

# 第 21 讲：视频压缩简介

---

- 视频压缩方法

- 帧内压缩
- 条件刷新和运动补偿
- 3D变换编码（DCT和小波）
- 2D变换+频率系数的预测

- ITU-T和ISO标准概述

- ISO MPEG 1 标准

- 输入视频格式和数据结构
- 图像类型和压缩方法

- ISO MPEG 2 标准

- 输入视频格式和数据结构
- 隔行视频压缩扩展
- 档次和级别

# 帧内编码

---

- 最简单的视频编码方法
- 帧内编码具有以下优点：
  - 随机访问每个帧，
  - 非常低的延迟，
  - 低复杂度（无多帧存储和运动估计）
- 帧内方法的压缩效率有限。

# 帧内编码 – 应用

- 用于摄像机中的消费数码视频格式DV及其变体DVCPRO和DVCAM以及用于专业电视制作的DVCPRO-50和DVCPRO HD采用25Mbits / s至50Mbits / s的帧内视频编码。它们使用基于DCT的类似JPEG的帧内压缩。通过允许在一帧内对量化表进行优化，在标称5: 1的压缩比条件下，DV可获得比JPEG更高的质量。它还使用自适应场间压缩，即如果没有检测到运动，则将隔行帧的两个场压缩在一起。
- Motion JPEG 2000 (MJ2K) 使用JPEG 2000来逐帧压缩运动图像或视频。除了实现比JPEG更好的压缩效率之外，JPEG 2000还支持单个嵌入式比特流的无损，近无损和有损编码，这使得MJ2K对于电影业具有较大的吸引力。

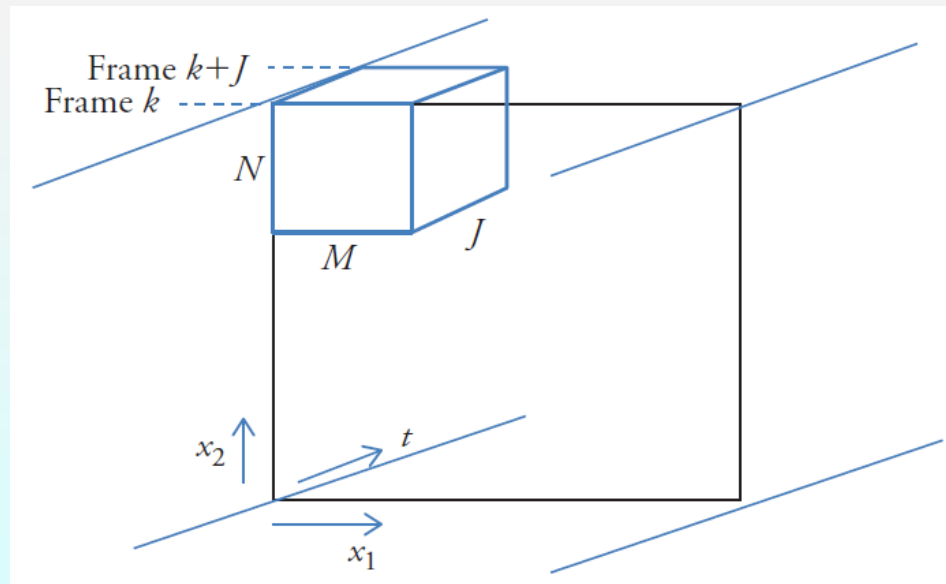
# 帧间编码: CR 与 MC

---

- **条件刷新 (CR)**是现代视频压缩标准跳帧（块）模式的基础：
  - 计算帧差（无运动补偿）
  - 只编码有变化的区域
- 所有现代视频压缩标准都采用**运动补偿 (MC)** 编码：
  - 在当前帧和参考帧之间进行**运动估计**
  - 计算**位移帧差**
  - 使用**2D DCT**对位移帧差进行编码。

# 帧间编码: 3D DCT

- 视频分为 **$M \times N \times J$** 的块，其中 $J$ 表示时间维度。
- 类似于**2D DCT编码**，变换系数要进行量化和编码。
- 3D DCT编码**不需要**单独的**运动估计**步骤。
- 然而，它需要编码器和解码器中**存储 $J$ 帧**。
- 注意：要对视频进行随机访问，**每 $J$ 帧**才可能随机访问一次。





