

图9-49描述了一个简单的运动检测器框图。这里的思想就是在低频亮度信号中比较帧间的变化。在很大程度上，运动检测器的性能决定图像的质量。运动信号（K）通常被整流，通过平均水平和垂直方向的几个样本使运动信号平滑，并乘以一个增益因子，然后在使用之前进行削波。运动检测器容易犯的唯一错误就是在图像的静态区域使用2D梳状滤波器。

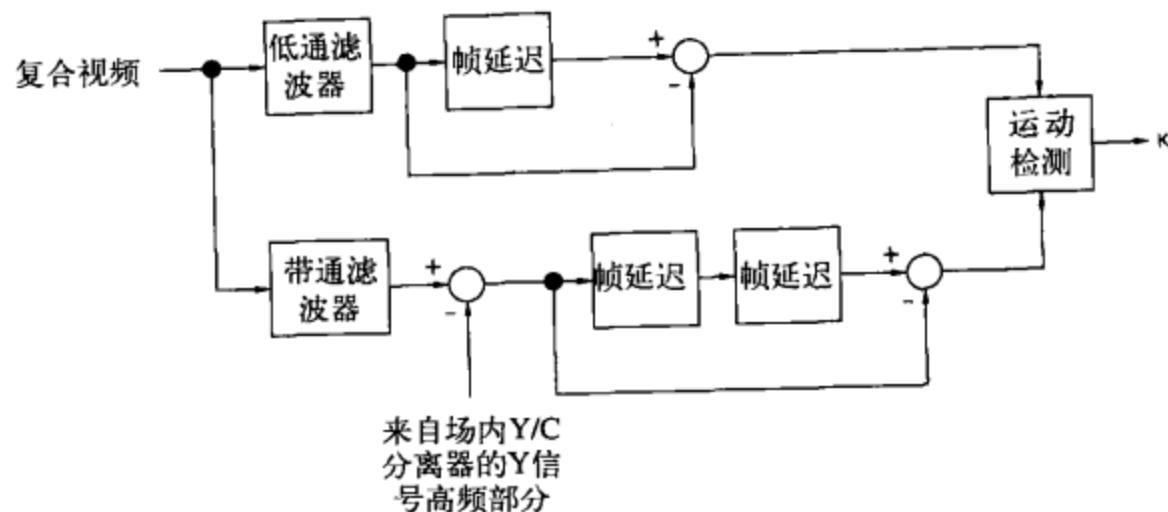


图9-49 NTSC制式的简单运动检测器框图

### 9.2.11 alpha通道支持

通过在NTSC/PAL解码器中集成一个额外的ADC，还可以数字化模拟alpha信号（也称为键），而且和视频数据一起流水，维持同步关系。这允许设计者改变解码器（可能拥有不同的流水线延迟），以适应具体的应用，而且不必考虑alpha通道的流水线延迟。alpha通常是线性的，模拟范围为0~100 IRE。alpha通道中不提供消隐基底或者同步信息。

### 9.2.12 解码器视频参数

为了指定NTSC/PAL解码器的相对质量，定义了许多工业级标准的视频参数。为了测量这些参数，NTSC/PAL解码器的输出（解码第8章中描述的各种不同视频测试信号）通过使用视频测试设备进行监测。表9-11示范了家用级和演播室级质量解码器的典型AC参数值以及个别参数的描述。

表9-11 NTSC和PAL解码器的典型AC视频参数

参 数	家用级质量	演播室质量0	单 位
微分相位	4	$\leq 1$	度
微分增益	4	$\leq 1$	%
亮度非线性特性	2	$\leq 1$	%
色调准确度	3	$\leq 1$	度
色饱和度准确度	3	$\leq 1$	%
SNR (遵循EIA/TIA RS-250-C)	48	>60	dB
色度-亮度串扰	$<-40$	$<-50$	dB
亮度-色度串扰	$<-40$	$<-50$	dB
H倾斜	$<1$	$<1$	%
V倾斜	$<1$	$<1$	%
Y/C采样偏离	$<5$	$<2$	ns
解调正交	$90 \pm 2$	$90 \pm 0.5$	度

有几个AC参数，例如短时间波形失真、组延迟和K因子，取决于模拟视频滤波器的质量，因此这里不对它们进行讨论。除了本节讨论的AC参数外，还有一些其他参数应该包含在解码器规范中，例如色同步捕捉与锁定频率范围，以及被解码YIQ或YUV视频信号的带宽。

另外还应该指定几个DC参数，如表9-12所示。尽管通常不指定同步锁相能力，但是除了时钟抖动，我们还是试图生成一列的同步锁相参数，如表9-13所示。

表9-12 NTSC和PAL解码器的典型DC视频参数

参 数	(M) NTSC	(B、D、G、H、I) PAL	单 位
同步输入幅度	40±20	43±22	IRE
色同步输入幅度	40±20	42.86±22	IRE
视频输入幅度(标称1V)	0.5~2.0	0.5~2.0	V

表9-13 NTSC和PAL解码器的典型同步锁相参数。这些参数假定视频信号的SNR>30dB，并且优于表9-12所示的DC参数

参 数	最 小	最 大	单 位
同步锁定时间 <sup>1</sup>		2	场
同步恢复时间 <sup>2</sup>		2	场
短期同步锁范围 <sup>3</sup>	±100		ns
长期同步锁范围 <sup>4</sup>	±5		μs
校正之前连续丢失行同步脉冲数量	5		同步脉冲
场相关性 <sup>5</sup>		±5	ns
短期副载波锁定范围 <sup>6</sup>	±200		Hz
长期副载波锁定范围 <sup>7</sup>	±500		Hz
副载波锁定时间 <sup>8</sup>		10	行
副载波准确度		±2	度

- 注：1. 从同步锁相过程开始到实现场相关性规范的时间。  
 2. 从同步锁相损失到实现场相关性规范的时间。  
 3. 维持场相关性规范的范围。短期范围假定行时间在两个连续行之间的改变量缓慢。  
 4. 维持场相关性规范的范围。长期范围假定行时间在一帧中的改变量缓慢。  
 5. 指示场样本准确度。对于使用VCO或者VCXO的同步锁相系统，这一规范与样本时钟抖动相同。  
 6. 副载波锁定时间和维持准确度规范的范围。短期时间假定副载波在2帧内改变量缓慢。  
 7. 副载波锁定时间和维护准确度规范的范围。长期时间假定副载波在24小时内改变量缓慢。  
 8. 在副载波瞬时180°相位偏移之后，上锁时间位于±2°之内。副载波频率的标称值为±500Hz。

## 1. 微分相位

微分相位失真，通常称为微分相位，指的是色度相位受亮度电平的影响有多大，换句话说，就是当亮度电平变化时色调偏移有多大。相位误差可能是正的也有可能是负的，因此微分相位表示为峰-峰测量值，以副载波相位的度数表示。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级解码器的微分相位参数接近于1°或者更小。

## 2. 微分增益

微分增益失真，通常称为微分增益，指的是色度增益受亮度电平的影响有多大，换句话说，

就是当亮度电平变化时色饱和度偏移有多大。信号有可能衰减也有可能放大，因此微分增益表示为两个电平之间的最大幅度变化，表示为最大色度幅度的百分比。

该参数使用一个相位和幅度均匀的色度叠加到不同亮度电平（例如调制斜坡测试信号或者复合测试信号的调制5级部分）上测量而得。演播室质量级解码器的微分增益参数接近于1%或者更小。

### 3. 亮度非线性特性

亮度非线性特性也称为微分亮度和亮度非线性失真，表示亮度增益受亮度电平的影响有多大，换句话说，就是解码的亮度电平和理想亮度电平之间的非线性关系。

459  
461通过使用一个非调制的5级或者10级阶梯测试信号，或者复合测试信息的调制5级部分，将最大级和最小级之间的差，表示为最大级的一个百分比，用来表示亮度非线性特性。尽管这个参数包含在微分增益和微分相位参数中，但是传统上这个参数单独指定。

### 4. 色度非线性相位失真

色度非线性相位失真指的是色度相位（色调）受色度幅度（色饱和度）的影响有多大，换句话说，就是当饱和度变化时色调偏移有多大。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，可以测量每个色度包的解码器输出。最大色调测量值和最小测量值之间的差就是峰—峰值。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

### 5. 色度非线性增益失真

色度非线性增益失真指的是色度增益受色度幅度（色饱和度）的影响有多大，换句话说，就是指已解码的色度幅度电平和理想色度幅度电平之间的非线性关系，这通常看作高饱和度色度信号的衰减。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，对解码器进行调整，从而使中间色度包（40 IRE）被正确解码。另2个解码色度包的幅度的测量值和标称值之间的最大差表示色度非线性增益失真，用IRE表示，或者表示为最坏包标称幅度的百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

### 6. 色度—亮度互调

色度—亮度互调通常称为交叉调制，表示亮度电平受色度的影响有多大。这可能是高饱和度色度电平削波或正交失真的结果，而且由于色饱和度的改变还可以表现为一个不规则的亮度变化。

通过使用一个调制基底测试信号，或者组合测试信号的调制基底部分，用解码的50 IRE亮度电平和被解码亮度电平之间的最大差来表示色度—亮度互调，色度—亮度互调可以用IRE表示，或者表示为一个百分比。这个参数通常不能单独指定，但是可以包含在微分增益和微分相位参数中指定。

### 7. 色调准确度

462  
464色调准确度表示被解码的色调值与理想的色调值之间有多紧密。相位误差可能是正的也可能是负的，所以色调准确度是最坏情况下的正测量值和最坏情况下的负测量值（相对于标称值）之间的差，以副载波相位的度数表示。这个参数使用EIA或EBU的75%彩条作为测试信号进行测量。

### 8. 色饱和度准确度

色饱和度准确度表示被解码的饱和度值与理想的饱和度值之间有多紧密，该参数使用EIA或EBU的75%彩条作为测试信号。该测量值可能是增益也有可能是衰减，所以色饱和度准确度是最坏情况下的增益测量值和最坏情况下的衰减测量值（相对于标称值）之间的差，表示为标称值的

百分比。

### 9. H倾斜

H倾斜也称行倾斜和行时间失真，导致行频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像左边和右边之间亮度的变化。对于数字解码器，H倾斜主要是模拟输入滤波器和传输媒介的伪影。H倾斜使用一个线条（例如NTC-7 NTSC复合测试信号中的线条）来测量，并且测量倾斜的峰—峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的第一个和最后一个微秒。

### 10. V倾斜

V倾斜也称为场倾斜和场时间失真，导致场频信号中出现倾斜的主要因素是白条。这种类型的失真导致图像的顶边和底边之间亮度的变化。对于数字解码器，V倾斜主要是模拟输入滤波器和传输媒介的伪影。V倾斜使用 $18\mu s$ 、100 IRE的白条（位于场中心的130行扫描行的中心）或者场方波测量。所测量的是倾斜的峰—峰变化（用IRE表示，或者白条幅度的百分比），但是忽略白条的前3行和最后3行。

## 参考文献

1. Benson, K. Blair, 1986, *Television Engineering Handbook*, McGraw-Hill, Inc.
2. Clarke, C.K.P., 1986, *Colour encoding and decoding techniques for line-locked sampled PAL and NTSC television signals*, BBC Research Department Report BBC RD1986/2.
3. Clarke, C.K.P., 1982, *Digital Standards Conversion: comparison of colour decoding methods*, BBC Research Department Report BBC RD1982/6.
4. Clarke, C.K.P., 1982, *High quality decoding for PAL inputs to digital YUV studios*, BBC Research Department Report BBC RD1982/12.
5. Clarke, C.K.P., 1988, *PAL Decoding: Multidimensional filter design for chrominance/luminance separation*, BBC Research Department Report BBC RD1988/11.
6. Drewery, J.O., 1996, *Advanced PAL Decoding: Exploration of Some Adaptive Techniques*, BBC Research Department Report BBC RD1996/1.
7. ITU-R BT.470-6, 1998, *Conventional Television Systems*.
8. *NTSC Video Measurements*, Tektronix, Inc., 1997.
9. Perlman, Stuart S. et. al., *An Adaptive Luma-Chroma Separator Circuit for PAL and NTSC TV Signals*, International Conference on Consumer Electronics, Digest of Technical Papers, June 6-8, 1990.
10. Sandbank, C. P., *Digital Television*, John Wiley & Sons, Ltd., 1990.
11. SMPTE 170M-1999, *Television-Composite Analog Video Signal-NTSC for Studio Applications*.
12. *Television Measurements, NTSC Systems*, Tektronix, Inc., 1998.
13. *Television Measurements, PAL Systems*, Tektronix, Inc., 1990.

