



北邮信息理论  
与技术教研中心

BUPPT Information Theory & Technology Center

## 第四章

# 信源编码与数据压缩



主讲人: 牛 凯

[niukai@bupt.edu.cn](mailto:niukai@bupt.edu.cn)



- 本章讨论在物理层决定有效性的最主要因素：信源编码和数据压缩技术。
- 信源编码是主要利用信源的统计特性，解除信源相关性，去掉信源冗余信息，从而达到压缩信源输出的信息率，提高系统有效性的目的。
- 第二代移动通信主要是语音业务，所以信源编码主要指语音压缩编码。第三代移动通信中的信源编码将不仅包含语音压缩编码，还包含各类图像压缩编码和多媒体数据压缩等方面内容。

## § 4.1 语音压缩编码



- 在本节中，我们将讨论语音压缩编码的基本原理与方法，以及在移动通信中的语音编码。



## 4.1.1 引言

语音压缩编码大致可以分为以下三类：

- 波形编码
- 参量编码
- 混和编码

以上三类编码，波形编码质量最高，其质量几乎与压缩处理之前相同，可以适用于公用骨干(固定)通信网。参量编码质量最差，不能用于骨干通信网，而仅适合于特殊通信系统，比如军事与保密通信系统。混和编码质量介于两者之间，目前主要用于移动通信网。



# 1. 波形编码的性能估计

利用信息论中连续(模拟)有记忆信源的信息率失真 $R(D)$ 函数理论可以分析波形编码的性能。

信息率失真 $R(D)$ 为:

$$R(D) = \frac{1}{2} \log_2 \frac{\sigma^2(1-\rho^2)}{D}$$

上式的计算结果如下表所示:

信噪比(dB)	35	32	28	25	23	20	17
$R(D)$ (bit/样点)	4	3.5	2.5	2.34	2	1.5	1
压缩倍数K	2	2.28	3.2	3.42	4	5.3	8



- 由上述分析结果可以得到如下结论：当语音质量达到进入公网要求标准时，即  $\frac{\sigma^2}{D} \approx 26dB$  ,其  $K \approx 3.4$  倍，若进一步考虑实际语音分布与主观因素的影响(因为正态分布  $R(D)$  其压缩倍数可以进一步增大，取  $K = 4$  (保守值)这时语音速率可以从未压缩的PCM 64Kbps降至1/4速率的16Kbps。目前已实用化的DPCM为32Kbps。



## 2. 参量编码的性能估计

语音可以采用各种不同形式的参量来表达。为了分析方便，采用最基本的参量“音素”。以英语音素为例进行分析。英语中共有音素 $2^7 = 128 \sim 2^8 = 256$  按照通常讲话速率，每秒大约平均发送10个音素。由信息量计算公式，对于等概率事件有 $I = \log_2 N$ ， $N$ 为总组合数，则：

$$I_1(\text{上限}) = \log_2 N = \log_2 (256)^{10} = 80\text{bps}$$

$$I_2(\text{下限}) = \log_2 N = \log_2 (128)^{10} = 70\text{bps}$$

最后可计算出压缩比 $K$ 为：

$$K = \frac{64\text{Kbps}}{70 \sim 80\text{bps}} \approx 914 \sim 800\text{倍}$$



### 3. 混合编码的性能估计

显然混合编码的理论压缩比是介于上述两类编码之间，且与语音质量需求有关。若要求混合编码偏重于个性特征，则其压缩比靠近波形编码的压缩比值，若要求混和编码偏重于共性，则其压缩比靠近于参量编码。



## 4.1.2 数字通信中的语音编码



- 高质量的混合编码是移动通信中的优选方案。
- 在低数据比特率、高压缩比的混合编码中，数据数据比特率、语音质量、算法复杂度与处理时延是四个主要参量。



# 1. 数据比特率(bps)

- 数据比特率越低压缩倍数就越大，可通信的话路数也就越多，移动通信系统也就越有效。
- 数据比特率降低，语音质量也随之相应降低，为了补偿质量的下降，可采用提高设备硬件复杂度和算法软件复杂度的办法。
- 降低比特速率另一种有效方法是采用可变速率的自适应传输，它可以大大降低语音的平均传送率。
- 还可以进一步采用语音激活技术，充分利用至少3/8的有效空隙，可获得大致约2.67dB的有效增益。



## 2. 语音质量

度量方法不外乎客观与主观两个角度：

- 客观度量可以采用信噪比、误码率、误帧率，相对而言简单、可行。
- 主观度量是由人耳主观特性来判断，比客观度量复杂。目前国际上常采用的主观评判方法称为MOS方法。



### 3. 复杂度与处理时延

- 语音编码硬件复杂度取决于DSP处理能力，而软件复杂度则主要体现在算法复杂度上。算法复杂度增大，也会带来更长的运算时间和更大的处理时延。

- 如右所示，我们给出几种已知低数据比特率语音编码的上述四个参数与性能比较表格。

编码器类型 \ 参数	数据比特率 (Kbps)	复杂度 (MIPS)	时延 (ms)	质量 (MOS)
脉码调制PCM	64	0.01	0	4.3
自适应差分脉码调制ADPCM	32	0.1	0	4.1
自适应自带编码	16	1	25	4
多脉冲线性预测编码	8	10	35	3.5
随机激励线性预测编码	4	100	35	3.5
线性预测声码器	2	1	35	3.1

## § 4.2 移动通信中的语音编码

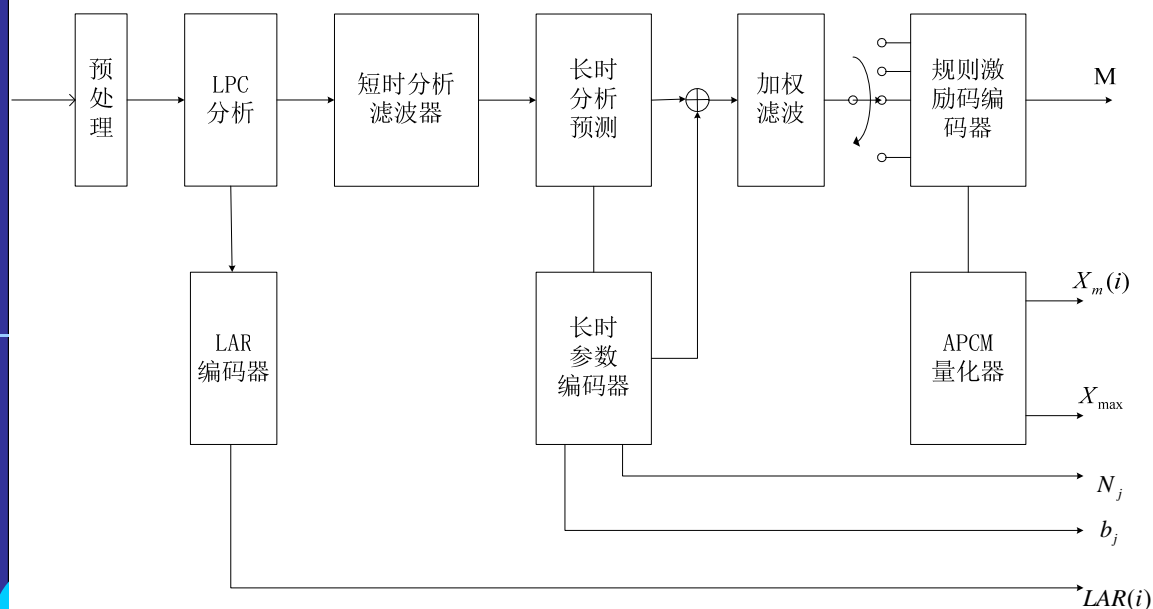


- 本节将结合第二代(2G)的GSM与IS-95系统以及第三代(3G)的WCDMA和CDMA2000等不同系统所采用的语音编码具体方案,着重从原理上来阐述移动通信中的语音编码。

## 4.2.1 GSM系统的RPE-LTP声码器原理



- RPE-LTP声码器采用等间隔，相位与幅度优化的规则脉冲作为激励源，以便使合成后的波形更接近原始信号。该方案结合长期预测以消除信号的冗余度，降低编码速率，同时其算法较简单，计算量适中且易于硬件实现。