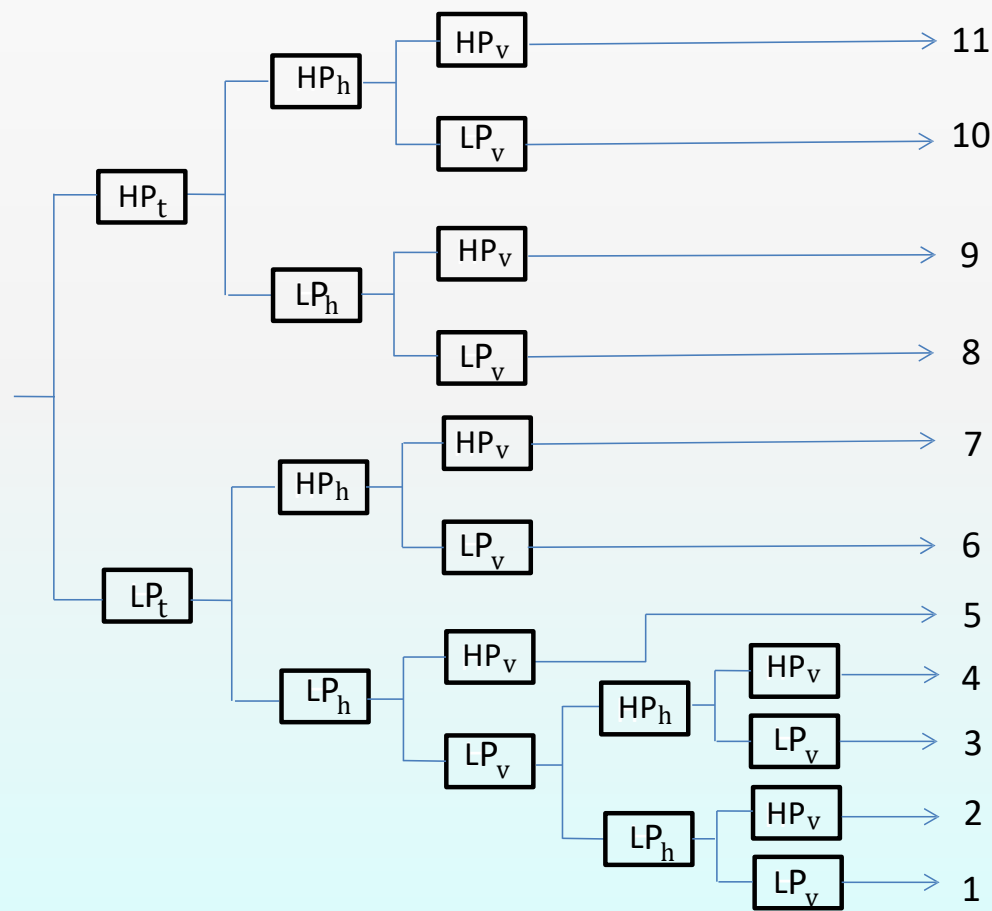
# 帧间编码: 2D 变换 + DPCM

---

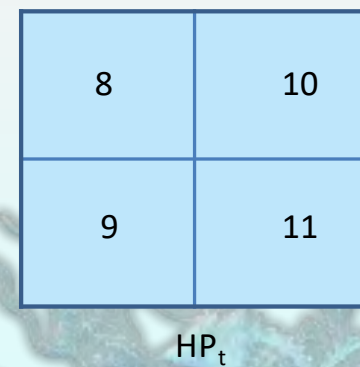
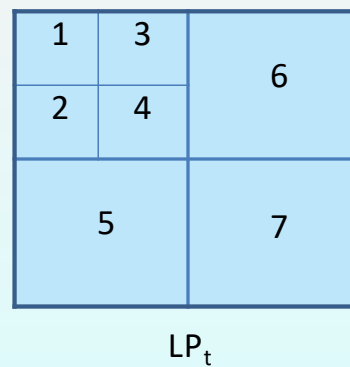
- **混合变换/ DPCM**编码在时间方向上实现2D变换系数的差分脉码调制 (DPCM) , 以**克服3D变换编码的多帧存储**需求。
- 它对于所有帧的每个空域 $N \times N$ 块执行**2D DCT**, 如JPEG。然后, 将DPCM编码器组中的每个编码器根据特定DCT系数的统计分布进行调整, 以应用于时间方向上的变换系数。也就是说, 对**时间方向上的各个DCT系数之差**进行量化和编码。
- 可以根据2D DCT系数的时域统计适当地采取DPCM自适应量化器来**实现与3D DCT编码性能相当**的结果[Roe 77]。

# 帧间编码: 3D 小波变换

## • 11子带 3D 小波分解

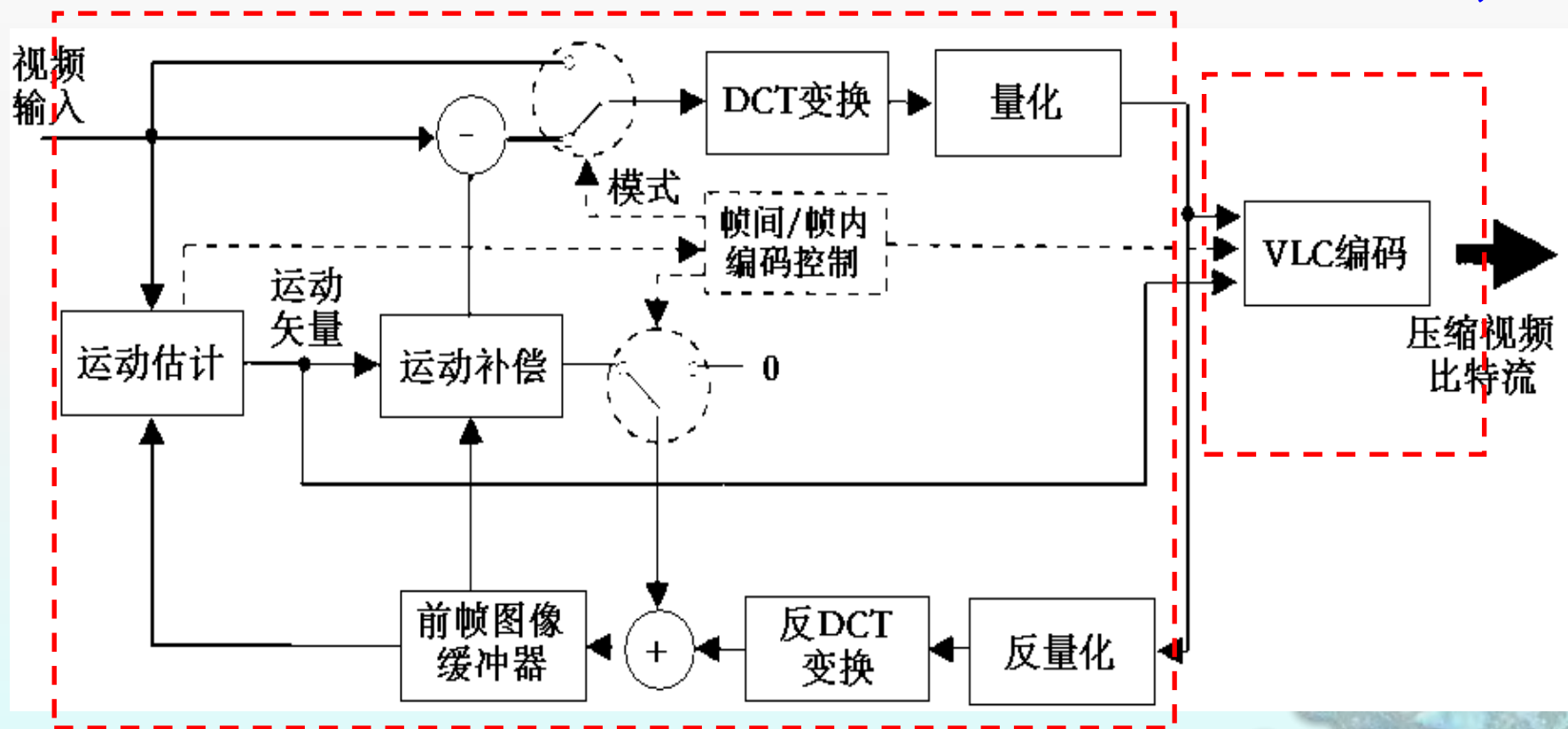


- (1,1)和 (1,-1) Haar滤波器的时间分解。
- 对时域低通和高通子带进行空间分解。



# 基于块的混合编码

无损压缩（无失真压缩）  
包括：游程编码，熵编码



有损压缩（率失真压缩）



# 视频压缩标准概述

- ITU-T 标准 – 视频编码专家组 (VCEG)
- ISO 标准 – 运动图像专家组 (MPEG)
- JCT-VC – 视频编码联合协作小组

**Table 8.1** International Standards for Video Compression

Standard	Approved By	First Edition	Short Description
H.261	ITU-T	1988	Obsolete
MPEG-1	ISO/IEC	1993	Obsolete
MPEG-2/H.262	ISO/IEC, ITU-T	1996	Digital broadcast
H.263	ITU-T	1996	Obsolete
MPEG4-AVC/H.264	ISO/IEC, ITU-T	2003	All purpose
MPEG HEVC/H.265	ISO/IEC, ITU-T	2013	All purpose