## 第10章 H.261和H.263

目前有多个视频会议标准，如表10-1所示。图10-1至图10-3举例说明了几个常见的视频会议系统的结构图。

表10-1 视频会议标准系列

	H.310	H.320	H.321	H.322	H.323	H.324	3G-324M
网络	宽带ISDN ATM LAN	窄带交换 数据ISDN	宽带ISDN ATM LAN	保证带宽分组 交换网络	不保证带宽分组 交换网络（以太网）	PSTN或 POTS	移动网络
视频编解 码器	MPEG-2 H.261	H.261 H.263 H.264	H.261 H.263	H.261 H.263	H.261 H.263 H.264	H.261 H.263	MPEG-4.2
音频编解 码器	MPEG-2 G.711 G.722 G.728	G.711 G.722 G.722.1 G.728	G.711 G.722 G.722.1 G.728	G.711 G.722 G.722.1 G.723.1 G.728 G.729	G.711 G.722 G.722.1 G.723.1 G.728 G.729	G.723	G.722.2 G.723.1
复用	H.222.0 H.222.1	H.221	H.221	H.221	H.225.0	H.223	H.223A/B
控制	H.245	H.231 H.242 H.243	H.242	H.230 H.242	H.245	H.245	H.245
多点		H.231	H.231	H.231	H.323		
数据	T.120	H.239 T.120	T.120	T.120	H.239 T.120	T.120	T.120
通信接口	AAL I.363 AJM I.361 PHY I.432	I.400	AAL I.363 AJM I.361 PHY I.400	I.400 和 TCP/IP	TCP/IP	V.34 调制解调器	移动通信

### 10.1 H.261

ITU-T H.261是第一个为视频会议系统而开发的视频压缩和解压标准。H.261最初是在窄带综合业务数字网（N-ISDN）上开展速率为 $p \times 64$  kbit/s的可视电话和视频会议等业务而制定的，其中 $p=1\sim30$ 。现在，H.261已经成为了所有视频会议标准的最低需求，如表10-1所示。

图10-4给出了一个典型的H.261编码器结构图。视频编码器提供了一个与其他信号（如控制信号、音频信号等）复用的独立数字视频比特流。视频解码器执行与编码器相反的过程。

如图3-7所示，H.261视频数据采用4：2：0的YCbCr格式。表10-2中给出了其主要的规格说明。H.261最大的图像速率是受限的，即在两个传送图像之间可能有0、1、2或3个非传送图像。

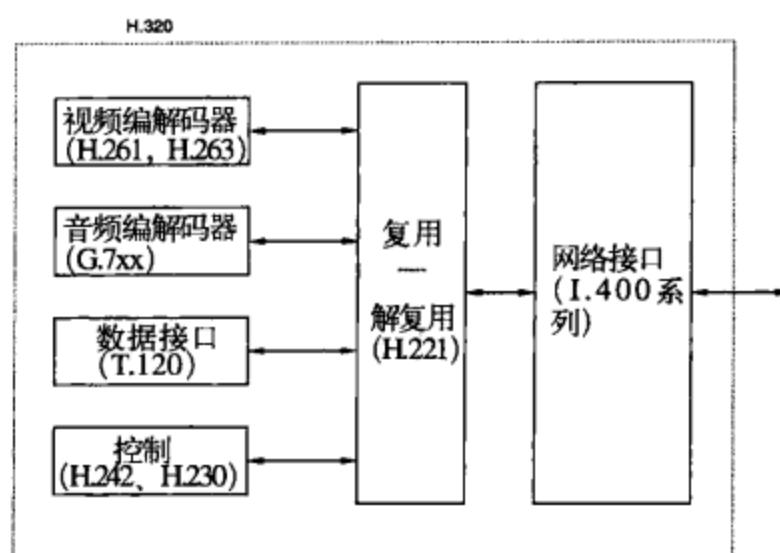


图10-1 典型的H.320系统

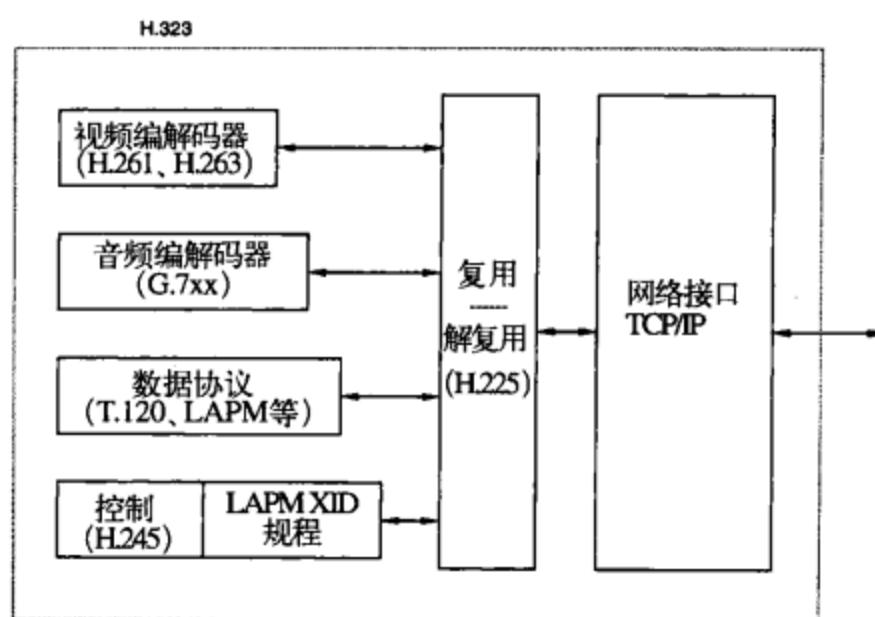


图10-2 典型的H.323系统

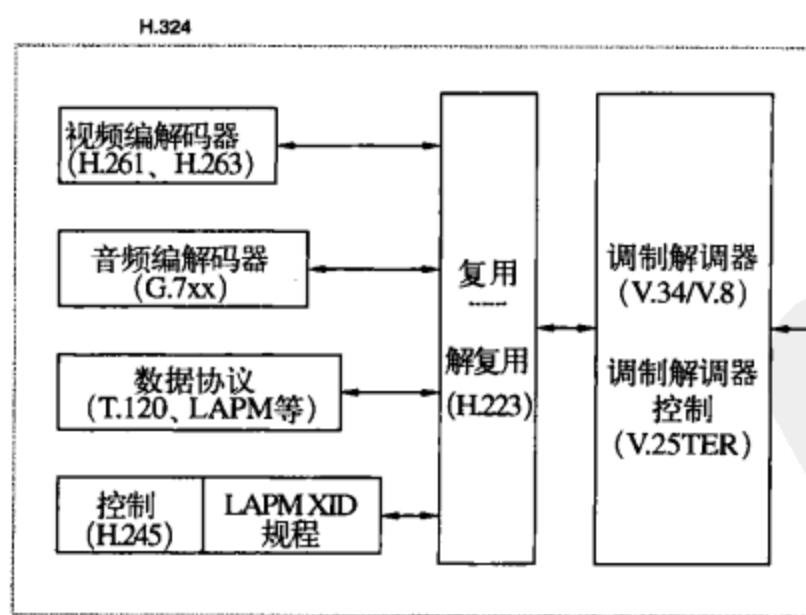


图10-3 典型的H.324系统

表10-2 H.261的YCbCr参数

参 数	CIF	QCIF
有效分辨率	$352 \times 288$	$176 \times 144$
帧率	29.97Hz	
YCbCr采样结构	4 : 2 : 0	
YCbCr编码形式	归一化量化DCT，每采样点8位表示	