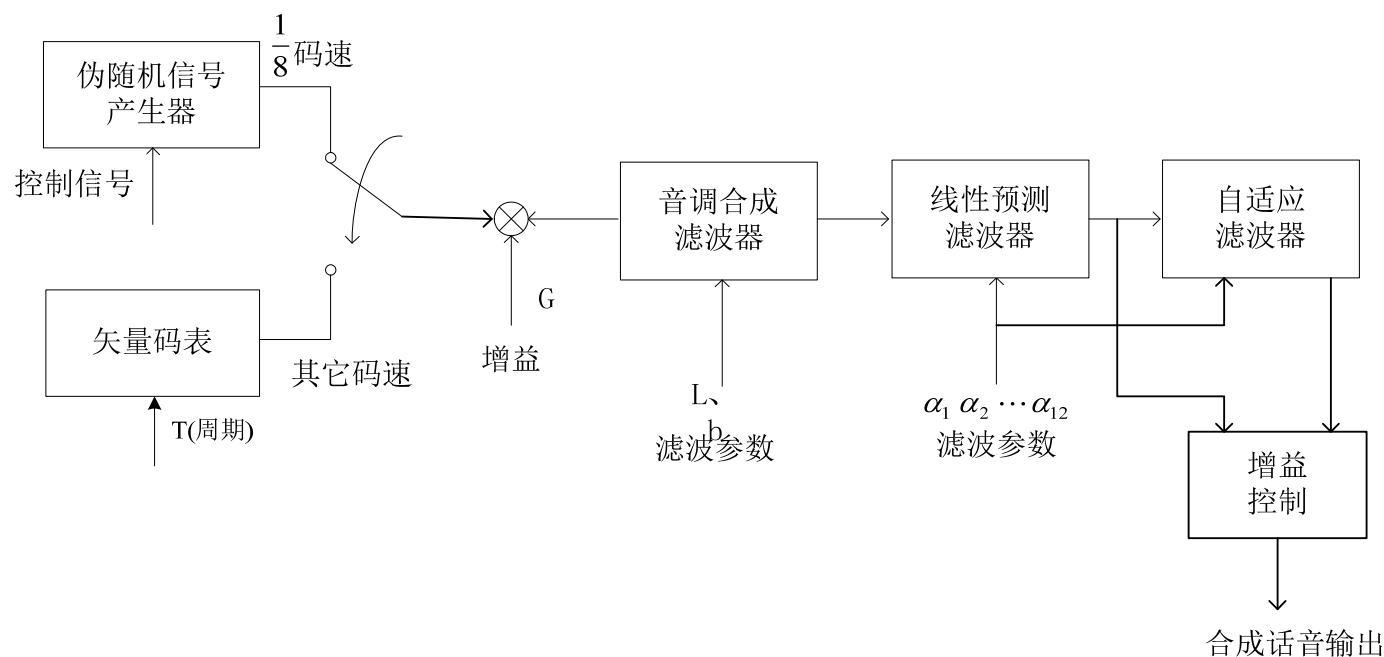


- REP-LTP编码器包括下列五个部分：预处理，线性预测分析，短时分析滤波，长时预测以及规则脉冲激励编码，其编码器原理如左图所示。

## 4.2.2 IS-96系统的QCELP声码器

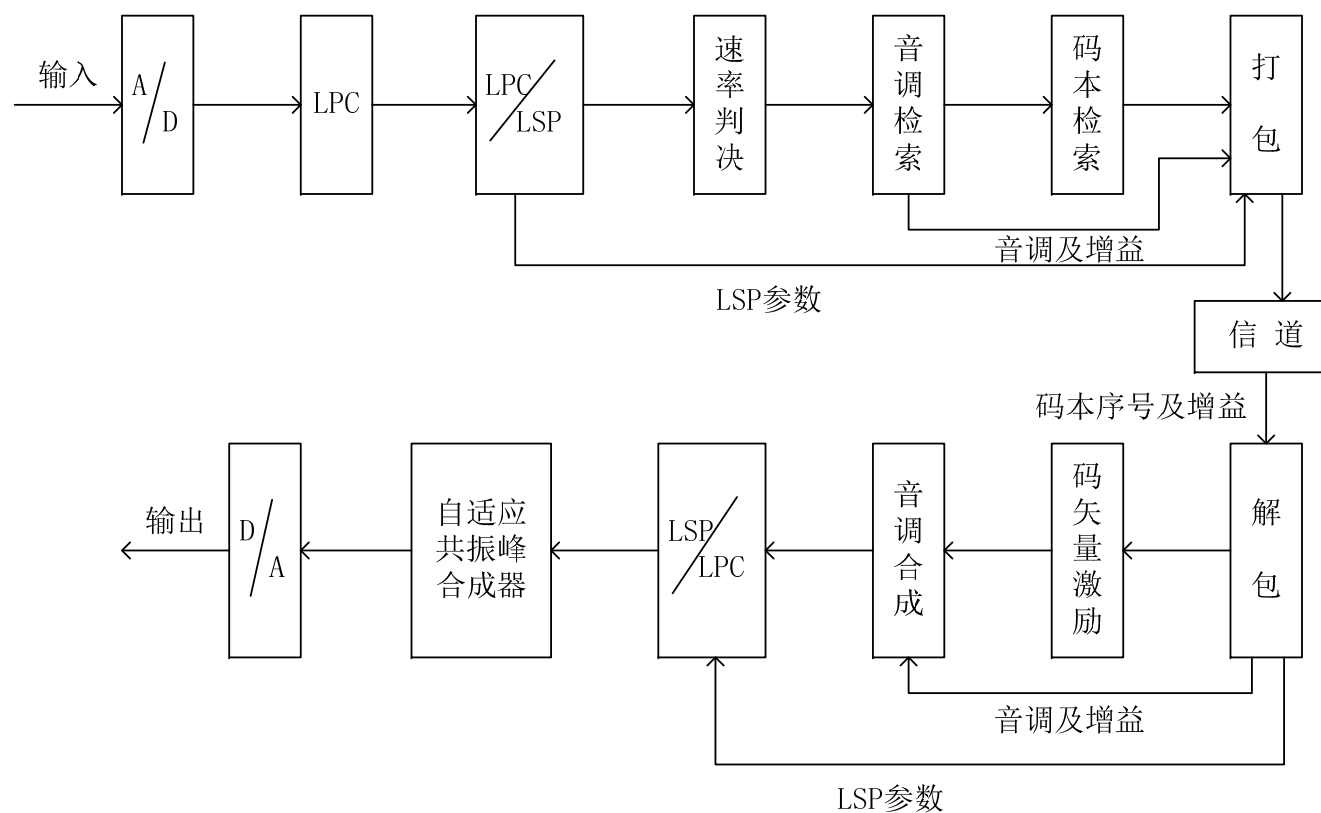


- QCELP声码器是Qualcomm公司提出的用于IS-96系统的语音编码标准。
- QCELP方案的编码原理如下图所示。





- TIA/EIA IS-96的QCELP语音编译码系统如下图所示：





### 4.2.3 CDMA2000系统的EVRC声码器

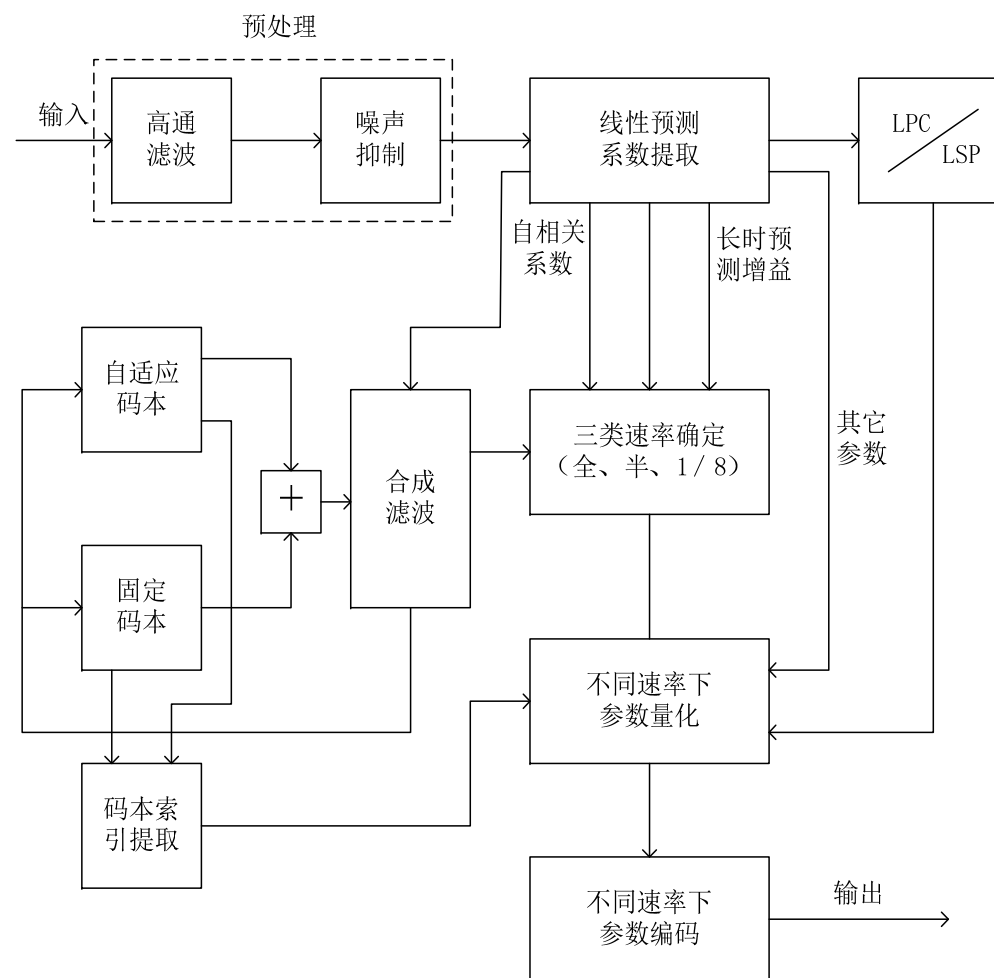


- EVRC(Enhanced Variable Rate Codec)即增强型可变速率语音编码器，是由美国电信工业协会TIA/EIA于1996年提出的CDMA2000系统的语音编码方案。
- EVRC编码器基于码激励线性预测，与传统CELP算法主要区别为：它能基于语音能量、背景噪声和其它语音特性动态调整编码速率。



- EVRC编码器结构如右所示，具体由以下几部分组成：

- 高通滤波器
- 线性预测器的参数提取模块
- 速率确定模块
- 参数量化模块
- 参数编码模块



## 4.2.4 WCDMA系统中的AMR声码器

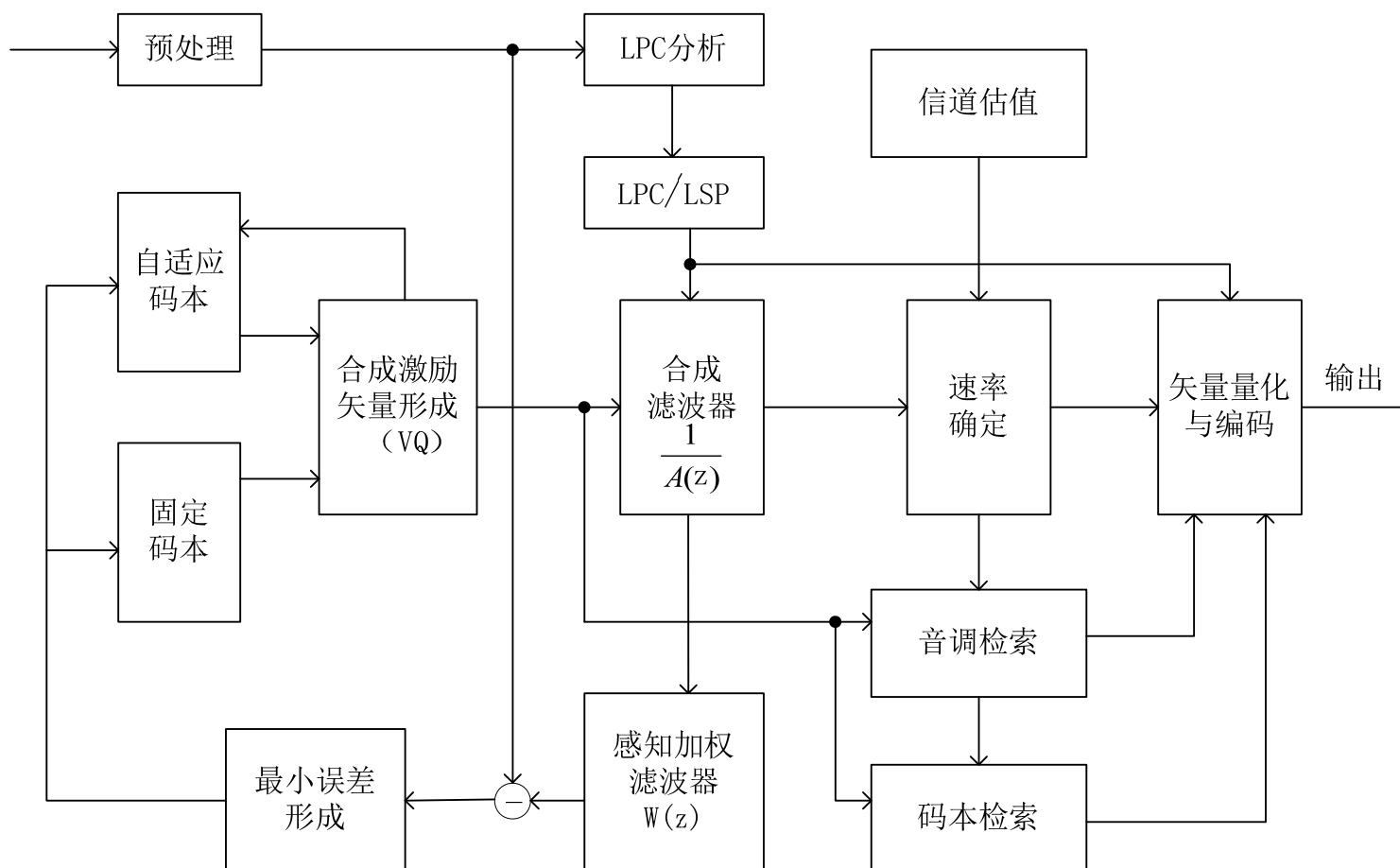


- AMR是第三代移动通信中WCDMA优选的语音编码方案，其基本思路是联合自适应调整信源和信道编码模式来适应当前信道条件与业务量大小。
- AMR编码自适应有两个方面：信源和信道。对于信道存在两类选择：全速率(FR): 22.8Kbps和半速率(HR): 11.4Kbps，而对于FR和HR不同信道模式分别有8种和6种信源编码速率。