# ISO MPEG-1 标准

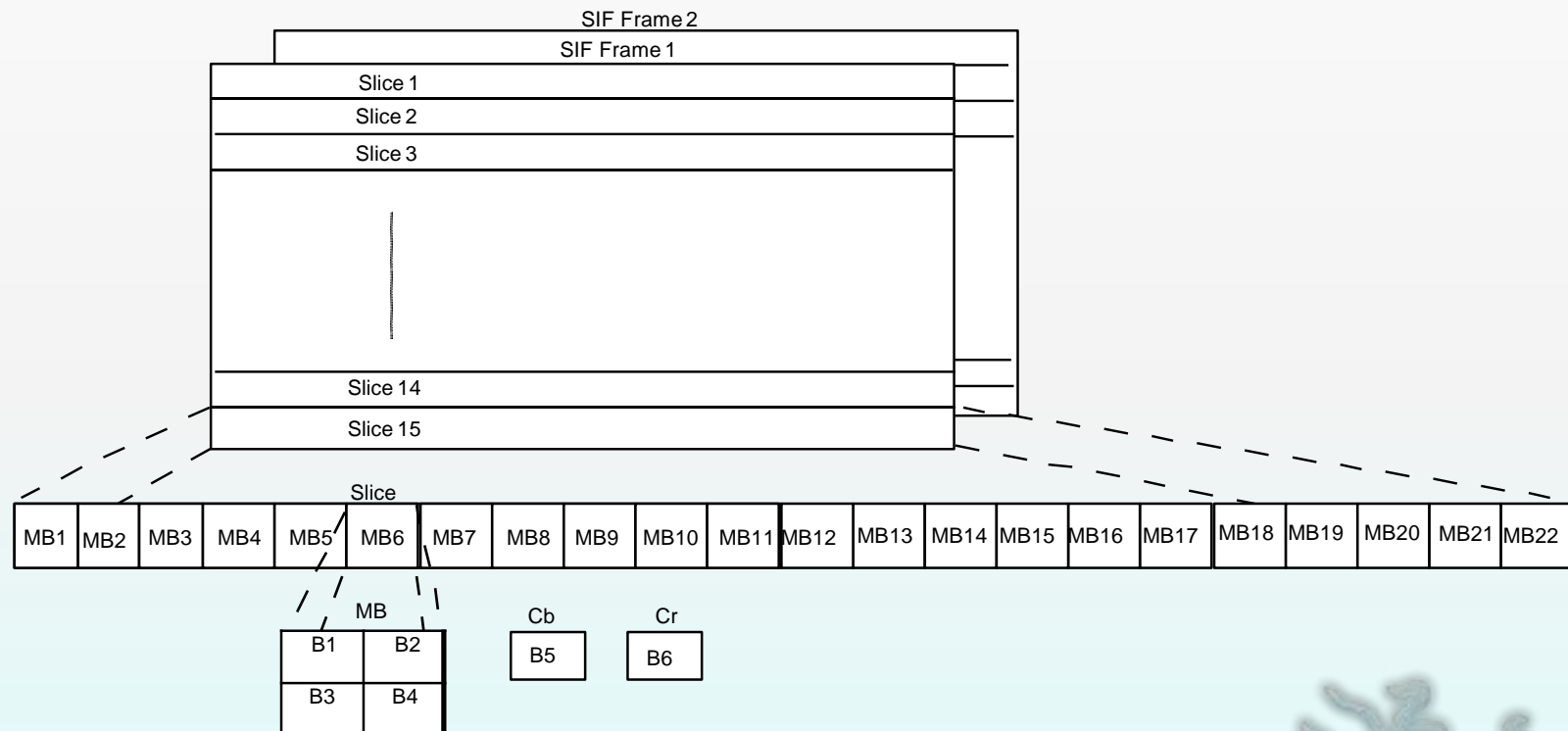
- **MPEG-1标准**规定了每个图像的压缩模式，例如INTRA，P和B型运动补偿，以及用于每个宏块的DCT系数的量化和编码的方法。
- 编码器功能如运动估计，压缩模式的选择，每个MB设置MQQUANT值以及码率控制都不作为标准的规范部分。
- 该标准仅指定由**图像组（GOP）**，**图像和MB级语法**组成的比特流语法，以传送每个MB的压缩模式（MTYPE），运动矢量和MQQUANT值以及可变长编码（VLC）的DCT系数。
- 编码器的**IDCT模块**应与解码器上的IDCT模块匹配，如IEEE标准1180-1990中规定的64位浮点IDCT，以避免预测过程中的错误传播。

# MPEG-1:目标比特率和输入格式

- **应用**：VideoCD（已淘汰）
- **目标比特率**：1.5 Mbps（视频为1.2 Mbps，其余为音频和系统）
- **输入视频格式**：SIF（标准输入格式）
  - 逐行扫描  $352 \times 288 \times 25$
  - 逐行扫描  $352 \times 240 \times 30$
  - 帧率 23.97, 24, 25, 29.97, 50, 59.94, 60 Hz
  - 颜色表示: Y, Cb, Cr, 4:2:0 色度子采样
- **从隔行扫描转换**：丢弃所有偶数或奇数场，对保留场沿水平维度进行2倍下采样
- **视频质量**：符合VHS录制的视频质量：