H.261支持的两种图像（或帧）类型如下。

- 内码帧（Intra）或I帧：没有使用参考帧进行预测的帧。
- 预测帧（Inter）或P帧：基于前一帧预测得到的帧。

### 10.1.1 视频编码层

如图10-4所示，编码器基本的功能有预测、块变换和量化。

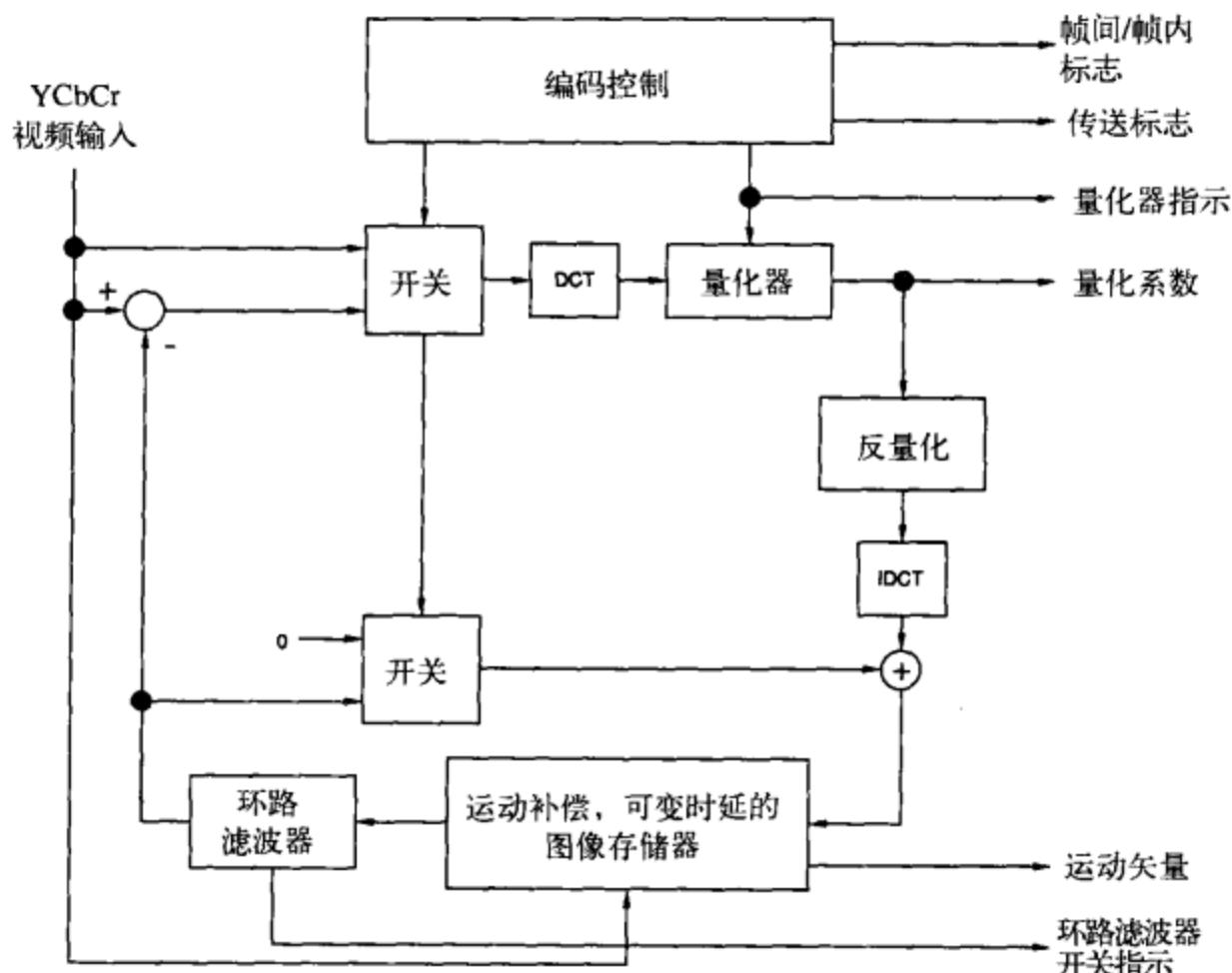


图10-4 典型的H.261编码器

预测误差（帧间模式）或者输入图像（帧内模式）被细分为多个8像素×8行的块，这些块被分成传送的和非传送的。如图10-5所示，4个亮度块（Y）和两个空间上对应的色差块（Cb和Cr）组成了一个16像素×16行的宏块（macroblock）。

H.261标准中没有给出模式选择和块传送的准则，这些准则可以作为编码策略的一部分动态地变化。编码器首先对传送块进行变换，并对变换后的系数进行量化，然后对量化后的系数进行可变长编码。

#### 1. 预测

通常，预测是在图像之间（inter-picture）进行，可以加入运动补偿和空域滤波器。采用预测的编码模式称为帧间编码（inter），而不采用预测的编码模式则称为帧内编码（intra）。

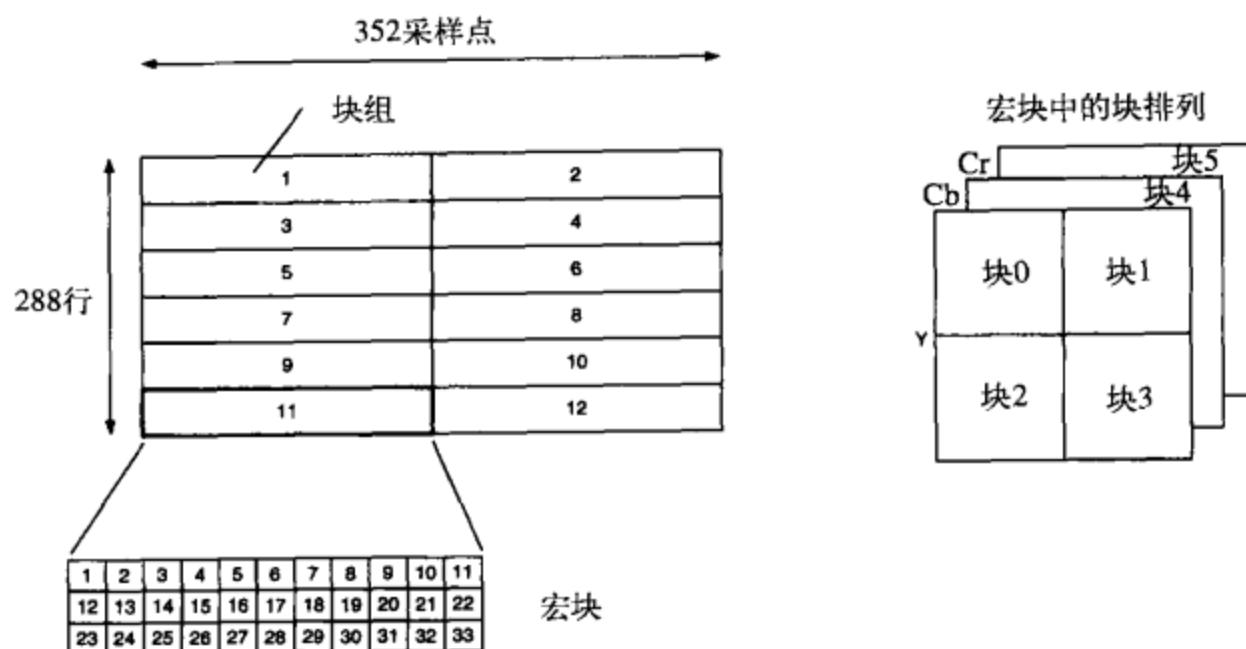


图10-5 H.261块组、宏块和块排列

## 2. 运动补偿

运动补偿在编码器中是可选的。解码器必须支持每个宏块拥有一个运动矢量。运动矢量是受限的，即所有参考像素都必须在已编码的图像区域内。

运动矢量的水平和垂直分量必须为不超过 $\pm 15$ 的整数值。宏块中所有4个Y块的运动矢量与宏块的运动矢量是相同的，而Cb和Cr块的运动矢量则为宏块运动矢量值的一半。

运动矢量的水平或垂直分量如果为正值，则表示预测是基于前一帧图像中位于被预测像素右方或者下方的像素。

## 3. 环路滤波器

可以采用二维空域滤波器来改进预测过程，该滤波器对 $8 \times 8$ 预测块内的像素进行操作。

滤波器可以分成水平和垂直函数，两者都是非递归的，且系数为0.25、0.5、0.25。特别地，在块边缘处，可能会有某一个抽头（tap）落入块区域之外，此时则将滤波器的系数改变为0、1、0。

对于一个宏块中的所有6个块，滤波器的开与关将根据宏块的类型来决定。

## 4. 离散余弦变换和反离散余弦变换

在编码器中，传送的块首先进行 $8 \times 8$ 的离散余弦变换（DCT）处理。反离散余弦变换（IDCT）的输出结果削波之后的范围在-256和255之间，采用9位编码表示。

H.261标准也没有定义计算变换的过程，但是反变换必须满足特定的误差容限。

## 5. 量化

在一个宏块中，除了intra-DC外，所有系数都采用相同的量化器。intra-DC通常采用步长为8且无死区的线性量化器量化。而其他系数则采用中心死区为0，步长为2, 4, 6, ..., 62之间任意值的线性量化器量化。

## 6. 重构图像削波

为了防止编码器和解码器环路中由于变换系数振幅的量化失真可能导致的运算溢出，H.261标准中采用了削波函数。由预测值和预测误差相加形成重构图像，削波函数用于对重构图像进行削波处理。削波器将像素值限定为小于等于0或者大于等于255。

## 7. 编码控制

尽管没有作为H.261标准的一部分而包含在其中，但是H.261标准还是采用了多个参数来控制编码视频数据的产生速率。这些参数包括：编码预处理过程、量化器、块重要性准则和时域亚采

470  
471

样等。时域亚采样是通过丢弃完整的图像来实现的。

### 8. 强制更新

为了防止IDCT变换的误差累积或者其他一些差错的影响，源编码采取了一种强制更新的方法。它规定每个宏块至少每传送132次就要以帧内模式传送一次（更新一次），以阻止编码中的误差累积。

## 10.1.2 视频比特流

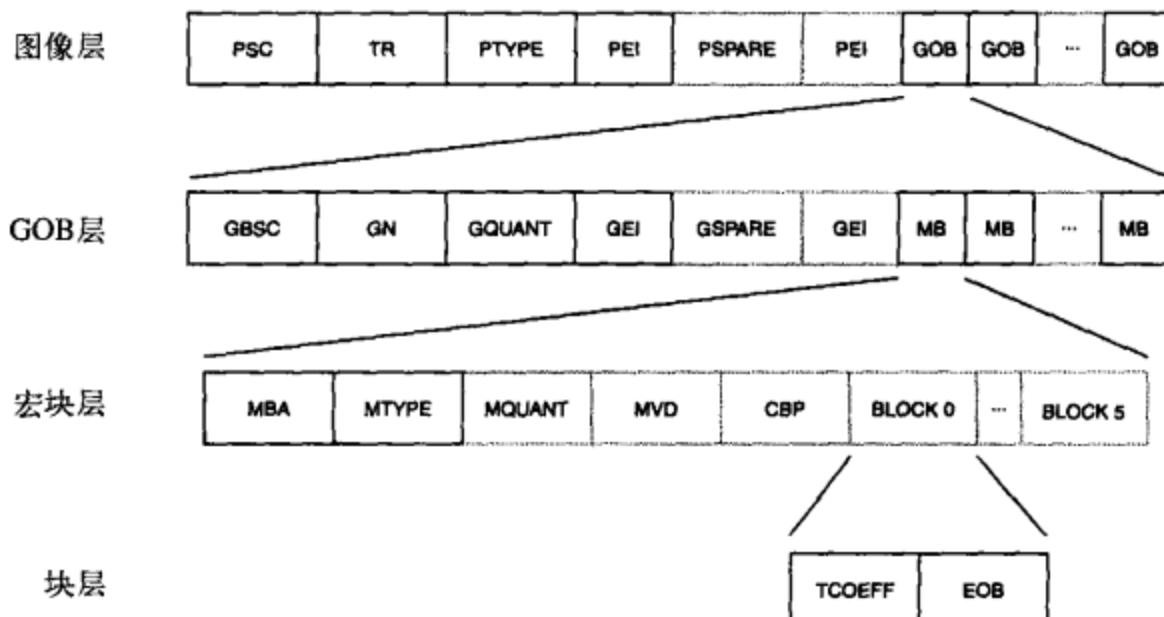
除非特别声明，视频比特流中最高有效位MSB (Most Significant Bit) 总是最先传送。MSB位于编码表的第1位，即最左边一位。同样，除非特别声明，所有的备用位均设置为1。

H.261的码流结构分为四个层次，从上到下依次是：

- 图像层 (Picture)
- 块组层GOB (Group of Blocks)
- 宏块层MB (Macroblock)
- 块层 (Block)