## 4.2.4 WCDMA系统中的AMR声码器



- AMR编码器结构如下图所示:



## § 4.3 图像压缩编码



- 在第一、二代移动通信中主要是语音业务，从2.5G开始就逐步引入数据业务，第三代业务推广为含语音、数据与图像的多媒体业务。为了适应第三代业务的需求，本节介绍图像压缩编码。



## 4.3.1 图像编码标准简介

- 图像的信息量远大于语音、文字、传真和一般数据，它所占用的频带也比其它类型业务宽。经过四十余年的努力，图像编码已形成了如右所示的系列化标准。

标准	压缩比与数据比特率	应用范围
JPEG	2~30倍	有灰度级的多值静止图片
JPEG-2000	2~50倍	移动通信中静止图片、数字照相与打印、电子商务
H.261	$p \times 64\text{Kbps}$ , 其中 $p=1,2,\dots,30$	ISDN视频会议
H.263	8Kbps~1.5Mbps	POTS视频电话、桌面视频电话、移动视频电话
MPEG-1	不超过1.5Mbps	VCD、光盘存贮、视频监控、消费视频
MPEG-2	1.5Mbps~35Mbps	数字电视、有线电视、卫星电视、视频存贮、HDTV
MPEG-4	8Kbps~35Mbps	交互式视频、因特网、移动视频、2D/3D计算机图形



- 目前制定视频压缩编译码国际标准的有两大国际组织：一个是ITU-T(以前称CCITT)，即国际电联的电信标准部，它制定的标准通常称为建议标准，一般用H.26X表示。另一个是ISO/IEC，即国际标准化组织和国际电工委员会，它所制定的一般就称为标准。通常采用JPEG和MPEGX表示。
- 目前，视频压缩编码大致可以分为两代，第一代视频压缩编码包括JPEG、MPEG-1、MPEG-2、H.261、H.263等等；第二代视频压缩编码包括JPEG-2000、MPEG-4、MPEG-7、H.264等等。



## 4.3.2 静止图像压缩标准JPEG

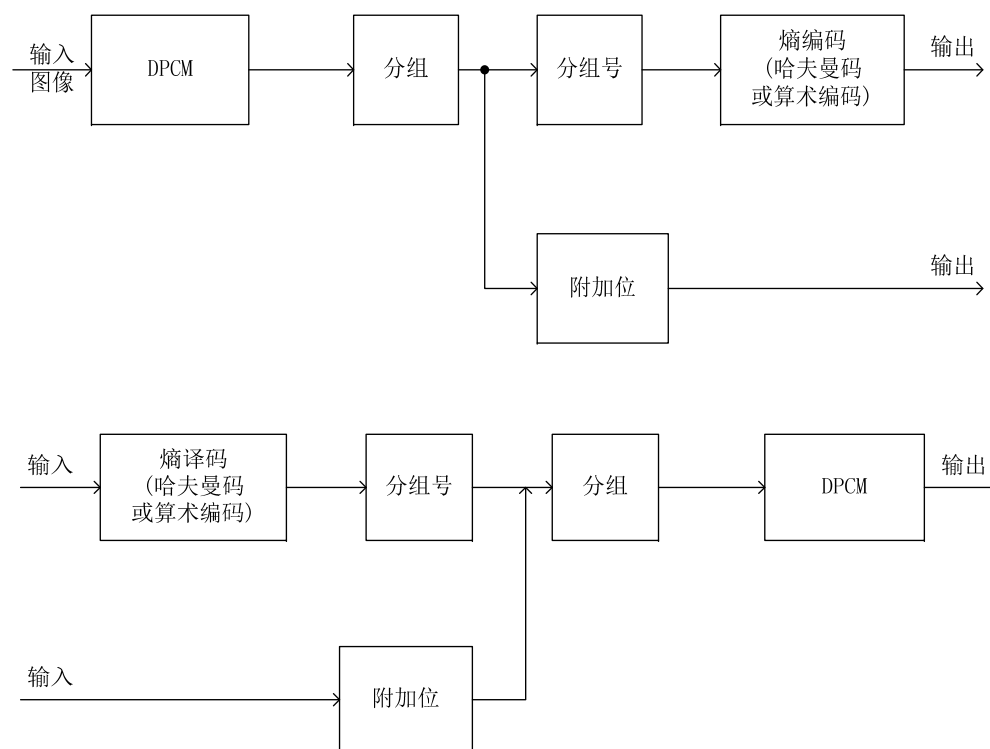
- 对于静止图像，国际标准化组织ISO和原来的国际电报电话咨询委员会CCITT(现改名为ITU-T)以及国际电工委员会IEC共同组织了一个图片专家联合小组(Joint Photographic Experts Group)研究制定标准，称它为JPEG标准。
- JPEG标准分为两类：基于DPCM与熵编码的无失真编码系统；基于离散余弦变换DCT的限失真编码系统。





# 1.基于DPCM的无失真编码

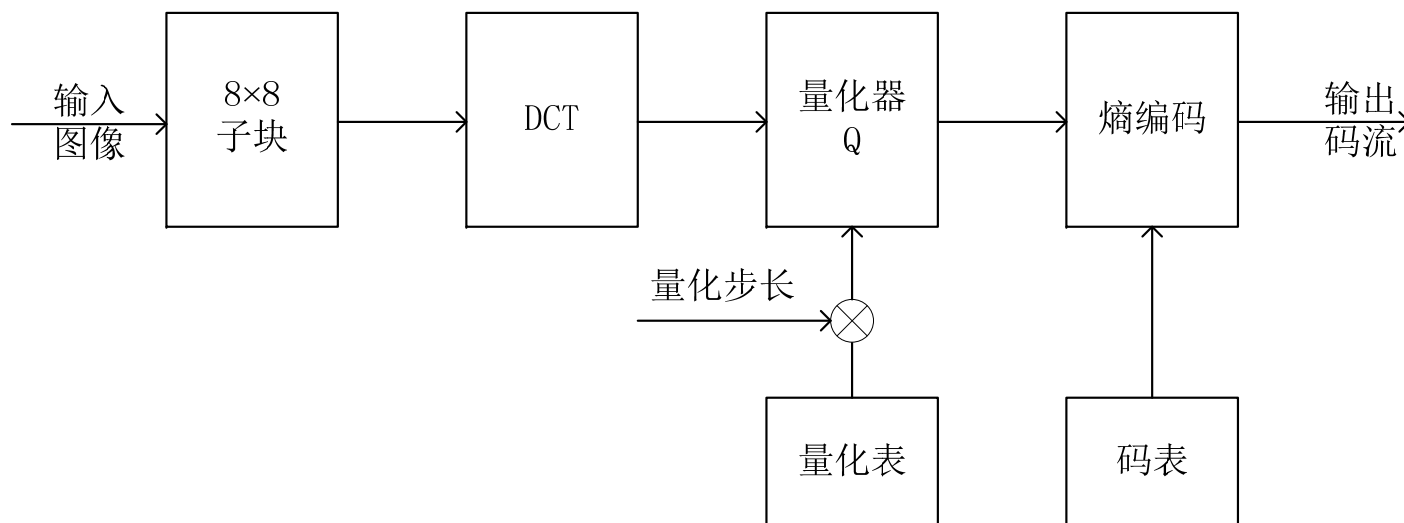
- 无失真编码又称为无损信源编码，它是一种不产生信息损失的编码，一般其压缩倍数比较低，为4倍左右。
- JPEG无失真编码的发送与接收系统实现原理如右图所示。
- 无失真JPEG编码从原理上看，主要是以DPCM为基础，再加上哈夫曼编码或者算术编码的熵编码方式。





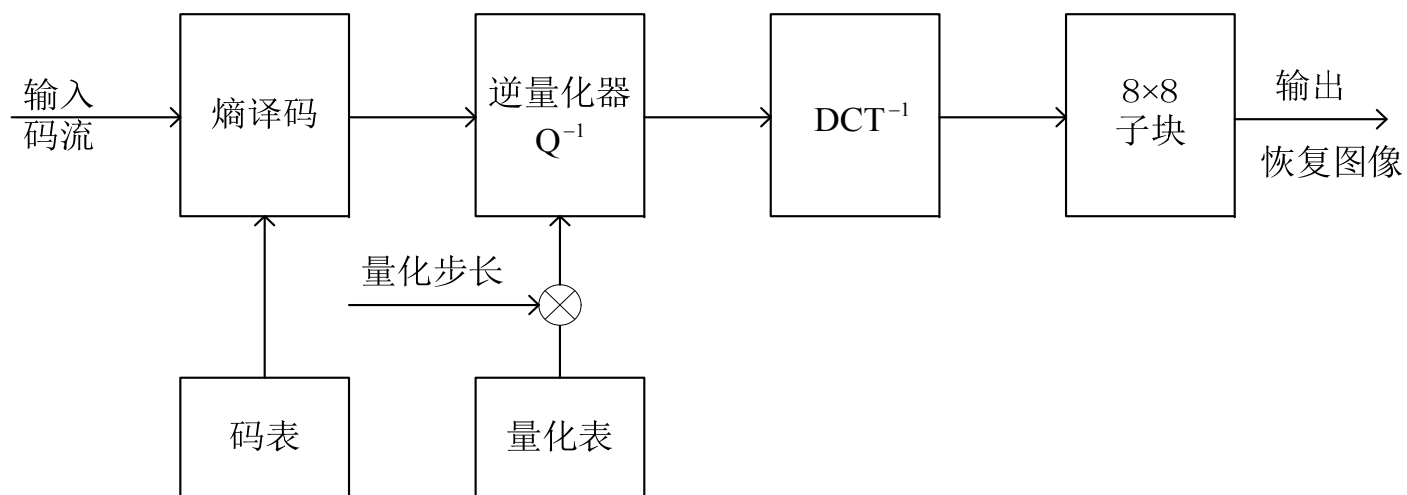
## 2. 基于离散余弦变换DCT的限失真编码

- 限失真编码属于有损信源编码，以离散余弦变换DCT为基础，再加上限失真量化编码和熵编码，它能够以较少的比特数获得较好的图像质量。
- 限失真JPEG编码器原理如下所示：