# MPEG-1数据结构和比特流

- 图像组, 图像, 片, 宏块 (MB) 和块层



编码单元: MB – 四个 $8 \times 8$ 亮度块, 一个 $8 \times 8$  Cr块, 一个 $8 \times 8$  Cb块。

变换单元:  $8 \times 8$  块



# MPEG 1 – 数据结构

---

1) **序列**, 由多个图像组组成

2) **图像组 (GOP)**, 由多幅图像组成

3) **图像**, 由多个片组成

- I-图像: 帧内编码
- P-图像: 相对于之前的I或P图像
- B-图像: 正向, 反向或双向
- D-图像: 只包含每个块的直流分量

4) **片**, 由多个宏块组成  
主要用来错误恢复。

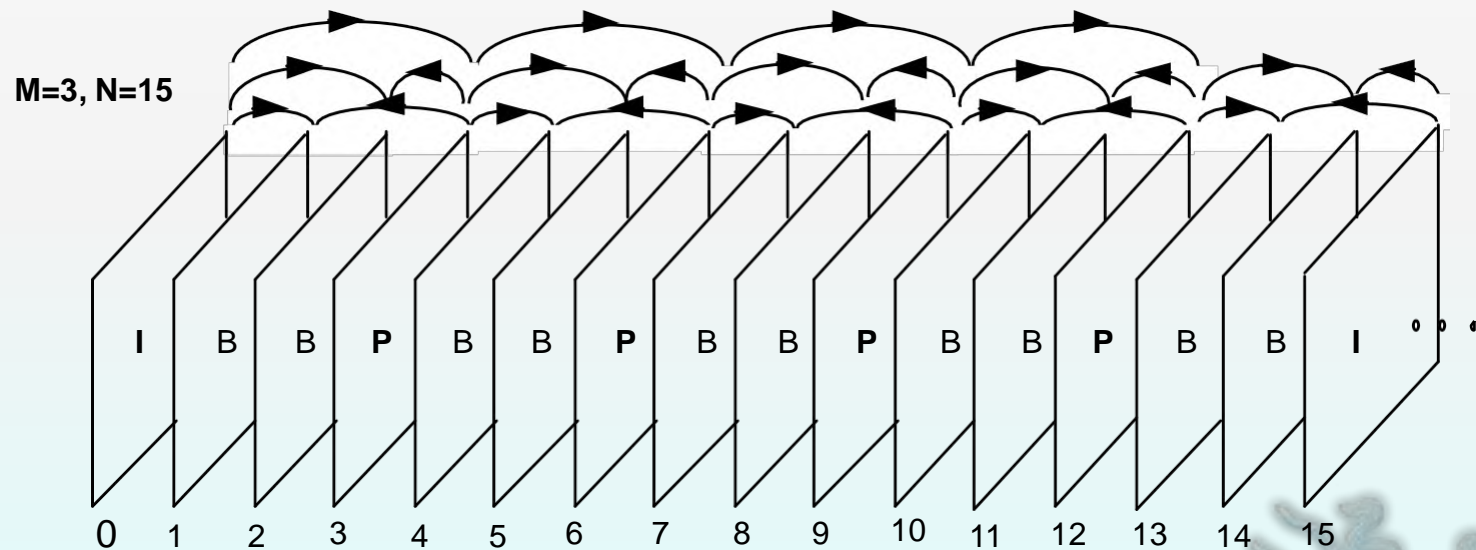
5) **宏块**, 由多个块组成  
一个 MB 包括 4 Y, 1 Cr 和 1 Cb 块 (与 H.261 MB 相同)

6) **块**,  $8 \times 8$  采样阵列



# 图像组 (GOP)

- **GOP**可以包含**I**、**P**和**B**图像；
- $N = \text{GOP}$ 中的图片数;  $M = \text{预测距离}$ ;
- 连续**锚定图像**之间的B图像数量为 $M-1$ 。



— 显示顺序      0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 ...  
— 编码顺序      0 3 1 2 6 4 5 9 7 8 12 10 11 15 13 14 ...

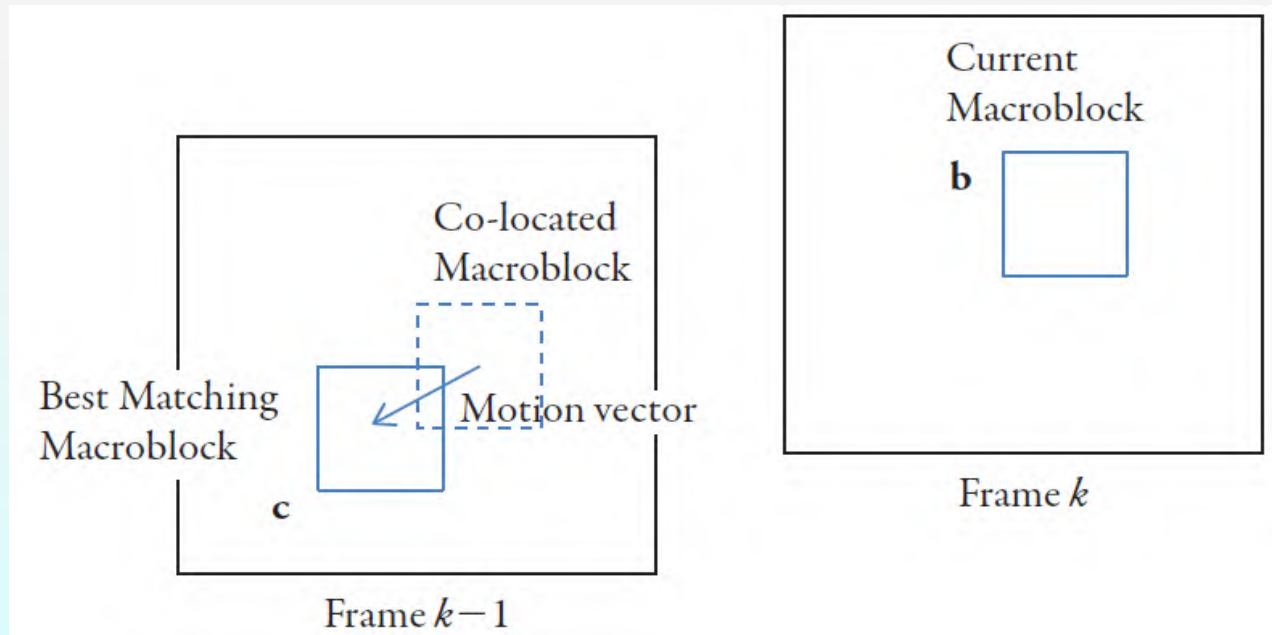
# 帧内帧

---

- 不参考任何其他图像的编码
  - 类似JPEG的DCT编码
  - 在MPEG-1中，每组图像（GOP）以I图像开头
  - 提供对视频比特流的随机访问
  - I图像消耗了大部分的比特预算
  - MPEG-1通过在每个MB的语法中引入量化器尺度参数“**MQUANT**”来允许空间自适应量化。
  - “**帧内**” **MB**：用当前量化矩阵编码。
  - “**帧内-A**” **MB**：量化矩阵由MQUANT缩放，在头数据中传输。
- MQUANT取整数值1到31。