### 1. 图像层

如图10-6所示，图像数据由图像头和块组 (GOB) 数据组成。被丢弃图像的图像头将不需要传送。



472

图10-6 H.261视频比特流层次结构

#### □ 图像起始码 (PSC)

PSC (Picture Start Code) 长度为20位，值为0000 0000 0000 0001 0000。

#### □ 时间参考 (TR)

TR (Temporal Reference) 是一个5位的二进制数，表示32种可能值。TR的生成是通过将前一图像中的TR值递增1，然后再加上非传送图像（在29.97Hz频率下）数目而得到的。运算仅在5个最低有效位LSB (Last Significant Bit) 上进行。

#### □ 类型信息 (PTYPE)

PTYPE (Type Information) 采用6位编码来表示有关图像的信息：

位1表示分割屏幕指示器

0=关, 1=开

位2表示文档相机指示器

0=关,1=开

位3表示冻结图像解除

0=关,1=开

位4表示源格式

0=QCIF,1=CIF

位5表示可选的静止图像模式

0=开,1=关

位6表示备用

#### 额外插入信息 (PEI)

PEI (Extra Insertion Information) 采用1位表示, 当设置为1时表示接下来将出现可选数据字段。

#### 备用信息 (PSPARE)

如果PEI设置为1, 在PEI之后将紧随一个9位的额外插入信息, 其中8位为PSPARE (Spare Information), 第9位为PEI位, 表示是否有另一个9位的额外插入信息在之后, 如此反复。

### 2. 块组层

每一个图像都被分成了多个块组 (GOB), CIF图像由12个GOB组成, 而QCIF图像则由3个GOB组成 (见图10-5)。GOB由176样本×48行的Y块和对应的88×24的Cb和Cr块数据组成。

如图10-6所示, GOB数据由GOB头和宏块数据组成。在图10-5中, 在CIF或者QCIF中顺序编号的图像起始码之间, 每一个GOB头将被传送一次, 即使GOB中没有宏块数据也需传送。

#### 块组起始码 (GBSC)

GBSC (Group of Blocks Start Code) 长度为16位, 值为0000 0000 0000 0001。

#### 组编号 (GN)

GN (Group Number) 是一个4位二进制值, 表示块组的位置。这4位是图10-5中编号的二进制表示, 编号13、14、15保留以备将来使用。

#### 量化器信息 (GQUANT)

GQUANT (Quantizer Information) 是一个5位二进制值, 表示块组使用的量化器, GQUANT一直有效, 直到被接下来的MQUANT所取代。GQUANT取值为1~31。

#### 额外插入信息 (GEI)

GEI (Extra Insertion Information) 用1位表示, 当设置为1时, 表示接下来将出现可选数据字段。

#### GSpare

如果GEI设置为1, 则在GEI之后将紧随一个9位的额外插入信息, 其中8位为GSpare, 第9位为GEI位, 表示在之后是否有另一个9位的额外插入信息, 如此反复。

### 3. 宏块层

如图10-5所示, 一个GOB被分成33个宏块。每一个宏块由一个16样本×16行的Y块和对应的8×8的Cb块和Cr块组成。

宏块的数据由宏块头和块数据组成 (见图10-6)。

#### 宏块地址 (MBA)

MBA (Macroblock Address) 是一个变长的码字, 表示宏块在块组中的位置。图10-5中给出了宏块的传送顺序。对于块组GOB中的第一个宏块, 其MBA是图10-5中的绝对地址, 对于接下来的宏块, 它们的MBA是其绝对地址与最近传送的宏块的MBA之间的差值。表10-3中给出了MBA的码表。

表10-3 H.261中MBA的可变长编码表

MBA	编 码			MBA	编 码		
1	1			17	0000	0101	10
2	011			18	0000	0101	01
3	010			19	0000	0101	00
4	0011			20	0000	0100	11
5	0010			21	0000	0100	10
6	0001	1		22	0000	0100	011
7	0001	0		23	0000	0100	010
8	0000	111		24	0000	0100	001
9	0000	110		25	0000	0100	000
10	0000	1011		26	0000	0011	111
11	0000	1010		27	0000	0011	110
12	0000	1001		28	0000	0011	101
13	0000	1000		29	0000	0011	100
14	0000	0111		30	0000	0011	011
15	0000	0110		31	0000	0011	010
16	0000	0101	11	32	0000	0011	001
				33	0000	0011	000
MBA填充				0000	0001	111	
起始码				0000	0000	0000	0001

对于在紧接在GOB头之后或者已编码的宏块之后的填充位（称为MBA填充），也可以有一个码字。这个码字在解码器中将被丢弃。

表10-3中同时也给出了起始码的码字。MBA通常包含在被传送的宏块中。当宏块中没有包含图像的MBA信息时，该宏块将不会被传送。

#### □ 类型信息 (MTYPE)

MTYPE (Type Information) 是一个变长的码字，它包含有关于宏块和宏块中数据元素的信息。表10-4中列出了宏块类型、包含的元素和变长的码字。MTYPE通常也包含在被传送的宏块中。

表10-4 H.261中MTYPE的可变长编码表

预测模式	MQUANT	MVD	CBP	TCOEF	编 码		
intra				×	0001		
intra	×			×	0000	001	
inter			×	×	1		
inter	×		×	×	0000	1	
inter + MC		×			0000	0000	1
inter + MC		×	×	×	0000	0001	
inter + MC	×	×	×	×	0000	0000	01
inter + MC + FIL		×			001		
inter + MC + FIL		×	×	×	01		
inter + MC + FIL	×	×	×	×	0000	01	

#### □ 量化器 (MQUANT)

只有当MTYPE中指示MQUANT存在时，MQUANT (Quantizer) 才会出现。MQUANT是一

个5位的码字，它表示块组中当前和接下来的块要采用的量化器，MQUANT一直有效，直到被下一个MQUANT所取代。MQUANT的码字与GQUANT的码字是一致的。

#### □ 运动矢量数据 (MVD)

正如MTYPE所指示的那样，所有的运动补偿 (MC) 宏块都包含MVD (Motion Vector Data)。MVD是通过将当前宏块矢量减去前一个宏块矢量所获得的。当出现下列情况时，前一个宏块的矢量被视为0：

- (1) 估计宏块1、12和23的MVD时；
- (2) 当前宏块MBA与前一宏块的MBA差值小于1时；
- (3) 前一宏块的MTYPE不是运动补偿时。

运动矢量数据由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-5中给出了这些变长码。

表10-5 H.261中MVD的可变长编码表

矢量差分	编 码			矢量差分	编 码		
-16和16	0000	0011	001	1	010		
-15和17	0000	0011	011	2和-30	0010		
-14和18	0000	0011	101	3和-29	0001	0	
-13和19	0000	0011	111	4和-28	0000	110	
-12和20	0000	0100	001	5和-27	0000	1010	
-11和21	0000	0100	011	6和-26	0000	1000	
-10和22	0000	0100	11	7和-25	0000	0110	
-9和23	0000	0101	01	8和-24	0000	0101	10
-8和24	0000	0101	11	9和-23	0000	0101	00
-7和25	0000	0111		10和-22	0000	0100	10
-6和26	0000	1001		11和-21	0000	0100	010
-5和27	0000	1011		12和-20	0000	0100	000
-4和28	0000	111		13和-19	0000	0011	110
-3和29	0001	1		14和-18	0000	0011	100
-2和30	0011			15和-17	0000	0011	010
-1	011						
0	1						

#### □ 编码块模式 (CBP)

如果MTYPE指示CBP存在，将出现一个变长的CBP (Coded Block Pattern)。CBP表示宏块中哪一个块至少传送了一个变换系数。模式编码如下所示：

$$P_0 P_1 P_2 P_3 P_4 P_5$$

对于块[n]，其任意系数出现，则 $P_n=1$ ，否则 $P_n=0$ 。图10-5中给出了块编号方式 (十进制)。表10-6中给出了CBP编码的码字。

表10-6 H.261中CBP的可变长编码表

CBP	编 码		CBP	编 码	
60	111		62	0100	0
4	1101		24	0011	11
8	1100		36	0011	10

(续)

CBP	编 码		CBP	编 码	
16	1011		3	0011	01
32	1010		63	0011	00
12	1001	1	5	0010	111
48	1001	0	9	0010	110
20	1000	1	17	0010	101
40	1000	0	33	0010	100
28	0111	1	6	0010	011
44	0111	0	10	0010	010
52	0110	1	18	0010	001
56	0110	0	34	0010	000
1	0101	1	7	0001	1111
61	0101	0	11	0001	1110
2	0100	1	19	0001	1101
35	0001	1100	38	0000	1100
13	0001	1011	29	0000	1011
49	0001	1010	45	0000	1010
21	0001	1001	53	0000	1001
41	0001	1000	57	0000	1000
14	0001	0111	30	0000	0111
50	0001	0110	46	0000	0110
22	0001	0101	54	0000	0101
42	0001	0100	58	0000	0100
15	0001	0011	31	0000	0011 1
51	0001	0010	47	0000	0011 0
23	0001	0001	55	0000	0010 1
43	0001	0000	59	0000	0010 0
25	0000	1111	27	0000	0001 1
37	0000	1110	39	0000	0001 0
26	0000	1101			

#### 4. 块层

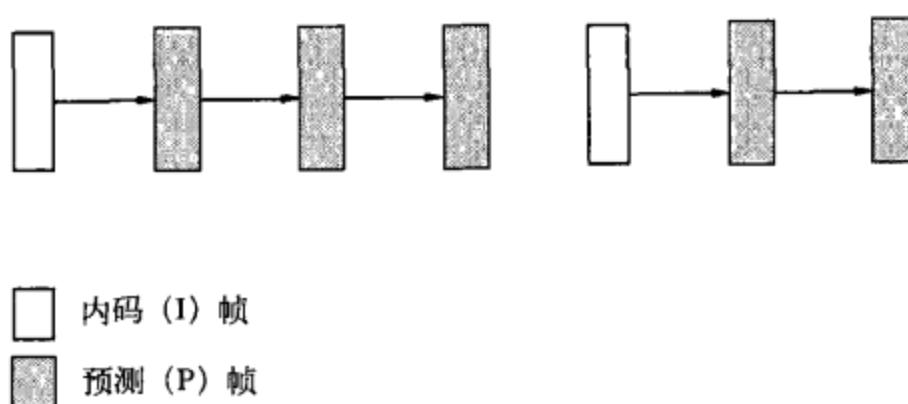
一个宏块由4个Y块和一个Cb块、一个Cr块组成（参见图10-5）。

如图10-6所示，一个8样本×8行的块的数据由变换系数的码字和块结束（EOB，end of block）标识组成。图10-5中给出了块传送的顺序。

##### □ 变换系数 (TCOEFF)

当MTYPE指示为intra时，宏块中的所有6个块的变换系数数据都将出现。否则，MTYPE和CBP将联合指出哪些块有系数数据传送。量化的DCT系数的传送顺序如图7-57所示。