- 限失真JPEG译码器原理如下所示：



### 4.3.3 准活动图像视频压缩标准H.26X

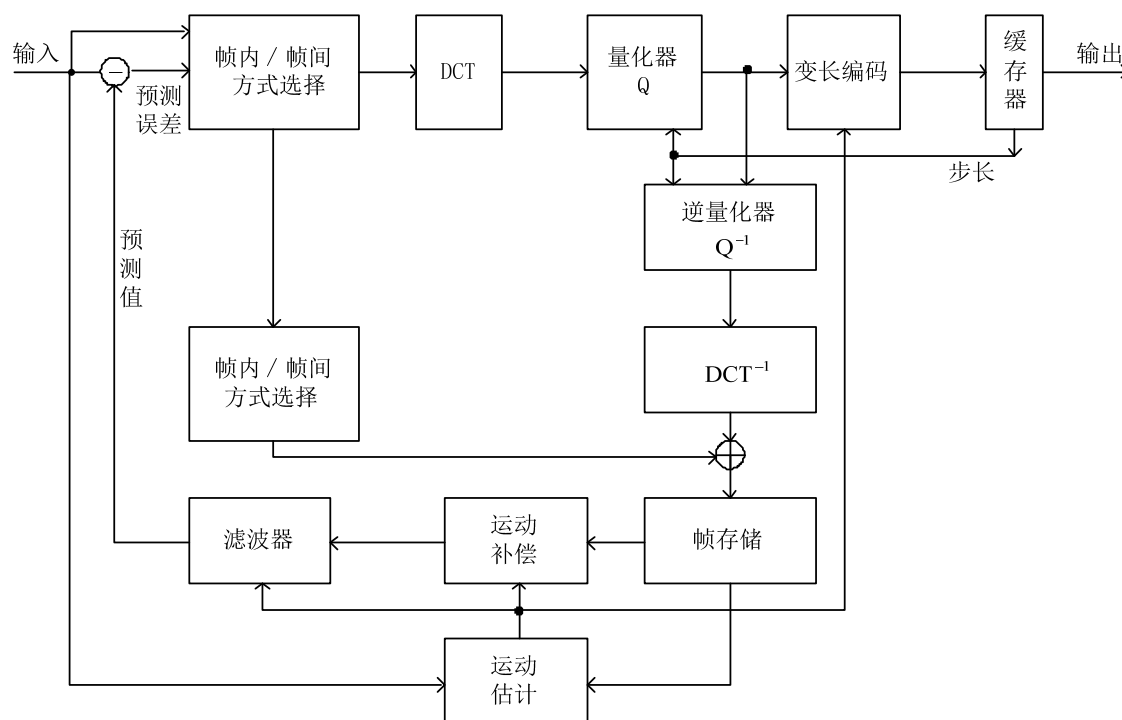


- 编码标准H.26X是由ITU-T制定的建议标准，自上世纪八十年代中期开始现已制定了H.261、H.262、H.263，最近刚完成了H.264建议标准。其中H.262和MPEG-2视频编、译码标准是同一个标准，这是两大国际组织的共同成果。而H.264也是两大组织联手制定的，被称为“MPEG-4 Visual Part 10”，也就是“MPEG-4 AVC(Advanced Video Coding)”，2003年3月被正式确定为国际标准。



# 1. H.261编码标准简介

- H.261主要用于传输会议电话及可视电话信号，它将码率确定为： $p \times 64\text{Kbps}$ ，其中 $p=1,2,\dots,30$ 。其对应的数据比特率为 $64\text{Kbps} \sim 1.92\text{Mbps}$ 。



- H.261编码器原理图如左所示。





## 2. H.263编码标准简介

- H.263系列适合于PSTN、无线网络和因特网。  
H.263信源编码算法的核心仍然是H.261标准中所采用的编码算法，其原理框图也与H.261基本上一样。
- H.263与H.261的区别如下：
  - H.261只能工作于CIF与QCIF两类格式，而H.263则可工作于5种格式：CIF、QCIF、SubQCIF、4CIF、16CIF
  - H.263吸收了MPEG等标准中有效、合理的部分
  - H.263在H.261基本编码算法基础上又提供了四种可选模式，以进一步提高编码效率

## 4.3.4 活动图像视频压缩标准MPEG



- 这类标准是由国际标准化组织ISO和国际电工委员会于1998年成立的一个研究活动图像的专家组MPEG(Moving Picture Experts Group)负责制定的。现已制定了MPEG-1、MPEG-2、MPEG-4以及补充标准MPEG-7与MPEG-21等，其中MPEG-2与MPEG-4是与ITU-T联合研制的。
- 在MPEG系列标准中，MPEG-1、MPEG-2属于第一代视频压缩标准，而MPEG-4则属于第二代视频压缩标准。

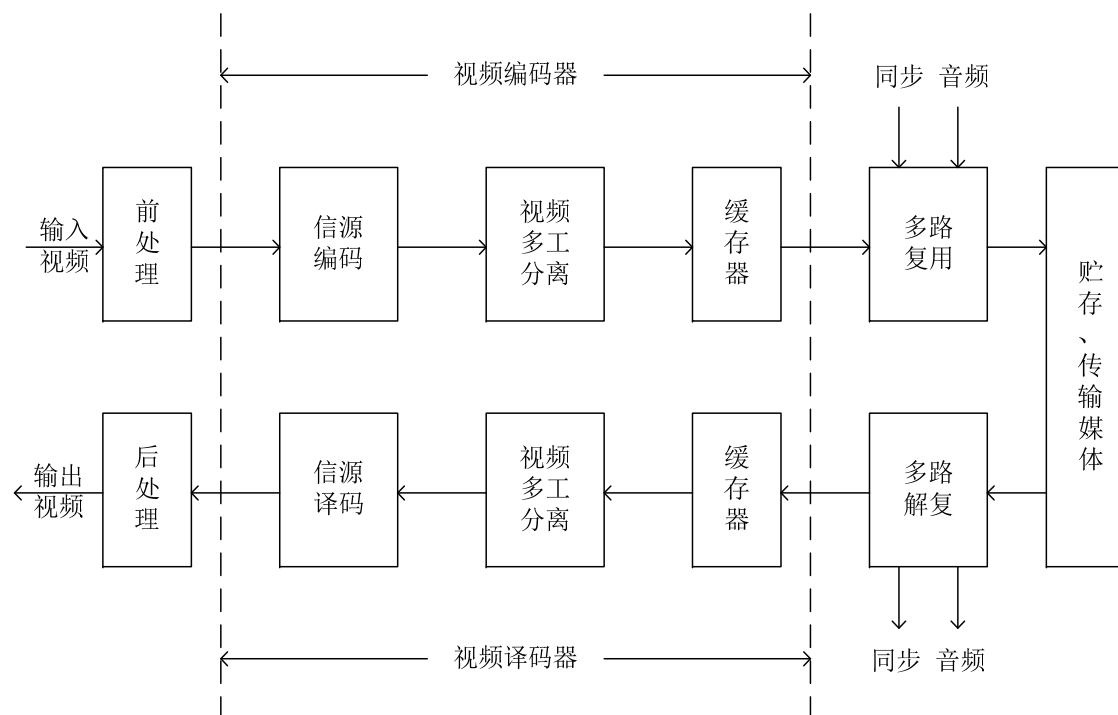




# 1. MPEG-1编码标准简介

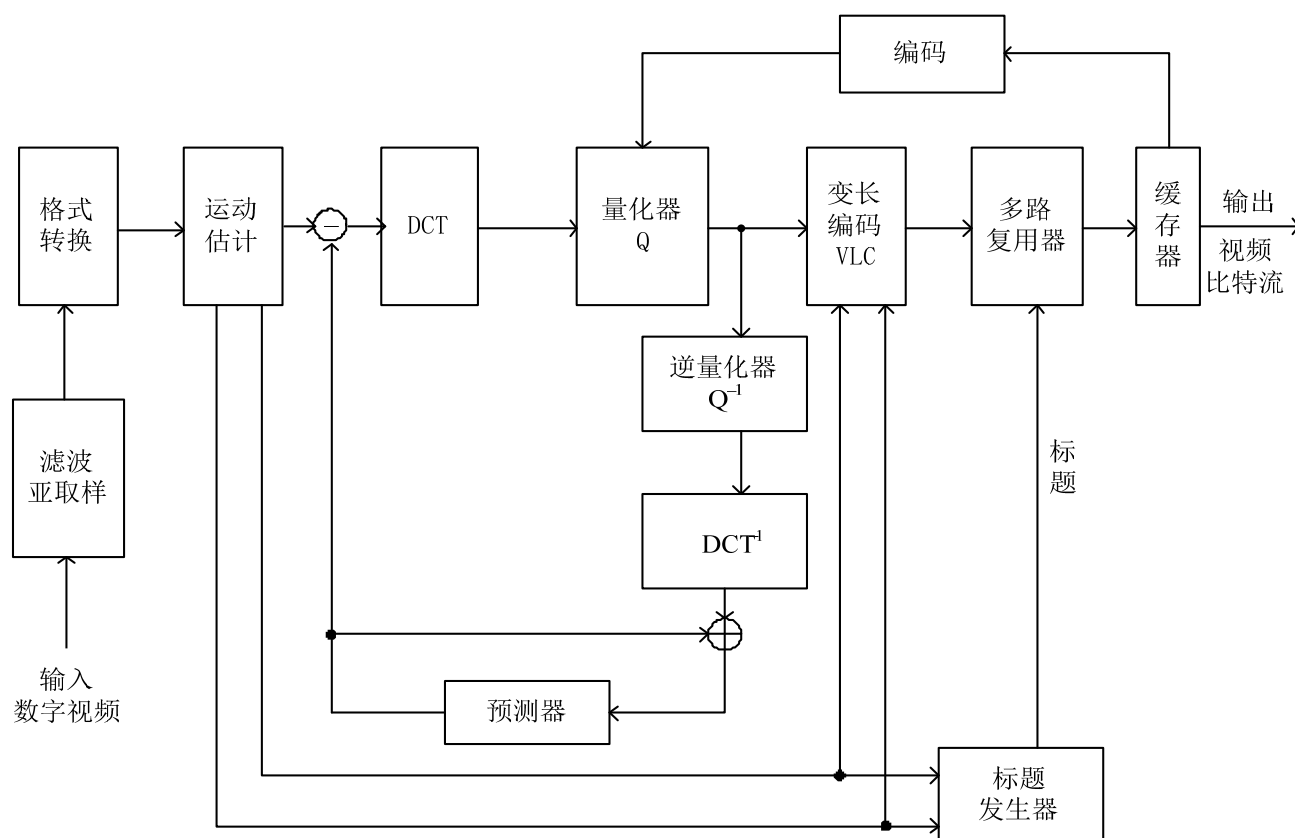
- MPEG-1主要是针对1.5Mbps速率的数字存贮媒体运动图像及其伴音制定的国际标准，用于CD-ROM的数字视频以及MP3等。