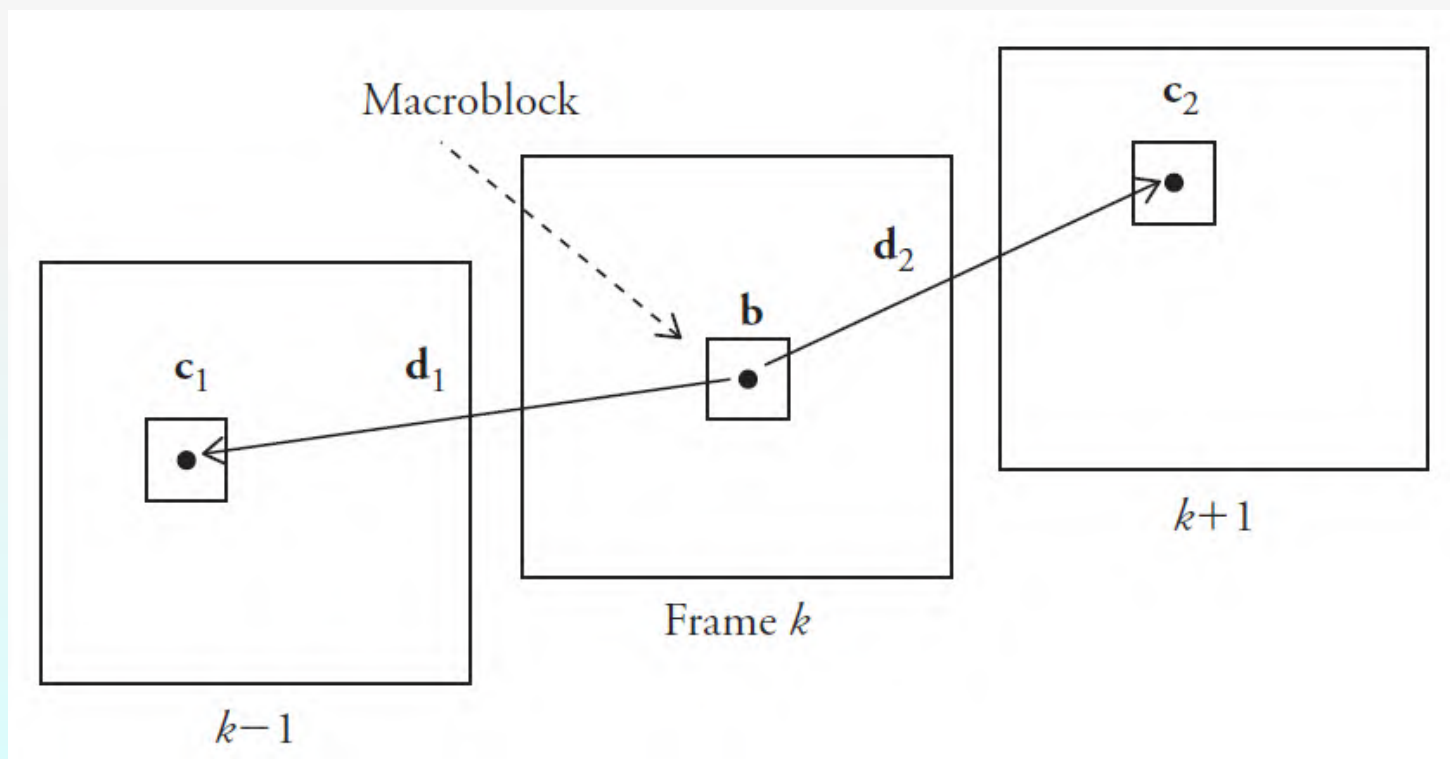
# 预测帧

- 反向运动估计 (正向运动补偿)
- 对运动补偿误差 (位移帧差) 进行编码
- P图像还支持帧内编码宏块。



# 双向预测帧

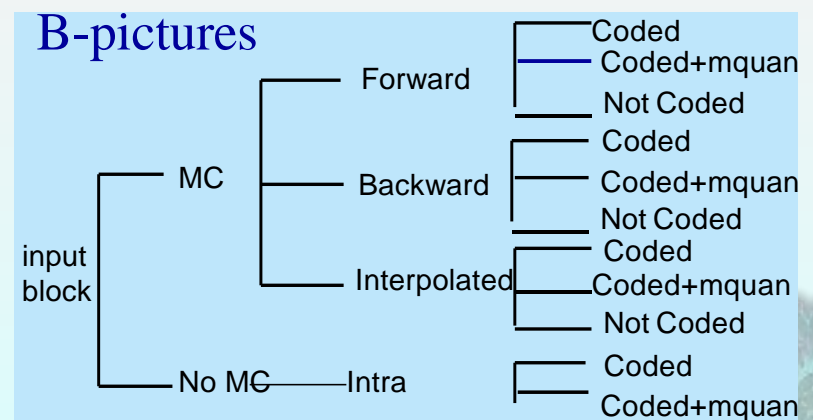
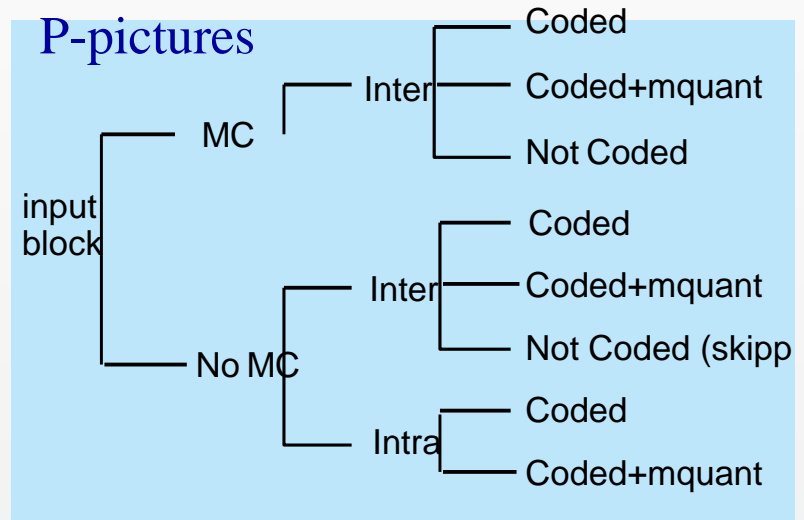
- 双向运动补偿
- **B图像**还支持**帧内**和**P型MC**编码的宏块。



# MPEG-1中的宏块类型

**Table 8.2** Macroblock Types in MPEG-1

I-pictures	P-pictures	B-pictures
Intra	Intra	Intra
Intra-A	Intra-A	Intra-A
	Inter-D	Inter-F
	Inter-DA	Inter-FD
	Inter-F	Inter-FDA
	Inter-FD	Inter-B
	Inter-FDA	Inter-BD
	Skip	Inter-BDA
		Inter-I
		Inter-ID
		Inter-IDA
		Skip





# 运动矢量编码

---

- GOP中的**I, P或B图像的数量**（即，N和M的值）是依赖于应用的。
- 该标准规定**每132张图像**中的一张图像必须是**I图像**，以避免由于编码器和解码器之间的IDCT不匹配而导致的**错误扩散**。
- 使用**DPCM**和**霍夫曼编码**对运动矢量进行编码。
- 运动矢量用**一半（0.5）像素精度**表示。
- 矢量可被表示的最大长度可以逐帧变化，以实现灵活性。
- 指向图像外像素的运动矢量是不允许的。

# MPEG-1示例编码器

- MPEG-1编码器包括用于运动估计的模块，压缩模式（MTYPE）的选择以及在每个MB处设置MQQUANT的值，以及运动补偿预测，量化器和去量化器，DCT和IDCT，可变长编码（VLC），多路复用器和缓冲调节器。
- 编码器端必须对解码器端处理过程进行复制，使得其产生的运动补偿预测与解码器相同，很显然，这是基于解码的重建帧（用于预测后续帧）。
- MPEG委员会开发了一个非规范模型编码器，称为仿真模型3（SM3），以验证标准。

# MPEG-1 编码器 (cont'd)

以下是SM3中提出的非规范性选择:

- 两个锚定帧之间B帧数量的常见选择是 $M = 1$ 或 $M = 2$ 。(MPEG-1中使用B帧是可选的。)
- 对每个MB估计运动矢量。SM3采用**对数搜索**和**望远镜搜索**方法，首先找到最佳的整数（全像素）运动矢量。望远镜搜索是在具有B帧的多帧上进行运动估计时减小搜索空间的方法。为了避免大的搜索范围，**望远镜搜索**将从帧到帧获得的最佳运动矢量估计进行级联。也就是说，以**零向量为中心搜索**从当前图像到先前图像的第一个运动矢量，然后**以前面步骤的最佳估计为中心搜索从前一图像向更先前图像的最佳估计**，等等。最后，对最终的全像素运动估计进行半像素更新。

# MPEG-1 编码器 (cont'd)

- 确定**每个MB的压缩模式MTYPE**有两种方法：
  - i) 一种**穷尽搜索**的方法使用每个允许的类型对每个MB进行编码，然后选择产生最佳R-D折中的一个。
  - ii) 更快的方法**依次进行一系列决定**，例如在P图像的情况下：
    - 确定是否传输运动矢量。
    - 如果没有选择MC，则决定是否使用没有运动矢量的帧内模式或帧间模式。
    - 如果选择了帧间模式，则决定残差是否足够大以进行编码。
    - 确定量化器参数是否满意或需要更改。
- 基于MB的**归一化空间活动性**来更新MQQUANT值。
- 通过一个被称为视频缓冲验证器（VBV）的解码器参数模型来实现**码率控制**。编码器监视模型缓冲区的状态，以更新每个图像类型的比特分配。缓冲器越大，编码器的灵活性就越大，但解码延迟也越大。



# ISO MPEG-2 标准

- **MPEG-2标准** (ISO / IEC-13818) 为数字电视, 有线/卫星电视, DVD和高清隔行视频编码的4-15 Mbit / s的广播质量隔行视频指定了**比特流语法和解码语义**, 15-30 Mbit / s用于HDTV应用。
- **视频编码竞争测试**始于1991年11月, 标准视频部分于1993年11月形成委员会草案, 于1996年出版。
- **MPEG-2 标准包括:**
  - ISO/IEC 13818-1: MPEG-2 系统
  - ISO/IEC 13818-2: MPEG-2 视频
  - ISO/IEC 13818-3: MPEG-2 音频;
  - ISO/IEC 13818-4: MPEG-2 一致性
  - ISO/IEC 13818-5: MPEG-2 软件
  - ISO/IEC 13818-6: MPEG-2 DSM-CC
  - ISO/IEC 13818-7: AAC 音频



# MPEG-2的新特性总结

- 引入了用于隔行视频的**帧和场图像**类型
- 引入了用于隔行视频的**新型MC预测**模式
- 为每帧图像的MB引入了**场/帧DCT选项**
- 除了之字形扫描外，还允许进行**交错扫描**。
- 允许对DCT系数进行**更精细的量化**。
- 允许更精细地**调整量化器比例因子**。
- 允许**单独的VLC表**用于帧内MB的DCT系数。
- 除了支持4: 2: 0色度格式，也支持**4: 2: 2**和**4: 4: 4**格式。
- 支持**容错编码和传输**（传输流）
- 引入**时间，空间和SNR可扩展视频编码**选项
- 引入**六个档次**（语法子集）和**四个级别**（对参数值的约束）。