如表10-7所示，大多数常见的连续零值（RUN）和紧接之后的非零值（LEVEL）的组合采用变长编码方式编码。因为CBP指示了哪些块没有系数数据，因此EOB不可能出现在第一个系数中。最后一位s表示非零值（LEVEL）的符号：0表示为正，1表示为负。



(RUN, LEVEL) 的其他组合采用一个20位的码字编码：其中6位表示转义 (ESC)，6位表示 RUN，8位表示LEVEL，如表10-8所示。

变长编码采用两个编码表：一个用于inter、inter + MC和inter + MC + FIL块中的第一个传送的 LEVEL；另一个用于除intra块中的第一个LEVEL之外的所有其他LEVEL，而对于intra块中的第一个LEVEL，则采用一个8位的定长编码。

除intra-DC外，所有的系数都有一个重构非零值 (reconstruction level, REC)，取值范围为 -2048~2047。重构非零值可通过下面的方程获得，其结果已经经过削波处理。QUANT取值从1到31，QUANT一直有效，直到被另一个GQUANT或者MQUANT所取代。

QUANT=奇数：

若 LEVEL>0

$$REC = QUANT \times (2 \times LEVEL + 1)$$

若 LEVEL<0

$$REC = QUANT \times (2 \times LEVEL - 1)$$

QUANT=偶数：

若 LEVEL>0

$$REC = (QUANT \times (2 \times LEVEL + 1)) - 1$$

若 LEVEL<0

$$REC = (QUANT \times (2 \times LEVEL - 1)) + 1$$

若 LEVEL=0

$$REC = 0$$

表10-7 H.261中TCOEFF的可变长编码表

RUN	LEVEL	编 码			
EOB		10			
0	1	1s	如果第一个系数在块中*		
0	1	11s	如果第一个系数不在块中		
0	2	0100	s		
0	3	0010	1s		
0	4	0000	110s		
0	5	0010	0110	s	
0	6	0010	0001	s	
0	7	0000	0010	10s	
0	8	0000	0001	1101	s

(续)

RUN	LEVEL	编 码			
0	9	0000	0001	1000	s
0	10	0000	0001	0011	s
0	11	0000	0001	0000	s
0	12	0000	0000	1101	0s
0	13	0000	0000	1100	1s
0	14	0000	0000	1100	0s
0	15	0000	0000	1011	1s
1	1	011s			
1	2	0001	10s		
1	3	0010	0101	s	
1	4	0000	0011	00s	
1	5	0000	0001	1011	s
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
2	1	0101	s		
2	2	0000	100s		
2	3	0000	0010	11s	
2	4	0000	0001	0100	s
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0100	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0011	0s		
4	2	0000	0011	11s	
4	3	0000	0001	0010	s
5	1	0001	11s		
5	2	0000	0010	01s	
5	3	0000	0000	1001	0s
6	1	0001	01s		
6	2	0000	0001	1110	s
7	1	0001	00s		
7	2	0000	0001	0101	s
8	1	0000	111s		
8	2	0000	0001	0001	s
9	1	0000	101s		
9	2	0000	0000	1000	1s
10	1	0010	0111	s	
10	2	0000	0000	1000	0s
11	1	0010	0011	s	
12	1	0010	0010	s	
13	1	0010	0000	s	
14	1	0000	0011	10s	
15	1	0000	0011	01s	

(续)

RUN	LEVEL	编 码			
16	1	0000	0010	00s	
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
ESC		0000	01		

注：在intra宏块中不采用。

对于intra-DC块，第一个系数通常为采用步长为8且无死区的量化器量化的变换值，其结果为一个8位编码值n。黑色（black）的编码值为0001 0000（16），白色（white）的编码值为1110 1011（235）。变换值1024的编码值为1111 1111。编码值0000 0000和1000 0000没有使用。除了255以外，所有编码值的解码值都为8n，而255的解码值为重构变换值1024。

477  
478

### 10.1.3 静止图像传送

H.261允许传送分辨率为当前所选视频格式分辨率4倍的静止图像。如果视频格式为QCIF，可以传送CIF分辨率的静止图像；如果视频格式为CIF，可以传送分辨率为 $704 \times 576$ 的静止图像。

表10-8 H.261 RUN和LEVEL编码

RUN	编 码	LEVEL	编 码
0	0000 00	-128	禁止
1	0000 01	-127	1000 0001
⋮	⋮	⋮	⋮
63	111111	-2	1111 1110
		-1	1111 1111
		0	禁止
		1	0000 0001
		2	0000 0010
		⋮	⋮
		127	0111 1111

## 10.2 H.263

ITU-T H.263改善了H.261在低码率情况下的视频质量。

H.263视频编码器提供了一个与其他信号（如H.223等）复用的独立数字视频比特流。视频解码器执行与编码器相反的过程。表10-9中列出了H.263有关YCbCr视频数据的主要规格。除了这些规格外，编解码器双方也可能协商一个定制的图像尺寸。图3-7中给出了4：2：0的YCbCr采样。

如下所示，H.263支持7种帧（或图像）类型，其中前两种是强制的（基准H.263）。

- Intra或I帧：没有采用参考帧进行预测的帧。
- Inter或P帧：基于前一个帧预测得到的帧。
- PB帧和改进的PB帧：一个代表两个帧的帧，且是基于前一个帧预测得到的帧。
- B帧：基于两个参考帧（前一帧和后一帧）预测得到的帧。
- EI帧：有一个同步帧的帧，同步帧大小与其相同或者稍小。
- EP帧：有两个参考帧（前一帧和同步帧）的帧。

表10-9 基准H.263的YCbCr参数

参 数	16CIF	4CIF	CIF	QCIF	SQCIF
有效分辨率 (Y)	$1408 \times 1152$	$704 \times 576$	$352 \times 288$	$176 \times 144$	$128 \times 96$
帧率			29.97Hz		
YCbCr采样结构			4 : 2 : 0		
YCbCr编码形式			归一化量化PCM，每采样点8位表示		

### 10.2.1 视频编码层

图10-8给出了一个典型的H.263编码器结构图，其基本的功能有预测、块变换和量化。

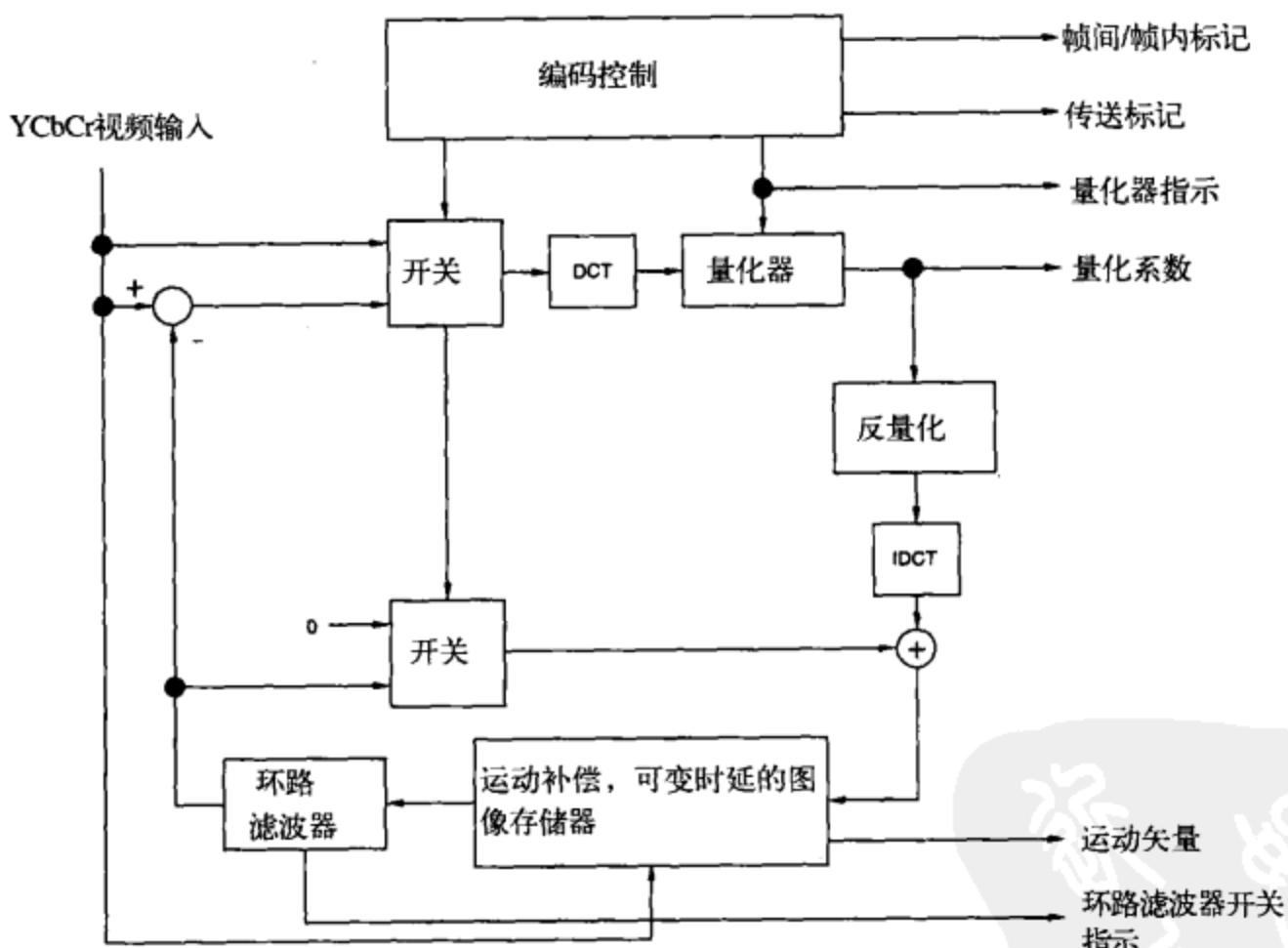


图10-8 典型的基准H.263编码器

预测误差或者输入图像被细分为多个 $8 \times 8$ 的块，这些块被分为传送的和非传送的。如图10-9所示，4个亮度块(Y)和两个空间上对应的色度块(Cb和Cr)组成一个 $16 \times 16$ 的宏块。

H.263标准中没有给出模式选择和块传送的准则，这些准则可以作为编码策略的一部分动态地改变。编码器首先对发送块进行变换，然后对变换后的系数进行量化，最后对量化结果进行可变长编码。