- MPEG-1视频编译码系统的原理性框图如右所示。



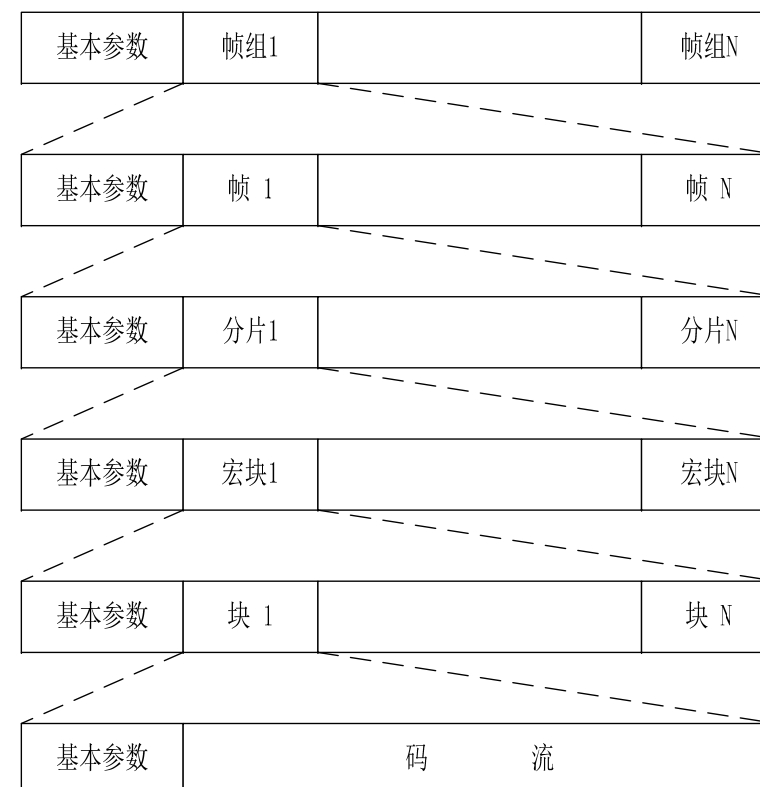


- 上述框图中，核心部件是视频编解码器，视频编码器的结构如下图所示。





分层名称	功能	序列层
块层	进行离散余弦变换DCT的基本单元	帧组层
宏块层	预测单元	帧层
分片层	同步恢复单元	分片层
帧(图片)层	基本编码单元	宏块层
帧组(图片组)层	视频随机存取单元	块层
视频序列层	节目内容随机存取单元	



- MPEG-1视频流采用分层式数据结构，其分层方法及其功能如上表所示。
- MPEG-1视频流分层结构图如上所示。



- MPEG-1中视频图像分成四种帧类型：I、P、B与D帧。
  - I帧为帧内编码帧(Intracoded frame)，编码时采用类似H.261的DCT编码
  - P帧为预测编码帧(Predictively coded frame)，采用前向运动补偿预测和误差的DCT编码，由其前面的I帧或P帧进行预测
  - B帧，为双向预测编码帧(bidirectionally predictively coded frame)，采用双向运动补偿预测和误差DCT编码
  - D帧为直流编码器(DC Coded frame)，它只包含每个块的直流分量



## 2. MPEG-2编码标准简介

- ISO/IEC的MPEG组织于1995年推出MPEG-2标准，它是主要针对数字视频广播、高清晰度电视HDTV和数字视盘等制定的4~9Mbps运动图像及其伴音的编码标准。
- MPEG-2与MPEG-1的差异如下：
  - MPEG-2专门设置了“按帧编码”和“按场编码”两类模式，并相应地对运动补偿和DCT方法进行了扩展
  - MPEG-2压缩编码在一些方面进行了扩展
  - 空间分辨率、时间分辨率、信噪比可分为不同等级以适合不同等级用途需求，并可给予不同等级优先级
  - 视频流结构具有可分级性
  - 输出码率可以是恒定也可以是变化的以适应同步与异步传输



- MPEG-2视频是一个多格式系统，允许对4种源格式进行5种类型11种单独技术规范进行编码
- 11种单独技术规范如右表格所示

等级	简单规范 (无B帧，不可缩放)	主规范 (B帧，不可缩放)	SNR缩放(B帧，SNR缩放)	空间可缩放的规范(B帧，空间或SNR可缩放)	高级规范(B帧，空间或SNR可缩放)
高层1 1920×1 152×60		80Mbps			100Mbps
高层2 1440×1 152×60		60Mbps		60Mbps	80Mbps
高层3 720×57 6×30	15Mbps	15Mbps	15Mbps		20Mbps
低层 352×28 8×30		4Mbps	4Mbps		

## 4.3.5 第二代视频压缩编码标准



- 本小节介绍三类代表性标准，即已应用于移动通信的JPEG-2000、MPEG-4编码标准和H.264编码标准。



# 1. JPEG-2000编码标准简介

JPEG-2000主要特点如下：

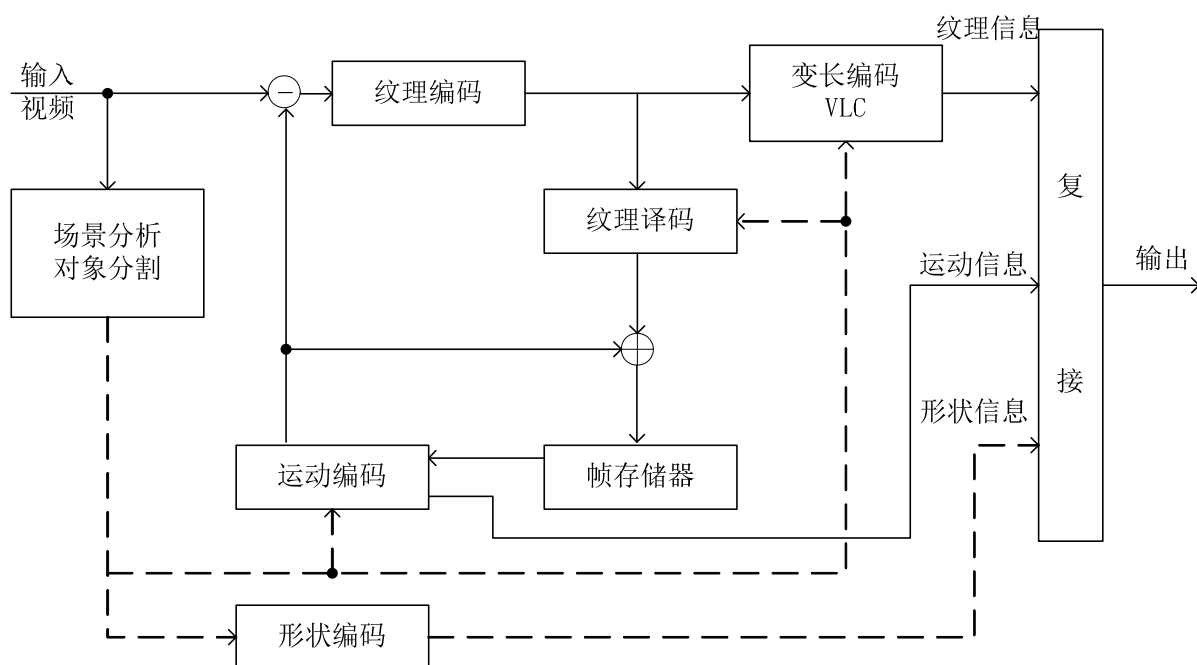
- 用以小波变换为主的多分辨率编码方式代替JPEG中采用的传统DCT变换
- 采用了渐进传输技术(progressive transmission)
- 用户在处理图像时可以指定感兴趣区域(ROI, region of interest), 对这些区域可以选取特定的压缩质量和解压缩质量
- 利用预测法可以实现无损压缩(lossless compression)
- 具有误码鲁棒性(Robustness to bit error), 抗干扰性好
- 考虑了人眼的主观视觉特性, 增加了视觉权重和掩膜





## 2. MPEG-4编码标准简介

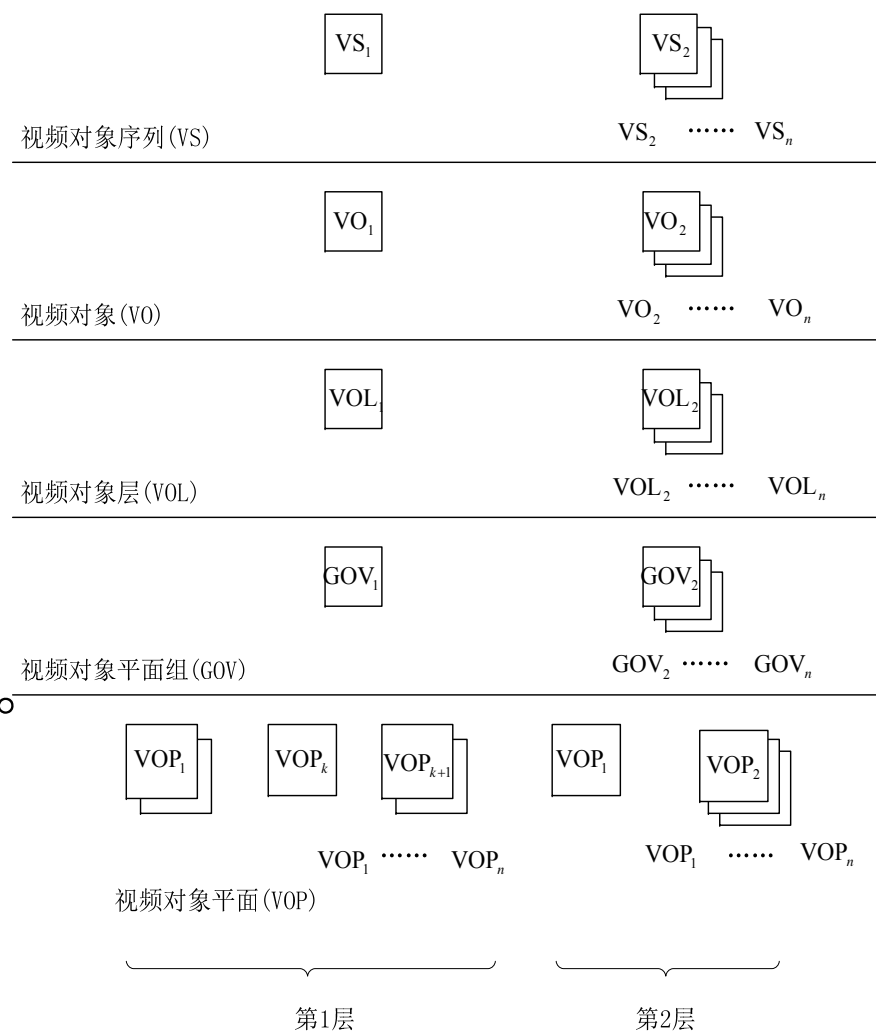
- 视频编码大体上可以分为两代：第一代基于像素的方法；第二代基于内容的方法。MPEG-4是基于对象方法。