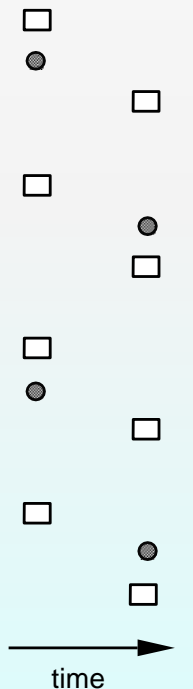
# MPEG-2: 输入视频格式

- 4:2:0, 4:2:2, 4:4:4格式中样值的垂直/时域位置

interlaced 4:2:0

top\_field\_first=1

top field  
bottom field



interlaced 4:2:0

top\_field\_first=0

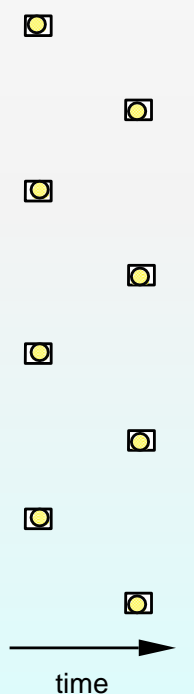
top field  
bottom field



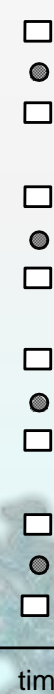
interlaced 4:2:2/4:4:4

top\_field\_first=1

top field  
bottom field

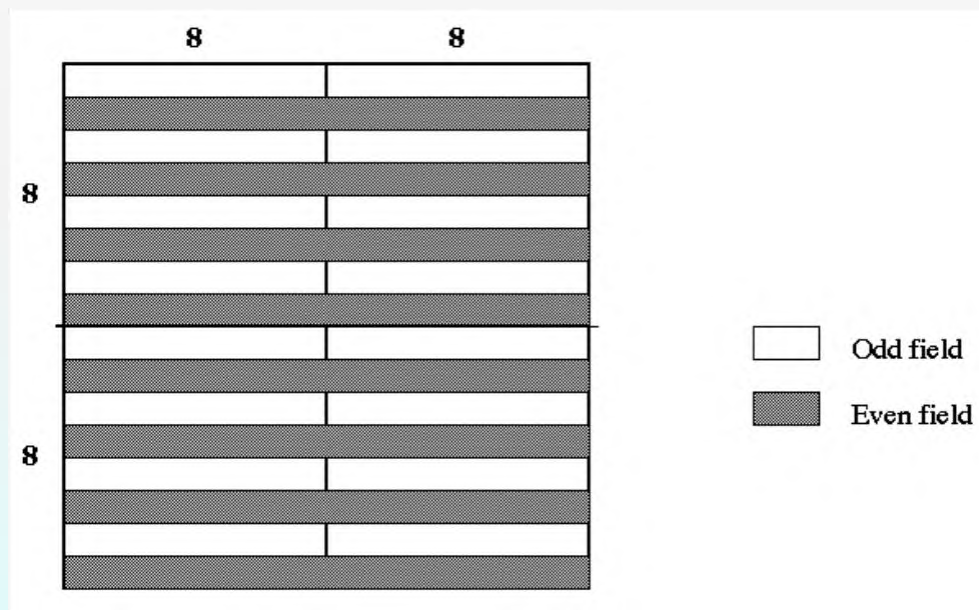


progressive

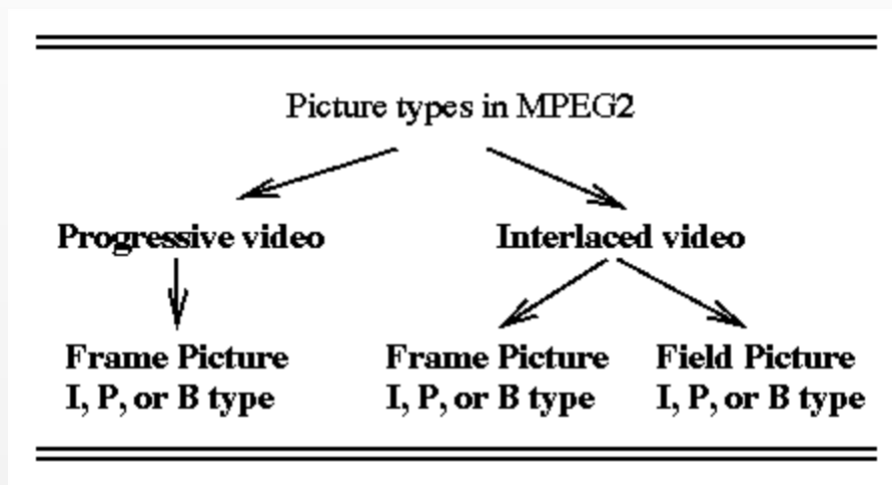


# 隔行视频的新图片类型

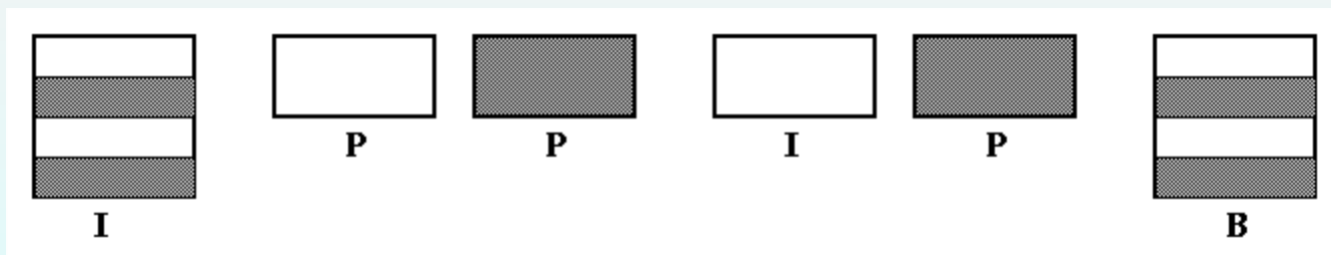
- **帧图像**：来自偶数和奇数场的行被交织以形成帧图像。帧图像可以是I，P或B型。
- **场图像**：偶数和奇数场被视为单独的图像。每个场图可以是I，P或B型。
  - 如果奇数场是P (B) 图像，则偶数场必须也是P (B) 图像。
  - 如果奇数场是I图像，那么偶数场可以是I或P图像。



# MPEG-2中的图像类型



- 场图像按照**显示的顺序**进行传输。



- 一个图像组可以由**场图像和帧图像的任意混合**组成。



# 用于隔行视频的MC预测模式

---

- 帧图像

- 基于帧的预测模式（P或B类型，与逐行视频相同）
- 基于场的预测模式
- 双重预测模式  
可以在宏块的基础上在同一帧图像内的帧/场/双重预测模式之间进行切换。

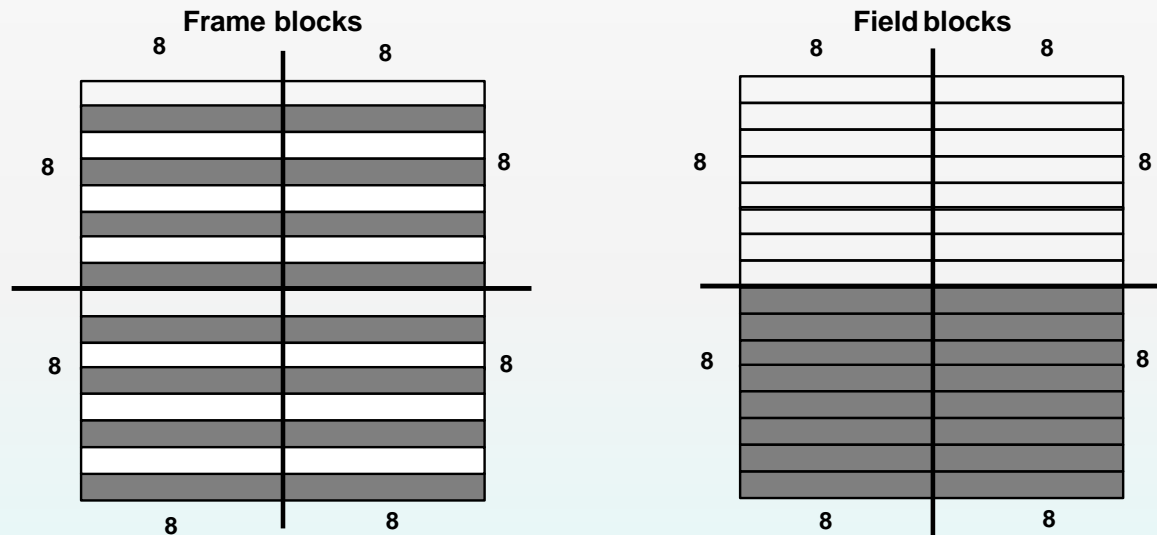
- 场图像

- 两种类型的基于场的预测模式
- 双重预测模式



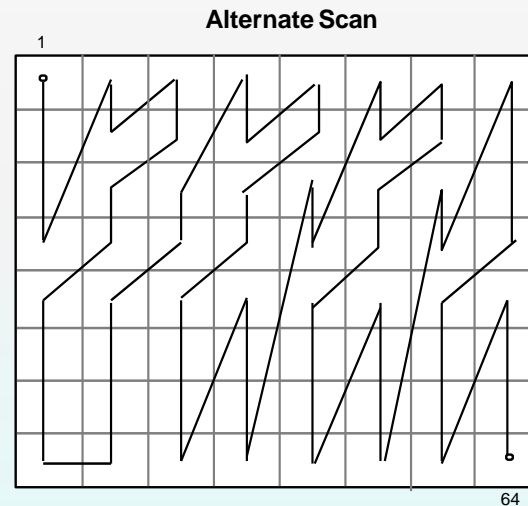
# 帧图像的场/帧 DCT 选项

- 计算帧和场块组织方式在**垂直方向上的相关性**。选择具有较高相关性的块组织方式。



- 例如，可以为包含**高运动的MB**选择**场DCT**，而**帧DCT**可适用于**运动很少或不运动但包含高空间变化**的MB。

- 除了之字形扫描，MPEG2允许进行以下可选扫描，称为“**交错扫描**”（用于**隔行扫描**）。



# DCT系数的更精细量化

- 帧内宏块

- DC系数的量化权重可以是8, 4, 2或1，即对于DC系数，允许11比特（全）分辨率。（注意：这个权重在MPEG-1中固定为8）。
- AC系数被量化到 $[-2048, 2047]$ 的范围内。（MPEG-1中该范围是 $[-256, 255]$ ）。