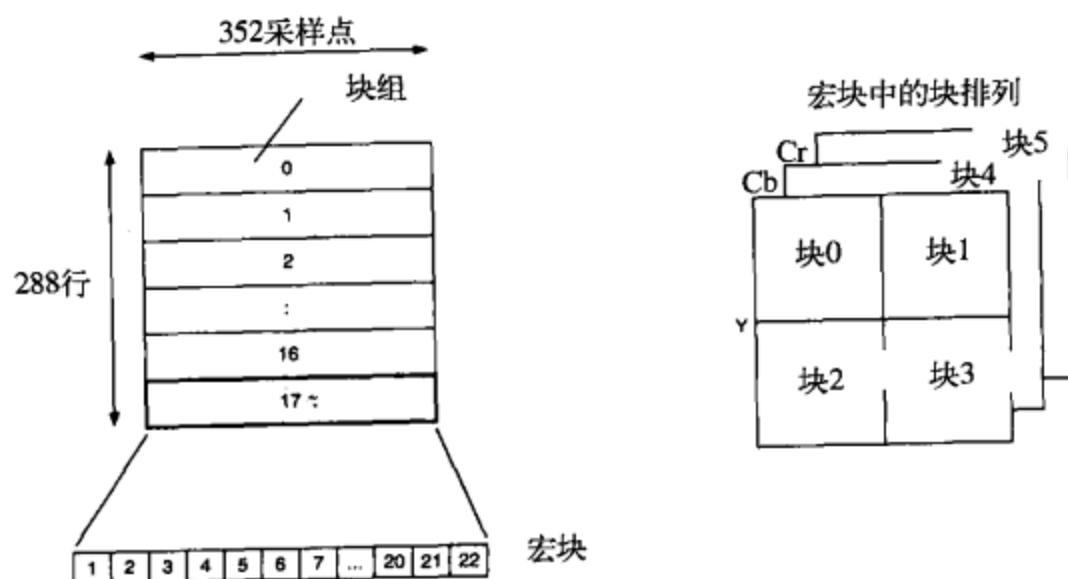


图10-9 H.263块组、宏块和块排列

### 1. 预测

预测发生在图像之间 (inter-picture)，可能包括运动补偿。采用预测的编码模式称为帧间编码 (inter)，而不采用预测的编码模式则称为帧内编码(intra)。

帧内编码模式用于图像层 (对于intra是I帧，对于inter是P帧) 或者P帧中的宏块层。在可选的PB帧模式中，B帧通常采用帧间编码模式。

### 2. 运动补偿

运动补偿在编码器中是可选的。解码器必须支持每个宏块拥有一个运动矢量 (在可选的高级预测或者去块滤波器模式中，每个宏块有一个或4个运动矢量)。

在可选的PB帧模式中，每一个宏块可以有一个附加的矢量。在可选的改进的PB帧模式中，每一个宏块可以包含一个附加的前向运动矢量。在可选的B帧模式中，宏块在传送时可以有一个前向和后向运动矢量。

对于基准H.263，运动矢量是受限的，即所有被用作参考的采样点都必须在已编码的图像区域内，而对于许多可选模式，这一限制已经去除了。运动矢量的水平和垂直分量必须为 $-16\sim15.5$ 范围内的整数值或半整数值。某些可选模式将这一范围增大到 $[-31.5, +31.5]$ 或者 $[-31.5, +30.5]$ 。

运动矢量的水平或垂直分量如果为正值，则表示预测是基于前一帧图像中位于被预测采样点右方或者下方的采样点。然而，对于B帧中的后向运动矢量，运动矢量的水平或垂直分量如果为正值，表示预测是基于下一帧图像中位于被预测的采样点左方或者上方的采样点。

### 3. 量化

对于第一个intra系数，量化器的数目是1，其他所有系数的量化器数目为31。在一个宏块中，除了intra-DC外，所有系数都采用相同的量化器。通常intra-DC采用的步长为8，且量化采用的是无死区的线性量化器。而其他系数则采用中心死区为0，步长为2, 4, 6, …, 62之间的任意值的线性量化器量化。

### 4. 编码控制

H.263标准采用了多个参数来控制编码视频数据的产生速率，这些参数包括：编码器预处理过程、量化器、块重要性准则和时域亚采样等。但是，这些参数并没有作为其一部分而包含在H.263标准中。

### 5. 强制更新

这个功能是通过强制使用编码算法中的帧内模式 (Intra) 而实现的。为了防止IDCT变换的误

差累积或者其他一些差错的影响，源编码采取了一种强制更新的方法。它规定每个宏块至少每传送132次就要以帧内模式传送一次（更新一次），以防止编码中的误码累积。

### 10.2.2 视频比特流

除非特别声明，视频比特流的最高有效位MSB总是最先传送。MSB位于编码表的第一位，即最左边一位。同样，除非特别声明，所有的备用位均设置为1。

H.263的视频流复用结构划分为四个层次，从上到下依次是：

- 图像层P (Picture)
- 块组层GOB (Group of Blocks) 或者条带 (Slice)
- 宏块层MB (Macroblock)
- 块层B (Block)

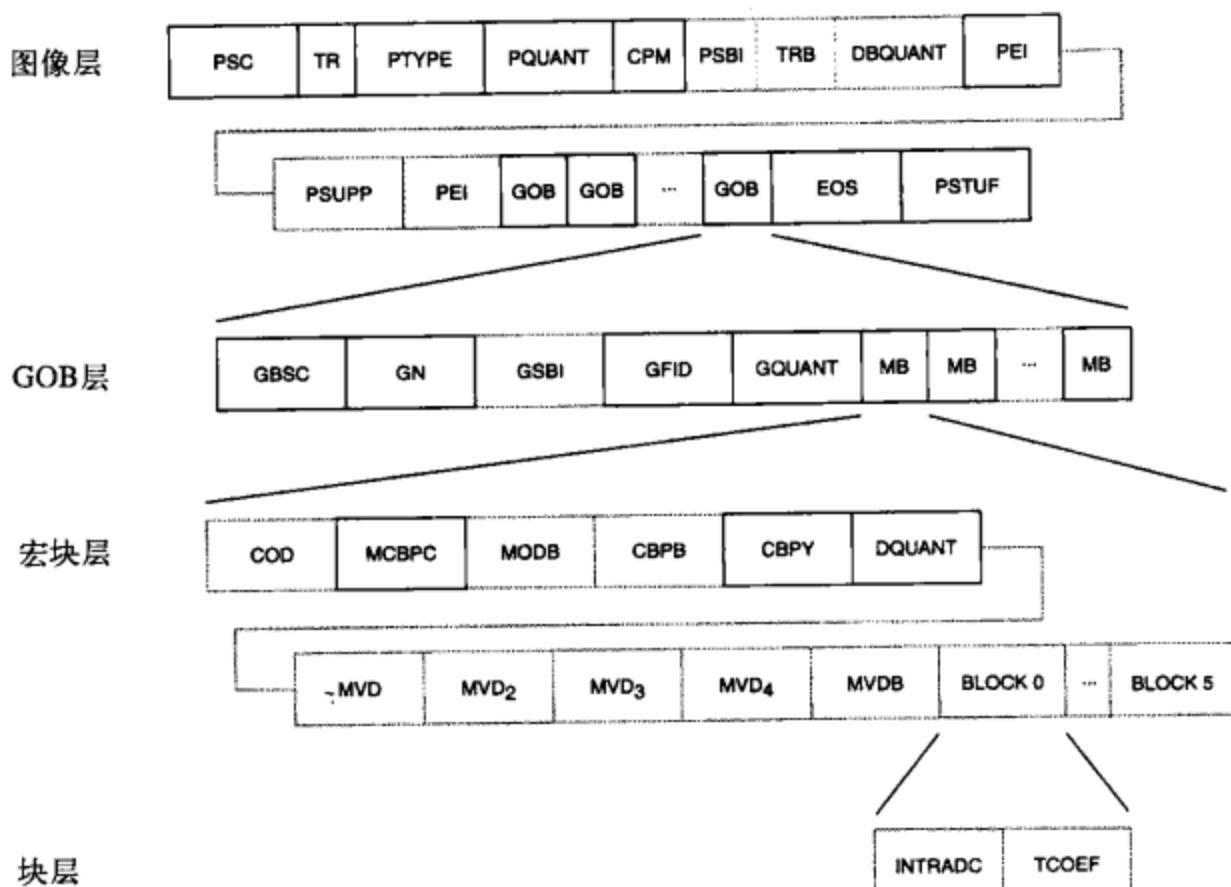


图10-10 基准H.263的视频比特流层次结构（在图像层无可选项PLUSPTYPE相关的字段）

#### 1. 图像层

如图10-10所示，图像数据由图像头、块组数据和序列结束（EOS, end-of-sequence）标识以及填充位（PSTUF）组成。被丢弃图像的图像头不需要传送。

##### 图像起始码 (PSC)

PSC (Picture Start Code) 长度为22位，值为0000 0000 0000 0000 1 00000。PSC必须是字节对齐的；因此，起始码之前必须添加0~7个0位，以确保起始码的第一位是字节的第一位，即最高有效位。

##### 时间参考 (TR)

TR (Temporal Reference) 是一个8位的二进制数，表示256种可能值。TR的生成是通过将前一帧传送的图像中的TR值递增1，然后再加上非传送图像（在29.97Hz频率下）数目而得到的。运算仅在8个LSB上进行。

如果指定了定制图像时钟频率 (PCF)，那么采用扩展时间参考 (ETR) 和TR共同组成一个10位的编码，其中TR存储在8个LSB上，ETR存储在两个MSB上。在这种情况下，运算在10个LSB上进行。

在PB帧和改进的PB帧模式中，TR仅对P帧有效。

#### □ 类型信息 (PTYPE)

PTYPE (Type Information) 采用13位来表示有关图像的信息：

位1表示1

位2表示0

位3表示分割屏幕指示器

0=关, 1=开

位4表示文档相机指示器

0=关, 1=开

位5表示冻结图像解除

0=关, 1=开

位6~8表示源格式

000=保留

001=SQCIF

010=QCIF

011=CIF

100=4CIF

101=16CIF

110=保留

111=扩展的PTYPE

如果位6~8不是111，那么下面5位将出现在PTYPE中：

位9表示图像编码类型

0=帧内, 1=帧间

位10表示可选的非限制运动矢量模式

0=关, 1=开

位11表示可选的基于语法的算术编码模式

0=关, 1=开

位12表示可选的高级预测模式

0=关, 1=开

位13表示可选的PB帧模式

0=正常图像, 1=PB帧

如果第9位设置位0，那么第13位必须设置为0。第10~13位是编码器和解码器之间协商的可选模式。

#### □ 量化器信息 (PQUANT)

PQUANT (Quantizer Information) 是一个5位二进制编码（取值为1~31），它表示图像采用的量化器，PQUANT一直有效，直到被下一个GQUANT或者DQUANT所取代。

#### □ 连续出现多点 (CPM)

CPM (Continuous Presence Multipoint) 是一个一位的值，表示编码器是否采用可选的连续出现多点和视频复用模式，0=关, 1=开。如果PLUSPTYPE不存在，那么CPM紧接着PQUANT，否则CPM将紧接着PLUSPTYPE。

**□ 图像子比特流指示器 (PSBI)**

PSBI (Picture Sub-Bitstream Indicator) 是一个可选的2位二进制编码，只有当CPM指示采用可选的连续存在多点和视频复用模式时，PSBI才会出现。

**□ PB帧中B帧时间参考 (TRB)**

当PTYPE或者PLUSTYPE指示帧类型为PB帧或者改进的PB帧时，PB帧中的B帧TRB才会出现。TRB是一个3位或者5位表示数 $[n + 1]$ 的二进制编码，其中 $n$ 为从最近的I帧（或P帧）或者PB帧（或改进的PB帧）的P部分开始和PB帧（或改进的PB帧）的B部分之前非传送图像的数目（时钟频率为29.97Hz或者由CPCFC指示的定制图像时钟频率）。当采用定制图像时钟频率时，TRB的值扩展到5位。