- 左图给出了对于一个任意形状的视频对象进行通用编码的原理框架。主要包含三部分：纹理、形状和运动三个编码模块。

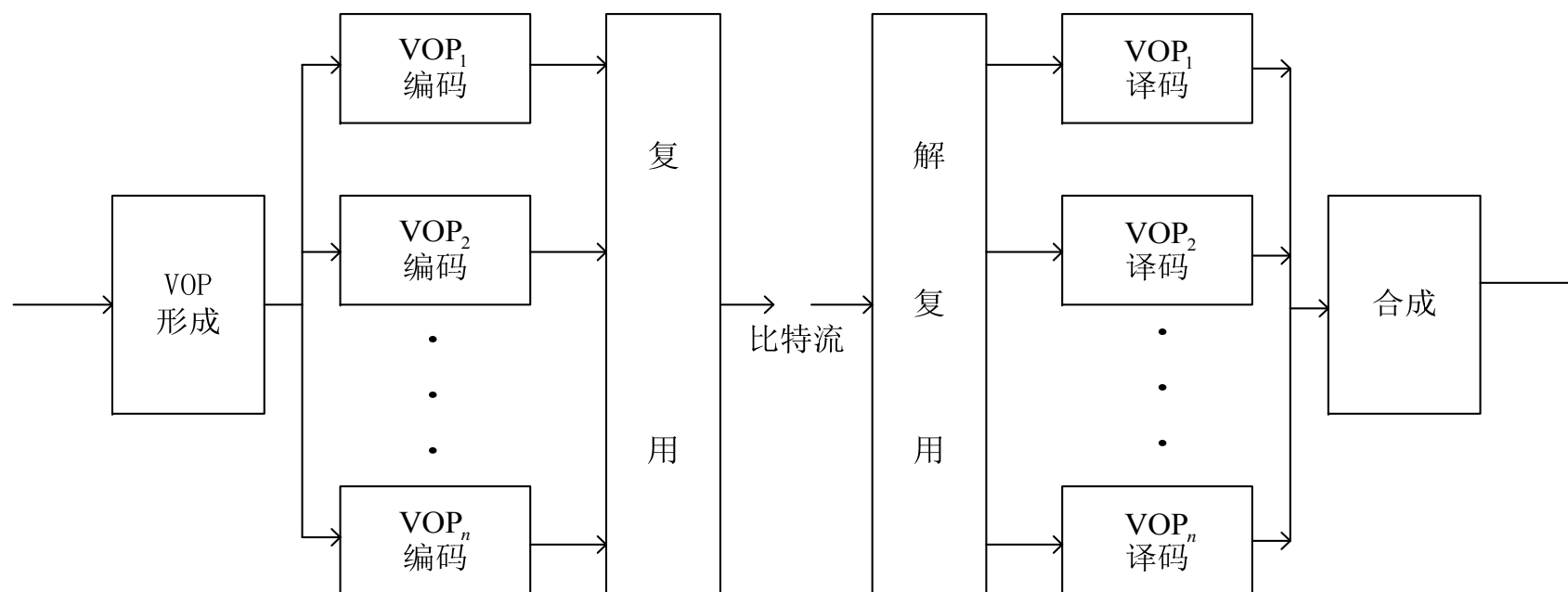


- MPEG-4标准中定义的中心概念是AV对象，其编码机制是基于 $16 \times 16$ 的像素宏块来设计的。
- MPEG-4视频码流提供了对视频场景的分层描述，如右图所示。





- 下面给出一个用VOP形成模块在MPEG-4视频验证模型VM中的框架构成原理：





- MPEG-4视频编译码的主要特点包括：
  - 图像信息处理的基本单元，由第一代像素块像素帧转变到以纹理、形状和运动三类主要数据的取样值构成视频对象平面VOPi
  - 视频编码基础转变成既取决于原有的客观统计特性，而更重要的则是取决于视频对象、内容的各种主、客观以及图像瞬时特性
  - 基于对象、基于内容
  - 对于不同的信源与信道，以及各个VO以及VOPi在总体图像中的重要性和地位，可以分别采用不同等级的保护与容错措施
  - 图像处理中具有时间、空间可伸缩性(尺度变换)



### 3. 新一代的视频编、译码标准H.264

- ITU-T与ISO/IEC联手成立的VCEG在H.263及其改进型与MPEG-4的基础上进行技术融合、改进和优化，共同提出H-264建议标准
- VCEG从图像质量与实时性两方面出发，给出以下几个方面要求：
  - 数据比特率与图像质量
  - 差错恢复
  - 时延
  - 语法定义
  - 复杂性
  - 网络友好性



- H.264与以往编码的主要差异有：
  - 运动估值和运动补偿
  - 采用内部(intra)预测
  - 采用系数变换技术
  - 采用变换系数量化
  - 熵编码
  - 在扫描顺序、去块滤波器、新的图片类型、熵编码模式和网络适应层等方向，都有与以往编码不一样的自己特色



## 4.4 我国音视频标准

- 数字音视频技术中，最核心的是音视频编解码算法和标准。目前，我国主要采用国外制定的标准，如：杜比AC-3、MPEG-2和H.264，使用这些标准需要向国外相关专利池缴付大量专利费，直接制约我国数字音视频产业的发展，也会影响移动通信业务的开展。
- 下面简要介绍近年来我国颁布的DRA数字音频与AVS数字视频国家标准。

## 4.4 我国音视频标准



- **4.4.1 DRA数字音频标准**
- DRA数字音频标准全称为《多声道数字音频编解码技术规范》(国标号: GB/T 22726-2008), 2009年2月, 由国家质量监督检验检疫总局和国家标准化管理委员会正式颁布。
- DRA是支持立体声和多声道环绕声的数字音频编解码, 最多可以支持64个正常声道和3个低频效果声道, 拥有压缩效率高, 音质好, 解码复杂度低和容错能力强的优点。



## 4.4 我国音视频标准



表4.6 DRA标准性能指标

采样速率范围	8-192kHz
量化精度	24bit
编码比特率	32-2304kbps
可支持声道数	正常声道64个，低频声道(LFE)3个
支持编码模式	VBR、CBR、ABR
音频帧长	1024采样点
算法复杂度	对于5.1声道384kbps码率数据，编解码复杂度为48MIPS。
压缩效果	128kbps立体声，MOS评分4.7；384kbps 5.1环绕声，MOS评分4.9；达到EBU(欧洲广播联盟)定义的“不能识别损伤”音频质量

## 4.4 我国音视频标准



- DRA 音频压缩算法基于人耳的听觉特性对声音信号进行量化和比特分配，属于感知音频编码，采用了变换编码的路线。
- DRA 算法实现了量化与熵编码独立优化，进一步提高了量化与熵编码的性能，采用了基于人耳听觉模型的自适应分块标量量化，并对量化指数进行了 Huffman 编码，主要技术特色在于可变分辨率滤波器组和熵编码。

## 4.4 我国音视频标准



- (1)可变分辨率滤波
- 准稳态片断具有高频域分辨率，而对瞬态信号具有高时域分辨率。传统算法往往采取折衷方法，但对于稳态信号和瞬态信号都不是最优。
- DRA算法采取了改进方法，对音频帧的瞬态信号范围进行分析，将稳态和瞬态信号分别处理。DRA对于稳态信号采用了高频域分辨率的滤波器组，使变换后的子带样本能量更加集中，有利于量化和熵编码；而对于瞬态信号则引入新的“瞬态窗函数”，提供了精细的时域分辨率，从而保留了足够听觉有效信息。

## 4.4 我国音视频标准



- (2)量化比特分配及熵编码
- 与同类音频编码器类似，DRA也采用心理学模型输出的量化掩蔽阈值分配量化噪音，使其尽可能被遮蔽而不被感知。
- 在对量化指数的熵编码中，根据每个量化指数的特性分配最优的码本指数，然后合并形成较大的段，共享一个码本指数，这样做可以尽量少的比特数传递码本向量信息。

## 4.4 我国音视频标准



- DRA已经被CMMB(China Mobile Multimedia Broadcasting)行业标准确立为必选音频标准, CMMB主要面向手机、PDA等小屏幕便携手持终端以及车载电视等终端提供广播电视服务。

## 4.4 我国音视频标准



- **4.4.2 AVS数字视频标准**
- AVS是我国自主制定，基于自主创新和部分公开技术的视频标准，技术方案简洁，实现复杂度低，是一套包含系统、视频、音频、媒体版权管理在内的完整标准体系。
- AVS通过简洁的一站式许可政策，解决了AVC专利许可问题死结，为数字音视频产业提供更全面的解决方案。

## 4.4 我国音视频标准



- AVS与MPEG-4、H.264标准具有相同的编码框架，但技术取舍的衡量指标各不相同，因而编码效率和复杂性也各有异同。
- 从编码效率来看，MPEG-4是MPEG-2的1.4倍，AVS和AVC相当，都是MPEG-2的两倍以上。从复杂度来看，H.264约为MPEG-2的9倍，AVS大致为MPEG-2的6倍。AVS的技术特点总结如下



## 4.4 我国音视频标准

- (1)自适应运动补偿
- 采用自适应宏块划分进行运动补偿是提高预测精度的重要手段之一。AVS将宏块划分最小限制为 $8 \times 8$ ，这一限制大大降低了编解码器的复杂度(30-40%)，而整体性能只降低大概在2-4%左右，达到较好折中。
- (2)帧内预测
- 与H.264类似，AVS也采用帧内预测技术，但基于 $8 \times 8$ 宏块进行，并且亮度只有5种预测模式，大大降低了预测模式选择的复杂度，且性能十分接近。





## 4.4 我国音视频标准

- **(3)多参考帧预测**
- 多参考帧预测使当前块能从前几帧图像寻找更好匹配。AVS限定最多采用两个参考帧，在不增大缓冲区的条件下提高了编码效率，实现方法比H.264更简洁。
- **(4)1/4像素插值**
- 运动矢量的精度是提高预测准确度的重要手段，其核心是插值滤波器的选择，AVS对1/2像素位置插值采用4阶滤波器可达到与H.264的6阶滤波器一致的性能。



## 4.4 我国音视频标准

- (5)整数变换
- AVS和H.264类似，都采用整数变换代替DCT变换，前者具有复杂度低、完全匹配等优点。AVS变换矩阵归一化在编码端完成，节省了译码缩放表。AVS采用了 $8 \times 8$ 、 $4 \times 4$ 两种变换，由于 $8 \times 8$ 变换比 $4 \times 4$ 变换的解相关能力更强，因此比只有 $4 \times 4$ 变换的H.264编码效率提高2%(约0.1dB)。
- (6)量化技术
- AVS与H.264都采用了量化与变换归一化结合的简化技术，由于变换归一化在编码端完成，因此译码器的反量化表与变换系数位置无关。



## 4.4 我国音视频标准

### •(7)B帧宏块模式

•AVS的B帧宏块采用空/时域结合的直接模式(Direct mode)以及运动矢量舍入控制技术，并且基于对称模式(Symmetric mode)在只编码一个运动矢量的条件下实现双向预测，比H.264中的B帧编码性能有所提高。