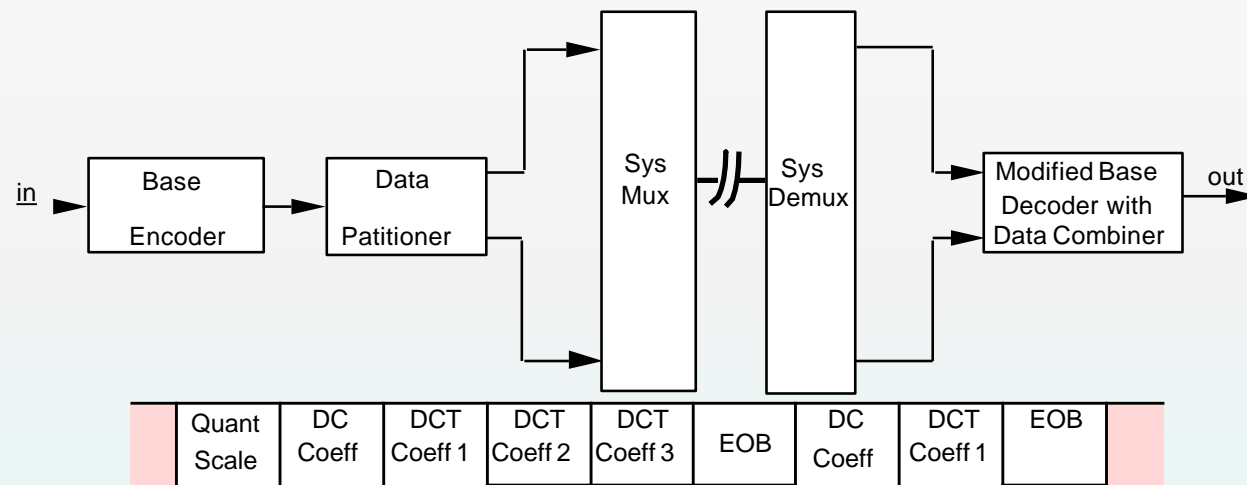
- 非帧内宏块

- 所有系数都被量化到 $[-2048, 2047]$ 的范围内。（MPEG-1中该范围是 $[-256, 255]$ ）。
- 精细调整MQQUANT

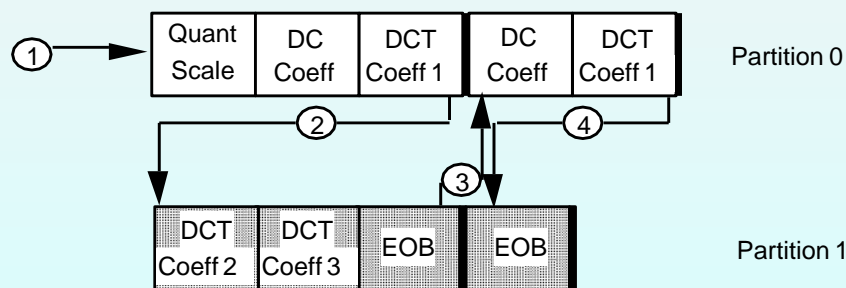
qnt_scl_code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
2*mquant (mpeg1)	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62
2*mquant (nonlin)	1	2	3	4	5	6	7	8	10	12	14	16	18	20	22	24	28	32	36	40	44	48	52	56	64	72	80	88	96	104	112

# 数据分区

- 将编码比特流分成**两个分区**：**分区0**是重要数据（例如，MV和DC系数），**分区1**是非重要数据。



- 例子:





# MPEG传输流

- **MPEG传输流 (MPEG-TS)** 是用于传输和存储**音频，视频**以及**程序**和**系统信息**协议 (PSIP) 数据，这些数据封装了基本流的数据包格式，可在信号降质时提供保持**完整性的纠错**和**码流同步**特征。
- 在MPEG-2第1部分系统 (ISO / IEC 13818-1或ITU-T H.222.0建议书) 中规定，并用于**广播系统**，如DVB，ATSC和IPTV。
- **数据包**是传输流中数据的**基本单位**。它以同步字节和头部开始。可能跟随一些附加字段，如在可选的自适应字段中的信号。数据包的其余部分由有效负载组成。
- MPEG-TS数据包的长度为**188字节**，但是通信媒介可以向数据包添加一些纠错字节。
  - ATSC传输增加了20字节的Reed-Solomon前向纠错，创建了208字节长的数据包
  - ISDB-T和DVB-T / C / S使用204字节。
- TS与为可靠媒体 (如DVD) 设计的节目流不同。

# MPEG-2 档次和级别

- 广播标准清晰度电视： **主要档次@主要级别**

↑ 级 别	HIGH		1920 pels/line 1152 lines/frame 60 frames/s 62.7 Msamples/s 80 Mbit/s			1920 pels/line 1152 lines/frame 60 frames/s 62.7 Msamples/s @ 83.5 Msamples/s * 100 Mbit/s for 3 layers
	HIGH-1440		1440 pels/line 1152 lines/frame 60 frames/s 47.0 Msamples/s 60 Mbit/s		1440 pels/line 1152 lines/frame 60 frames/s 47.0 Msamples/s 60 Mbit/s for 3 layers	1440 pels/line 1152 lines/frame 60 frames/s 47.0 Msamples/s @ 62.7 Msamples/s * 80 Mbit/s for 3 layers
	MAIN	720 pels/line 576 lines/frame 30 frames/s 10.4 Msample/s 15 Mbit/s	720 pels/line 576 lines/frame 30 frames/s 10.4 Msample/s 15 Mbit/s	720 pels/line 576 lines/frame 30 frames/s 10.4 Msample/s 15 Mbit/s for 2 layers		720 pels/line 576 lines/frame 30 frames/s 11.06 Msamples/s @ 14.75 Msamples/s * 20 Mbit/s for 3 layers
	LOW		352 pels/line 288 lines/frame 30 frames/s 3.04 Msamples/s 4 Mbit/s	352 pels/line 288 lines/frame 30 frames/s 3.04 Msamples/s 4 Mbit/s for 2 layers		
		SIMPLE non-scalable 4:2:0 (no B- pictures)	MAIN non-scalable 4:2:0	SNR scalable 4:2:0	SPATIAL scalable 4:2:0	HIGH non-scalable 4:2:2 scalable 4:2:0/4:2:2 * refers to 4:2:0 @ refers to 4:2:2

档次 →

# MPEG-2 示例编码器 (测试模型 5)

- $M = 1$  和  $M = 3$  的图像组; 帧图像
- 运动估计
  - 整数全像素搜索, 然后半像素更新
  - 帧/场/双重和正/反向运动估计
- 确定运动模式和运动类型
  - 帧/场/双重运动补偿类型
  - 正向/反向/插值运动补偿模式
- 帧/场DCT
- 宏块类型
- DCT系数的编码
  - 之字形扫描用于帧间编码; 交错扫描用于帧内编码
  - DCT系数的2D VLC, 帧间和帧内分别使用独立的编码表
- Mquant调整和码率控制