当图像时钟频率为29.97Hz时，最大非传送图像数目为6，而采用定制图像时钟频率时，这个数为30。

**□ PB帧中B帧量化器信息 (DBQUANT)**

当PTYPE或者PLUSTYPE指示帧类型为PB帧或改进的PB帧时，将会出现DBQUANT (Quantizer Information for B Frames in PB Frames)。DBQUANT是一个2位码字，表示QUANT和BQUANT之间的关系，如表10-10所示。除法运算采用截切方式。BQUANT范围为1~31。如果运算结果低于1或者高于31，将BQUANT分别截切到1和31。

表10-10 基准H.263中DBQUANT编码和QUANT/BQUANT关系

DBQUANT	BQUANT	DBQUANT	BQUANT
00	(5*QUANT)/4	10	(7*QUANT)/4
01	(6*QUANT)/4	11	(8*QUANT)/4

**□ 额外插入信息 (PEI)**

PEI (Extra Insertion Information) 采用1位表示，当设置为1时表示接下来将出现补充的增强信息PSUPP数据域。

**□ 补充的增强信息 (PSUPP)**

如果PEI设置为1，在PEI之后将紧随一个9位的额外插入信息，其中8位为PSUPP (Supplemental Enhancement Information)，第9位为PEI位，表示在之后是否有另一个9位的额外插入信息，如此反复。

**□ 序列结束 (EOS)**

EOS (End of Sequence) 是一个值为0000 0000 0000 0000 1 11111的22位字码。EOS必须为字节对齐，通过在编码之前插入0~7个零位，从而确保EOS编码的第一位为字节的第一位，并且是最有效位。

**□ 填充 (PSTUF)**

PSTUF (Stuffing) 是一个零位的变长字码。PSTUF的最后一位必须是字节的最后一一位并且是最有效位。

## 2. 块组层

如图10-9所示，每一个图像被细分成多个GOB。对于SQCIF、QCIF和CIF分辨率图像，一个GOB由16行组成，对于4CIF分辨率和16CIF分辨率的图像，一个GOB分别由32行和64行组成。因此，一个SQCIF、QCIF和CIF图像分别包含6 (96/16) 个、9 (144/16) 个和18 (288/16) 个GOB，每个GOB有一列宏块数据。而一个4CIF图像则有18 (576/32) 个GOB，每一个GOB有两列宏块数据。一个16CIF图像也有18 (1152/64) 个GOB，每一个GOB带有4行的宏块数据。GOB的编号方

式是从图像的顶部开始编号为0，然后垂直向下递增。

如图10-10所示，每一个GOB的数据由GOB头和宏块数据组成。宏块数据以递增的宏块编号顺序传送。对于每一个图像中编号为0的GOB，编码器不需要传送GOB头。解码器可以通知编码器仅传送非空的GOB头。

#### 块组起始码 (GBSC)

GBSC (Group of Blocks Start Code) 长度为17位，值为0000 0000 0000 0000 1。GBSC必须是字节对齐的，因此，必须在起始码之前加入0~7个零位，以确保起始码的第一位是字节的第一位，并且是最有效位。

#### 组编号 (GN)

GN (Group Number) 是一个5位二进制数，表示GOB的编号。组编号1~17用于标准图像格式。组编号1~24用于定制图像格式。组编号16~29被复制到条带头中。组编号30用于子比特流结束EOSBS (end of sub-bitstream) 指示符中，而组编号31则用于序列结束码EOS (end of sequence) 中。

#### GOB子比特流指示器 (GSBI)

GSBI (GOB Sub-Bitstream Indicator) 是一个2位的二进制编码，表示当前GOB到下一个图像或者下一个GOB起始码之间的子比特流数目。GSBI只有在采用连续出现多点和视频复用 (CPM) 模式时才会出现。

#### GOB帧ID (GFID)

GFID用来表示GOB中的帧ID，长度为2位。对于一个给定的帧，在每一个GOB头（或者条带头）中都必须是同一个值。通常，如果当前图像的PTYPE与前一个图像头中的PTYPE是相同的，那么当前帧的GFID必须与前一个帧的GFID一致；否则，当前帧的GFID必须采用与前一个帧GFID不同的值。

#### 量化器信息 (GQUANT)

GQUANT (Quantizer Information) 是一个5位二进制数，取值范围为1~31，表示块组中采用的量化器。GQUANT一致有效，直到被下一个GQUANT或者DQUANT所取代。

### 3. 宏块层

如图10-9所示，每一个GOB被细分成多个宏块。每一个宏块由 $16 \times 16$ 的Y块和对应的 $8 \times 8$ 的Cb和Cr块组成。宏块的编号方式是从左到右、从上往下递增的。宏块数据按照递增的宏块编号顺序传送。

宏块的数据由一个宏块头和块数据组成（参见图10-10）。

#### 编码宏块指示 (COD)

COD (Coded Macroblock Indication) 采用一位来表示块是否已被编码。0表示已经编码，1表示没有编码，并且剩下的宏块层为空。只有当图像不是采用帧间编码模式时COD才会出现。

如果没有编码，解码器解码时将宏块当作一个整块运动矢量为零且没有系数数据的帧间块 (inter-block) 来处理。

#### 宏块类型和色度编码块模式 (MCBPC)

MCBPC (Macroblock Type and Coded Block Pattern for Chrominance) 是一个变长的码字，表示宏块层类型和Cb与Cr的编码块模式。

表10-11和表10-12中列出了MCBPC的码字。填充位也有一个码字，但是在解码器解码时应该丢弃。在某些情况下，填充位不能出现在图像的第一个宏块中，这是为了避免与起始码混淆。表10-13和表10-14中列出了宏块类型 (MB类型)。

表10-11 基准H.263中I帧的MCBPC可变长编码表

宏块类型	CBPC (Cb, Cr)	编 码		
3	0, 0	1		
3	0, 1	001		
3	1, 0	010		
3	1, 1	011		
4	0, 0	0001		
4	0, 1	0000	01	
4	1, 0	0000	10	
4	1, 1	0000	11	
填充		0000	0000	1

表10-12 基准H.263中P帧的MCBPC可变长编码表

宏块类型	CBPC (Cb, Cr)	编 码			
0	0, 0	1			
0	0, 1	0011			
0	1, 0	0010			
0	1, 1	0001	01		
1	0, 0	011			
1	0, 1	0000	111		
1	1, 0	0000	110		
1	1, 1	0000	0010	1	
2	0, 0	010			
2	0, 1	0000	101		
2	1, 0	0000	100		
2	1, 1	0000	0101		
3	0, 0	0001	1		
3	0, 1	0000	0100		
3	1, 0	0000	0011		
3	1, 1	0000	011		
4	0, 0	0001	00		
4	0, 1	0000	0010	0	
4	1, 0	0000	0001	1	
4	1, 1	0000	0001	0	
填充		0000	0000	1	
5	0, 0	0000	0000	010	
5	0, 1	0000	0000	0110	0
5	1, 0	0000	0000	0111	0
5	1, 1	0000	0000	0111	1

表10-13 基准H.263中宏块类型和标准帧包含的数据

帧类型	宏块类型	名 称	COD	MCBPC	CBPY	DQUANT	MVD	MVD <sub>2,4</sub>
inter	没有编码	—	×					
inter	0	inter	×	×	×		×	
inter	1	inter+q	×	×	×	×	×	
inter	2	inter4v	×	×	×		×	×
inter	3	intra	×	×	×			

(续)

帧类型	宏块类型	名 称	COD	MCBPC	CBPY	DQUANT	MVD	MVD <sub>2-4</sub>
inter	4	intra+q	×	×	×	×		
inter	5	inter4v+q	×	×	×	×	×	×
inter	填充	—	×	×				
intra	3	intra		×	×			
intra	4	intra+q		×	×	×		
intra	填充	—		×				

表10-14a 基准H.263中宏块类型和PB帧包含的数据

帧类型	宏块类型	名 称	COD	MCBPC	MODB	CBPY
inter	没有编码	—	×			
inter	0	inter	×	×	×	×
inter	1	inter+q	×	×	×	×
inter	2	inter4v	×	×	×	×
inter	3	intra	×	×	×	×
inter	4	intra+q	×	×	×	×
inter	5	inter4v+q	×	×	×	×
inter	填充	—	×	×		

表10-14b 基准H.263中宏块类型和PB帧包含的数据

帧类型	宏块类型	名 称	CBPB	DQUANT	MVD	MVDB	MVD <sub>2-4</sub>
inter	没有编码	—					
inter	0	inter	×		×	×	
inter	1	inter+q	×	×	×	×	
inter	2	inter4v	×		×	×	×
inter	3	intra	×		×	×	
inter	4	intra+q	×	×	×	×	
inter	5	inter4v+q	×	×	×	×	
inter	填充	—					

色度编码块模式 (CBPC) 表示对于Cb或者Cr是否有一个非intra-DC变换系数被传送。当CBPC为1时，表示在块中 (Cb或者Cr) 出现了一个非intra-DC系数。

#### □ B块的宏块模式 (MODB)

对于宏块类型为0~4的宏块，如果PTYPE指示为PB帧，将出现B块的宏块模式MODB (Macroblock Mode for Blocks)。MODB是一个变长的码字，表示对于当前宏块，B系数或运动矢量是否被传送。表10-15列出了MODB的码字。对于改进的PB帧，MODB采用了与PB帧不一样的编码方式。

表10-15 基准H.263的MODB可变长编码表

CBPB	MVDB	编 码
		0
	×	10
×	×	11

□ B块编码块模式 (CBPB)

如果MODB指示有CBPB，将出现一个6位长的CBPB (Coded Block Pattern for B Block)。CBPB指示宏块中哪一个块至少有一个变换系数被传送。模式编码方式如下：

$P_0P_1P_2P_3P_4P_5$

对于块[n]，若任意系数出现，则 $P_n=1$ ，否则 $P_n=0$ 。图10-9中给出了块编号方式（十进制）。

□ 亮度编码块模式 (CBPY)

CBPY (Coded Block Pattern for Luminance) 是一个变长的码字，指定宏块中至少有一个非intra-DC变换系数被传送的Y块。然而，在先进帧内编码模式中，intra-DC采用与其他系数相同的方式来指示。

表10-16列出了CBPY的编码。对Y块，若任意非intra-DC系数出现，则 $Y_n=1$ 。图10-9中给出了Y块的编号方式（十进制）。

表10-16 基准H.263中CBPY的可变长编码表

CBPY (Y0, Y1, Y2, Y3)		编 码	
intra	inter		
0, 0, 0, 0	1, 1, 1, 1	0011	
0, 0, 0, 1	1, 1, 1, 0	0010	1
0, 0, 1, 0	1, 1, 0, 1	0010	0
0, 0, 1, 1	1, 1, 0, 0	1001	
0, 1, 0, 0	1, 0, 1, 1	0001	1
0, 1, 0, 1	1, 0, 1, 0	0111	
0, 1, 1, 0	1, 0, 0, 1	0000	10
0, 1, 1, 1	1, 0, 0, 0	1011	
1, 0, 0, 0	0, 1, 1, 1	0001	0
1, 0, 0, 1	0, 1, 1, 0	0000	11
1, 0, 1, 0	0, 1, 0, 1	0101	
1, 0, 1, 1	0, 1, 0, 0	1010	
1, 1, 0, 0	0, 0, 1, 1	0100	
1, 1, 0, 1	0, 0, 1, 0	1000	
1, 1, 1, 0	0, 0, 0, 1	0110	
1, 1, 1, 1	0, 0, 0, 0	11	

