成滤波器的参数。5.3kb/s的编码器采用代数码线预测激励（ACELP）；6.3kb/s的编码器则采用多脉冲最大似然量化(MP-MLQ)激励。根据传输编码参数,可重构激励源与合成滤波器进行解码,还原出来的数字语音信号经D/A转换器转换成模拟语音信号。

　　G.723编码器采用LPC合成－分析法和感觉加权误差最小化原理编码。G.723标准可在6.3kbps和5.3kbps两种码率下工作。对激励信号进行量化时，高速率（6.3kbps）编码器的激励信号采用多脉冲最大似然量化（MP&shy;-MLQ），低速率（5.3kbps）编码器的激励信号采用代数码本激励线性预测（ACELP）。其中，高码率算法（6.3kbps）具有较高的重建语音质量，而低码率算法（5.3kbps）的计算复杂度则较低[12]。与一般的低码率语音编码算法一样，这里的G.723标准采用的线性预测的合成分析法也就是我们通常所说的Analysis-by-Synthesis。

　　G.723建议采用的是定点运算。根据传输编码参数，可重构激励源与合成滤波器进行解码，还原出来的数字语音信号经D/A转换器转换成模拟语音信号。G.723算法对语音信号有很好的编解码效果， 同时也可处理音乐和其它声音信号，典型输入是64kbps（8k×8）或128kbps（8k×16）的A律或u律的PCM 采样语音信号。每次处理一帧语音信号，每帧240个采样点（30ms）。在5.3kbps的码率下，每帧语音被压缩成20个。编码器先对语音信号进行传统电话带宽的过滤，再将输入的16 bit 线性脉冲编码调制（PCM）码流分成长度为240个样点的语音帧，以帧为单位进行编码。首先把1帧信号分成4个长度为60个样点的子帧，接着进行高通滤波，这样就可以去掉直流分量；分别进行10阶线性预测编码（LPC）分析，从而得到各子帧的LPC参数，并把最后一个子帧的LPC参数转化成线谱对（LSP）参数，进行矢量量化编码，送到解码器。利用未量化的LPC参数构造短时感觉加权滤波器，对信号滤波后得到感觉加权的语音信号。每两个子帧（120样点）搜索一个开环基音值，并以此为依据，为每一个子帧构造一个谐波噪声成形滤波器，对感觉加权的语音信号进行滤波。每一子帧的LPC综合滤波器、感觉加权滤波器和谐波噪声成形滤波器级联起来，构成一个联合滤波器，利用它的冲击响应和开环基音周期，对每一子帧进行闭环基音搜索，对开环搜索的结果进行修正。同时通过一个5 阶基音预测器对信号进行预测，得到相应子帧的残差信号[13]。再进行固定码本搜索，也就是对每一子帧的残差信号进行矢量量化，先用实际信号减去预测信号得到残差信号，再用一个脉冲序列通过组成滤波器来模拟残差信号，在最小误差准则下，将得到的一系列参数，如滤波器系数、脉冲位置、脉冲幅度打包成一个比特流传送出去最后还要进行状态更新。

　　G.723解码器也是以帧为单位进行解码的。编码器输出的基音周期和差分值都被传送到解码器。首先通过激励解码器，基音解码器和LSP解码器对量化的LPC进行解码，然后构造LPC合成滤波器，对于每个子帧都需要进行自适应码本激励和固定码本激励的解码，然后输入到合成滤波器中，自适应后置滤波器由共振峰后置滤波器和前后向基音后置滤波器组成，激励信号输入到基音后置滤波器中，输出信号输入到合成滤波器中，其输出再输入到共振峰后置滤波器中，一个增益缩放单元保证共振峰后置滤波器的输入信号的能量电平。

　　G.723算法对语音信号有很好的编解码效果，同时也可处理音乐和其它声音信号,典型输入是64kb/s（8k×8）或128kb/s(8k×16)的A律或μ律的PCM采样语音信号。每次处理一帧语音信号,每帧240个采样点（30ms）。在5.3kb/s的码率下，每帧语音被压缩成20个字节传输;在6.3kb/s的码率下，每帧语音被压缩成24个字节传输。

G.723.1 是一种能够以非常低的位速率压缩语音或者音频信号的压缩技术，它是H.324标准系列的一部分。这种编译码方式可以使不同的通信设备利用一种标准化的通信协议进行通信。它可以用在以6.4或者5.3Kbps产生数字音频信号的数字式PBX/ISDN上。较高的比特速率可以提供较高的声音质量。较低的比特速率可以为系统设计人员提供更高的灵活性。ITU-T在它的G系列建议中详细介绍了这项标准。

## G.726

G.726有四种码率：64, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)，最为常用的方式是 32 kbit/s，但由于其只是 G.711速率的一半，所以可将网络的可利用空间增加了一倍。G.726具体规定了一个 64 kbpsA-law 或 µ-law PCM 信号是如何被转化为40, 32, 24或16 kbps 的 ADPCM 通道的。在这些通道中，24和16 kbps 的通道被用于数字电路倍增设备(DCME)中的语音传输，而40 kbps 通道则被用于 DCME 中的数据解调信号（尤其是4800 kbps 或更高的调制解调器）。

G.726 encoder 输入一般都是G.711 encoder的输出：64kbps A-law or u-law.其算法实质就是一个ADPCM， 自适应量化算法。

## G.729

G.729编码方案是电话带宽的语音信号编码的标准，对输入语音性质的模拟信号用8kHz、采样，16比特线性PCM量化。G.729A是ITU最新推出的语音编码标准G.729的简化版本。G.729协议使用的算法是共轭结构的算术码本激励线性预测（CS-ACELP），它基于CELP编码模型。由于G.729编解码器具有很高的语音质量和很低的延时，被广泛地应用在数据通信的各个领域，如VoIP和H.323网上多媒体通信系统等。

　 电话线路上的模拟语音信号，经话路带宽滤波（符合ITU-T G.712建议）后，被8kHz采样，量化成16bit线性PCM数字信号输入到编码器。该编码器是基于线性预测分析合成技术，尽量减少实际语音与合成语音之间经听觉加权后差分信号的能量为准则来进行编码的。

编码器的其主要部分有：

　　·线性预测分析和LPC系数的量化；

　　·开环基音周期估计；

　　·自适应码本搜索；

　　·固定码本搜索；

　　·码本增益量化。

　　G.729的解码也是按帧进行的，主要是对符合G.729协议的码流进行解码，得到相应的参数，根据语音产生的机理，合成语音。解码过程主要分为：参数解码；后滤波处理。

　　·参数解码

　　首先解码得到线谱对参数，并将线谱对参数转换为线性预测系数。然后解码出基音周期，获得自适应码本矢量V(n)。解码出固定码本矢量的四个脉冲的位置和符号，计算出固定码本矢量c(n)。解码出固定本预测增益gc和固定码本增益，接着得到激励信号。最后将激励信号输入到线性预测合成滤波器中，计算出重构语音。

　　·后滤波处理

　　后滤波处理主要是自适应后滤波。自适应后置滤波器是由三个滤波器级连而成：长时后置滤波器Hp(z)，短时后置滤波器Hf(z)，频谱倾斜补偿滤波器Ht(z)，后面接着进行一个自适应增益控制过程。后置 滤波器的系数每一个子帧更新一次。后置滤波能够有效地改善合成出的语音质量。