### •(8)熵编码

•AVS中的熵编码是H.264与MPEG-2熵编码的综合，既采用了MPEG-2的二维编码机制，又吸收了H.264利用上下文信息进行自适应编码的策略，从而将编码效率提高到与H.264相当，但却降低了复杂度。

### •(9)环路滤波

•由于AVS滤波点数、滤波强度分类数都比H.264少，因此大大减少了判断、计算的次数，从而降低了环路滤波的复杂度。

## 4.4 我国音视频标准



- H.264/AVC尽管具有良好的压缩效率，但其实现复杂度较高，很多技术都是各方利益妥协的产物，存在优化空间。AVS对视频编码的每项关键技术都进行了复杂性与效率的权衡，努力降低复杂度，并保证相近的编码效率，为3G/4G移动视频应用提供了很好的解决方案。

## 4.5 压缩感知

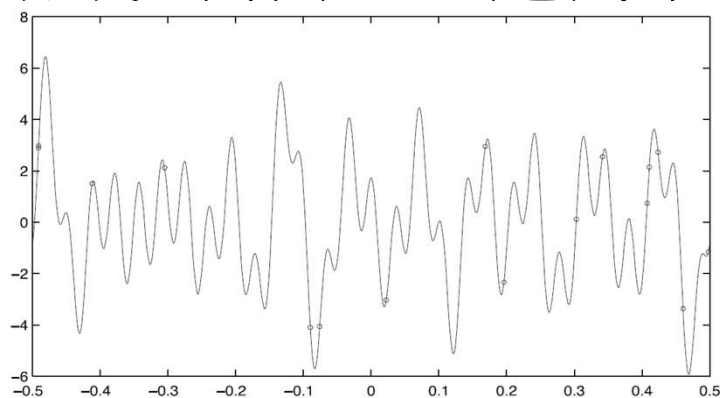


- 很多信号处理的对象是稀疏信号，即信号向量本身或信号向量在某个变换域仅有少数元素非零。
- 根据奈奎斯特采样定理，为了在接收端无失真恢复原始信号，传统信号处理方法要求对原始信号的采样率至少两倍于信号带宽。这样会增加处理复杂度和时延。
- 2006年，Candes、Romberg、陶哲轩[4.8]等人以及Donoho[4.9]提出了一项新的稀疏信号重建理论——压缩感知(Compressive Sensing)。
- 对稀疏信号进行压缩采样，利用较少的测量值重建信号，显著降低信号处理的复杂度与时延。目前已广泛应用于图像与视频处理领域。

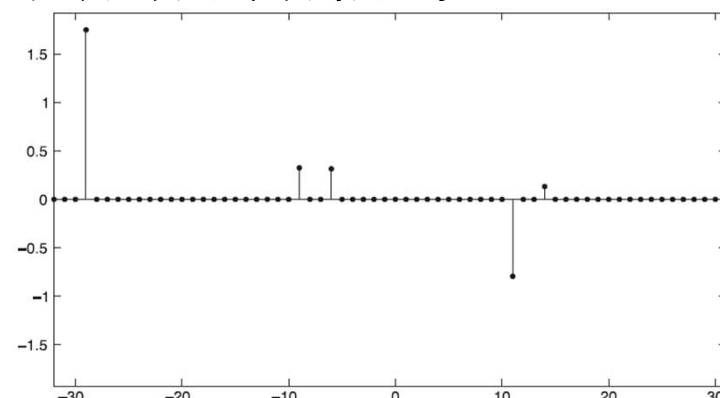


## 4.5 压缩感知

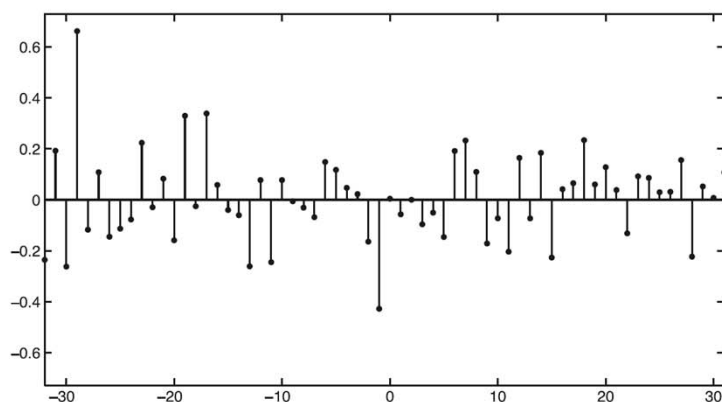
- 4.5.1 稀疏信号表示
- 所谓压缩感知，是指利用信号的稀疏性，从低维空间的观测向量重建高维空间的原始信号。



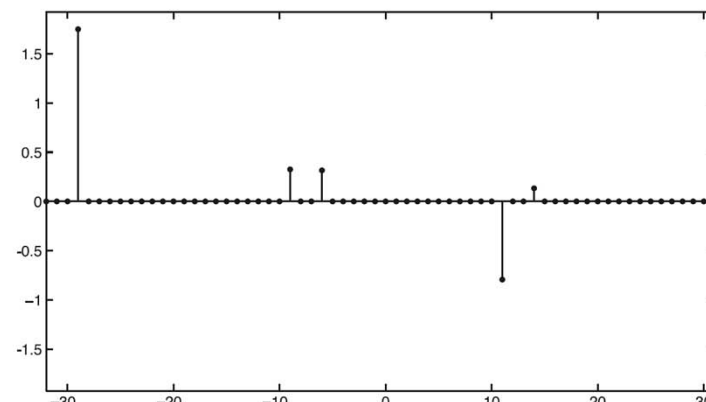
(a) 时域信号波形，经过 $M=16$ 个样点采样



(b) 经过 $N=64$ 点DFT变换后的信号频谱(稀疏度 $K=5$ )



(c) 采用2范数最小化(LS算法)重建的信号



(d) 采用1范数最小化重建的信号(精确恢复)



## 4.5 压缩感知

### • 1. 压缩感知模型

$$\mathbf{x} = \sum_{i=1}^N s_i \Psi_i = \Psi \mathbf{s}$$

- 对于可压缩信号，通常采用变换编码方法(例如DCT变换)进行处理，包括两步操作：
  - 首先以Nyquist速率采样所有信号样值；
  - 然后进行DCT变换，保留最大的K个元素，丢弃剩余元素
- 传统方法存在的问题
- 原始样值数量随着N 的增加而不断增大；
  - 即使大部分系数最终被丢弃而只保留K个系数，也需要执行完整的信号变换，因此存在计算与存储冗余；
  - 保留系数的序号需要编码，进一步提高了处理复杂度。



## 4.5 压缩感知



- 相比于传统的变换编码，压缩感知直接压缩信号而不需要采样操作。
- 给定观测矩阵 $\Phi$  ( $M \ll N$ )，观测向量 $\mathbf{y}$ 表示为

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{v} = \Phi \Psi \mathbf{s} + \mathbf{v} = \Theta \mathbf{s} + \mathbf{v}$$

- 问题转化为：
- (1) 设计合适的观测矩阵 $\Phi$ ，确保从观测向量恢复原始信号；
- (2) 设计信号重建算法，恢复信号



## 4.5 压缩感知



### • 2. 观测矩阵和观测维度约束

定理 4.1(RIP 特性): 如果信号  $\mathbf{x}$  是  $K$ -稀疏信号, 对于任意一个  $K$  稀疏的向量  $\mathbf{v}$  和任意的  $\varepsilon$ , 当且仅当观测矩阵  $\Theta$  满足如下条件:

$$1 - \varepsilon \leq \frac{\|\Theta \mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq 1 + \varepsilon \quad (4.5.4)$$

则  $M > K$  时压缩感知问题可以得到唯一解。

- 这个条件通常称作约束等距性(restricted isometry property, RIP), 它是决定信号重建成功与否的充要条件。
- 与RIP特性等价的判定条件是观测矩阵的非相关性, 即要求观测矩阵的任意一行不是稀疏基矩阵行向量的线性组合。

定理 4.2: 假设观测矩阵  $\Phi$  满足 RIP 特性, 则观测向量  $\mathbf{y}$  的维数  $M$  只需要满足  $M > C \cdot K \cdot \log(N / K)$ , 其中  $C \approx 0.28$ , 即可实现对原始信号的无失真恢复。

## 4.5 压缩感知



- 3. 优化模型

$$\begin{aligned} \mathcal{OP}1 : \hat{\mathbf{s}} &= \arg \min \|\mathbf{s}\|_2 \\ s.t. \mathbf{y} &= \mathbf{\Phi\Psi s} \end{aligned}$$

$$\begin{aligned} \mathcal{OP}2 : \hat{\mathbf{s}} &= \arg \min \|\mathbf{s}\|_0 \\ s.t. \mathbf{y} &= \mathbf{\Phi\Psi s} \end{aligned}$$

等价

$$\begin{aligned} \mathcal{OP}3 : \hat{\mathbf{s}} &= \arg \min \|\mathbf{s}\|_1 \\ s.t. \mathbf{y} &= \mathbf{\Phi\Psi s} \end{aligned}$$

## 4.5 压缩感知



- 4.5.2 信号重建算法

- 压缩感知的重建算法一般包括三类：贪婪算法、基追踪算法以及概率推断算法

- 1. 正交匹配追踪(OMP)算法

OMP 算法的具体流程描述如下：↵

(1) 设定迭代次数  $t=1$ ，计算观测值  $\mathbf{y} = \Phi \mathbf{x}$ ，初始化残差向量  $\mathbf{r}_0 = \mathbf{y}$ 、位置集合  $\Lambda_t = \emptyset$ 、近似向量  $\mathbf{a}_t = \mathbf{0}$ ，矩阵  $\Theta_0 = [\emptyset]$  为空矩阵；↵

For  $1 < t < K$  ↵

(2) 寻找  $\Phi$  中与  $\mathbf{r}_{t-1}$  内积最大的一列，记录位置索引

$$\lambda_t = \arg \max_{j=1, \dots, N} \left| \langle \mathbf{r}_{t-1}, \boldsymbol{\phi}_j \rangle \right|; \quad \leftarrow$$

## 4.5 压缩感知



(3)更新位置集合  $\Lambda_t = \Lambda_{t-1} \cup \{\lambda_t\}$ , 更新矩阵  $\Theta_t = [\Theta_{t-1}, \Phi_{\lambda_t}]$ ;

(4)利用最小二乘法求解如下问题

$$\mathbf{z}_t = \arg \min_{\mathbf{z}} \|\mathbf{r}_{t-1} - \Theta_t \cdot \mathbf{z}\|_2 \quad (4.5.8)$$

(5)更新近似向量  $\mathbf{a}_t = \Theta_t \mathbf{z}_t$  与残差向量  $\mathbf{r}_t = \mathbf{y} - \mathbf{a}_t$ ;

End For

(6)利用位置集合  $\Lambda_K$  恢复原始信号  $\hat{\mathbf{s}} = \Psi^T \mathbf{a}_K$ 。

OMP算法的优点在于收敛速度较快, 与MP算法有同样的恢复性能。由于每次迭代仅更新一个元素, 增加了运行的时间, 且迭代次数与稀疏度相关。因此需要较多的观测量以保证重构精确性, 如果稀疏度较大, 则运行时间也会随之增加。



## 4.5 压缩感知

- 2. 基追踪(*Basis Pursuit*) 算法
- 这一类算法的基本思想是根据1范数寻找信号稀疏解,
- 代表性算法包括LASSO (least absolute shrinkage and selection operator)算法、自适应LASSO算法以及Dantzig Selector(DS)算法

LASSO算法优化模型

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \mathbf{y} - \sum_{j=1}^N \boldsymbol{\phi}_j x_j \right\|^2 + \lambda \sum_{j=1}^N |x_j|$$

自适应LASSO算法优化模型

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \mathbf{y} - \sum_{j=1}^N \boldsymbol{\phi}_j x_j \right\|^2 + \lambda \sum_{j=1}^N w_j |x_j|$$

$\mathcal{OP}5$ : minimize  $\|\mathbf{x}\|_1$

Dantzig Selector(DS)算法

$$s.t. \left\| \boldsymbol{\Phi}^T (\mathbf{y} - \boldsymbol{\Phi} \mathbf{x}) \right\|_{\infty} \leq \lambda \sigma$$



## 4.5 压缩感知

### 3. 稀疏贝叶斯学习算法

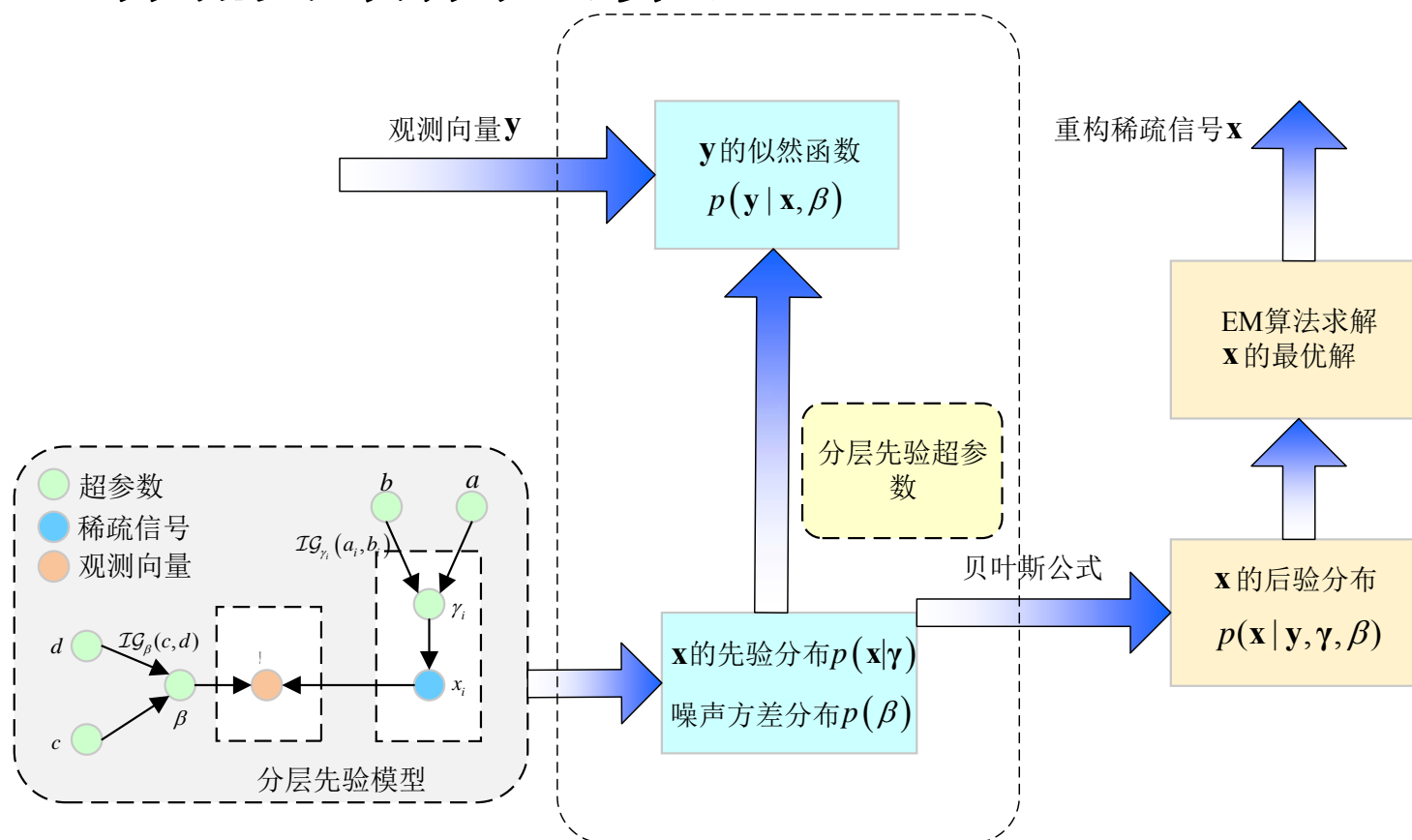


图4.17 稀疏贝叶斯学习算法流程图

## 4.5 压缩感知

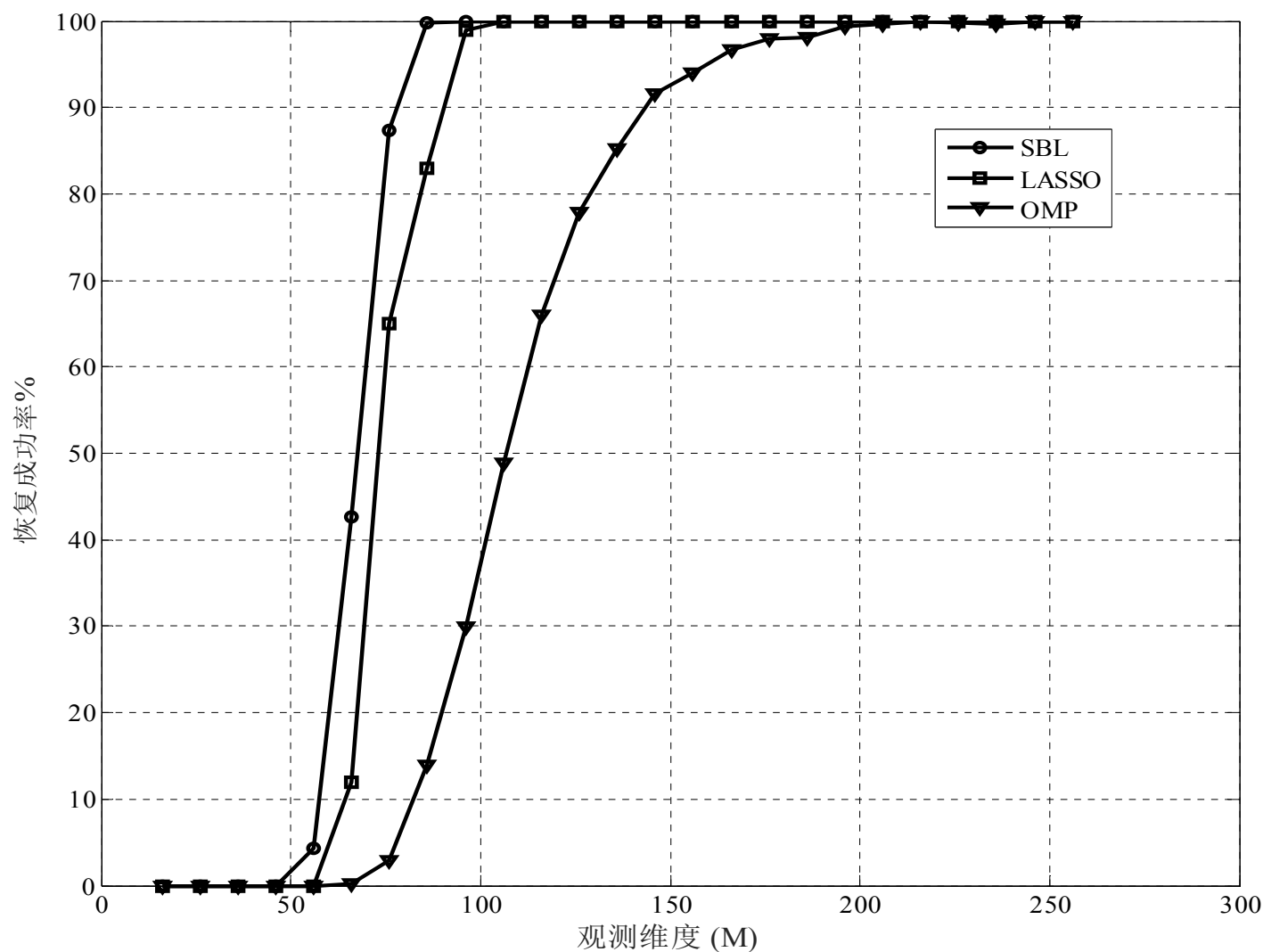


图4.18 不同观测值数目下三种算法恢复成功百分比(稀疏度 $K=16$ )

## 4.5 压缩感知



- 压缩感知理论突破了信号处理的金科玉律——奈奎斯特采样定律，是21世纪信号处理领域最耀眼的成果。
- 已经在磁共振(MRI)成像、图像处理等领域取得了有效应用。
- 在无线信号处理领域也有广泛应用，例如信道估计、大规模MIMO检测、多用户检测、频谱感知等。





## § 4.6 本章小结

- 本章主要讨论了移动通信物理层中系统有效性(数量)即信源编码问题。
- 首先从理论和实际两方面介绍了移动通信中语音编码，理论方面：语音压缩编码的基本原理、标准和性能指标；实际方面：重点介绍了移动通信中目前采用的各类语音编码方案：含GSM、IS-95、CDMA2000和WCDMA。
- 其次介绍主要用于多媒体的图像压缩编码，介绍了图像压缩编码原理、标准与主要技术指标，并重点介绍了静态图片的JPEG、活动图像MPEG和面向通信的H.263等原理性框图，对第二代视频编解码技术JPEG2000、MPEG4/H.264和我国自主音视频标准DRA、AVS进行了简要介绍。
- 最后概述了信号处理领域的新突破——压缩感知的基本理论与算法。

# 参考文献



- [4.1] 周炯槃, 《信源编码原理》, 人民邮电大学出版社, 1996
- [4.2] 吴伟陵, 《信息处理与编码(修订本)》, 人民邮电出版社, 2003.7
- [4.3] 吴伟陵, 《移动通信中的关键技术》, 北京邮电大学出版社, 2000.11
- [4.4] R. J. McEliece, The Theory of Information and Coding (2nd), Addison-Wesley Publishing Company, Inc. 2002
- [4.5] 3GPP Mandatory Speech Codec Speech Processing Functions, AMR Speech Codec: General Description (3G TS 26.071 Version 3.0.1), 1999
- [4.6] 3GPP 3G TS 26.101 Version 1.4.0 1999



谢谢！