□ 量化器信息 (DQUANT)

DQUANT是一个两位的码字，表示QUANT中值的改变。表10-17给出了该码字的不同值。

QUANT取值范围为1~31。如果DQUANT指示QUANT的值发生改变，并且改变后的结果小于

489

1或者大于31，那么分别将其设置为1和31。

表10-17 基准H.263中QUANT差分值的DQUANT编码

QUANT的差分值	DQUANT	QUANT的差分值	DQUANT
-1	00	1	10
-2	01	2	11

□ 运动矢量数据 (MVD)

在PB帧模式中，所有的宏块间和块内都包含MVD (Motion Vector Data)。

运动矢量数据由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了变长码。对于非限制运动矢量模式，可能会采用其他运动矢量编码方式。

### □ 运动矢量数据

在高级预测模式或者去块滤波器模式过程中，如果PTYPE和MCBPC指示三个码字MVD<sub>2</sub>、MVD<sub>3</sub>和MVD<sub>4</sub>存在，这三个运动矢量数据将会出现。每一个运动矢量数据都由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了这些变长码。

### □ B宏块运动矢量数据 (MVDB)

在PB帧和改进的PB帧模式过程中，如果MODB指示MVDB存在，MVDB将会出现。MVDB由一个表示水平分量的变长码字和一个表示垂直分量的变长码字组成。表10-18中给出了这些变长码。

表10-18 基准H.263中MVD、MVD<sub>2-4</sub>和MVDB的可变长编码表

矢量差分		编 码			
-16	16	0000	0000	0010	1
-15.5	16.5	0000	0000	0011	1
-15	17	0000	0000	0101	
-14.5	17.5	0000	0000	0111	
-14	18	0000	0000	1001	
-13.5	18.5	0000	0000	1011	
-13	19	0000	0000	1101	
-12.5	19.5	0000	0000	1111	
-12	20	0000	0001	001	
-11.5	20.5	0000	0001	011	
-11	21	0000	0001	101	
-10.5	21.5	0000	0001	111	
-10	22	0000	0010	001	
-9.5	22.5	0000	0010	011	
-9	23	0000	0010	101	
-8.5	23.5	0000	0010	111	
-8	24	0000	0011	001	
-7.5	24.5	0000	0011	011	
-7	25	0000	0011	101	
-6.5	25.5	0000	0011	111	
-6	26	0000	0100	001	
-5.5	26.5	0000	0100	011	
-5	27	0000	0100	11	
-4.5	27.5	0000	0101	01	
-4	28	0000	0101	11	
-3.5	28.5	0000	0111		
-3	29	0000	1001		
-2.5	29.5	0000	1011		
-2	30	0000	111		
-1.5	30.5	0001	1		
-1	31	0011			
-0.5	31.5	011			
0		1			

(续)

矢量差分		编 码			
0.5	-31.5	010			
1	-31	0010			
1.5	-30.5	0001	0		
2	-30	0000	110		
2.5	-29.5	0000	1010		
3	-29	0000	1000		
3.5	-28.5	0000	0110		
4	-28	0000	0101	10	
4.5	-27.5	0000	0101	00	
5	-27	0000	0100	10	
5.5	-26.5	0000	0100	010	
6	-26	0000	0100	000	
6.5	-25.5	0000	0011	110	
7	-25	0000	0011	100	
7.5	-24.5	0000	0011	010	
8	-24	0000	0011	000	
8.5	-23.5	0000	0010	110	
9	-23	0000	0010	100	
9.5	-22.5	0000	0010	010	
10	-22	0000	0010	000	
10.5	-21.5	0000	0001	110	
11	-21	0000	0001	100	
11.5	-20.5	0000	0001	010	
12	-20	0000	0001	000	
12.5	-19.5	0000	0000	1110	
13	-19	0000	0000	1010	
13.5	-18.5	0000	0000	1000	
14	-18	0000	0000	0110	
14.5	-17.5	0000	0000	0100	
15	-17	0000	0000	0100	
15.5	-16.5	0000	0000	0011	0

#### 4. 块层

在非PB帧模式下，一个宏块由4个Y块和一个Cb块、一个Cr块组成（参见图10-9）。如图10-10所示，一个8样本×8行的块的数据由intra-DC系数和变换系数的码字组成。图10-9中给出了块传送的顺序。

而在PB模式下，一个宏块则由4个Y块和一个Cb块、一个Cr块和6个B块的数据组成。

量化的DCT系数将按图7-57中的顺序传送。在修改量化(modified quantization)模式中，量化的DCT系数将按图7-58中的顺序传送。

##### □ 块内DC系数 (intra-DC)

intra-DC (DC Coefficient for Intra-Blocks) 是一个8位的码字。表10-19中列出了intra-DC的值及其对应的重构非零值。

在非PB帧模式下，如果MCBPC指示宏块的类型为3或者4，intra-DC系数将出现在宏块中的每一个块中。而在PB帧模式下，如果MCBPC指示宏块的类型为3或者4，intra-DC系数将出现在宏块中的每一个P块中（intra-DC系数不会出现在B块中）。

表10-19 基准H.263中intra DC的重构非零值

intra DC值	重构非零值
0000 0000	不采用
0000 0001	8
0000 0010	16
0000 0011	24
⋮	⋮
0111 1111	1016
1111 1111	1024
1000 0001	1032
⋮	⋮
1111 1101	2024
1111 1110	2032

#### □ 变换系数 (TCOEF)

在非PB帧模式下，TCOEF (Transform Coefficient) 由MCBPC或者CBPY决定是否存在。在PB帧模式下，TCOEF由CBPB决定是否在B块中出现。

事件由最后一个非零系数指示（如果块中还有其他非零系数，LAST = 0；否则LAST = 1）、非零系数前连续零的个数 (RUN) 和非零系数值 (LEVEL) 组合而成。

如表10-20所示，大多数常见事件采用变长码编码。s位指示非零系数的符号，0表示正，1表示负。

490  
494

表10-20 基准H.263中TCOEF的可变长编码表

LAST	RUN	LEVEL	编 码		
0	0	1	10s		
0	0	2	1111	s	
0	0	3	0101	01s	
0	0	4	0010	111s	
0	0	5	0001	1111	s
0	0	6	0001	0010	1s
0	0	7	0001	0010	0s
0	0	8	0000	1000	01s
0	0	9	0000	1000	00s
0	0	10	0000	0000	111s
0	0	11	0000	0000	110s
0	0	12	0000	0100	000s
0	1	1	110s		
0	1	2	0101	00s	
0	1	3	0001	1110	s
0	1	4	0000	0011	11s
0	1	5	0000	0100	001s

(续)

LAST	RUN	LEVEL	编 码		
0	1	6	0000	0101	0000
0	2	1	1110	s	
0	2	2	0001	1101	s
0	2	3	0000	0011	10s
0	2	4	0000	0101	0001
0	3	1	0110	1s	
0	3	2	0001	0001	1s
0	3	3	0000	0011	01s
0	4	1	0110	0s	
0	4	2	0001	0001	0s
0	4	3	0000	0101	0010
0	5	1	0101	1s	
0	5	2	0000	0011	00s
0	5	3	0000	0101	0011
0	6	1	0100	11s	
0	6	2	0000	0010	11s
0	6	3	0000	0101	0100
0	7	1	0100	10s	
0	7	2	0000	0010	10s
0	8	1	0100	01s	
0	8	2	0000	0010	01s
0	9	1	0100	00s	
0	9	2	0000	0010	00s
0	10	1	0010	110s	
0	10	2	0000	0101	0101
0	11	1	0010	101s	
0	12	1	0010	100s	
0	13	1	0001	1100	s
0	14	1	0001	1011	s
0	15	1	0001	0000	1s
0	16	1	0001	0000	0s
0	17	1	0000	1111	1s
0	18	1	0000	1111	0s
0	19	1	0000	1110	1s
0	20	1	0000	1110	0s
0	21	1	0000	1101	1s
0	22	1	0000	1101	0s
0	23	1	0000	0100	010s
0	24	1	0000	0100	011s
0	25	1	0000	0101	0110
0	26	1	0000	0101	0111
1	0	1	0111	s	
1	0	2	0000	1100	1s
1	0	3	0000	0000	101s

(续)

LAST	RUN	LEVEL	编 码		
1	1	1	0011	11s	
1	1	2	0000	0000	100s
1	2	1	0011	10s	
1	3	1	0011	01s	
1	4	1	0011	00s	
1	5	1	0010	011s	
1	6	1	0010	010s	
1	7	1	0010	001s	
1	8	1	0010	00s	
1	9	1	0001	1010	s
1	10	1	0001	1001	s
1	11	1	0001	1000	s
1	12	1	0001	0111	s
1	13	1	0001	0110	s
1	14	1	0001	0101	s
1	15	1	0001	0100	s
1	16	1	0001	0011	s
1	17	1	0000	1100	0s
1	18	1	0000	1011	1s
1	19	1	0000	1011	0s
1	20	1	0000	1010	1s
1	21	1	0000	1010	0s
1	22	1	0000	1001	1s
1	23	1	0000	1001	0s
1	24	1	0000	1000	1s
1	25	1	0000	0001	11s
1	26	1	0000	0001	10s
1	27	1	0000	0001	01s
1	28	1	0000	0001	00s
1	29	1	0000	0100	100s
1	30	1	0000	0100	101s
1	31	1	0000	0100	110s
1	32	1	0000	0100	111s
1	33	1	0000	0101	1000
1	34	1	0000	0101	1001
1	35	1	0000	0101	1010
1	36	1	0000	0101	1011
1	37	1	0000	0101	1100
1	38	1	0000	0101	1101
1	39	1	0000	0101	1110
1	40	1	0000	0101	1111
ESC			0000	011	

(LAST, RUN, LEVEL) 的其他组合采用一个22位的字来编码：其中7位表示转义 (ESC)，

1位为LAST，6位为RUN，剩余8位为LEVEL。表10-21中给出了RUN和LEVEL的编码。编码1000 0000被禁止使用，除非在修改的量化模式下。

表10-21 基准H.263的RUN、LEVEL编码

RUN	编 码	LEVEL	编 码
0	0000 00	-128	禁止
1	0000 01	-127	1000 0001
:	:	:	:
63	1111 11	-2	1111 1110
		-1	1111 1111
		0	禁止
		1	0000 0001
		2	0000 0010
		:	:
		127	0111 1111

除了intra-DC之外，其他所有系数都有一个重构非零值，取值范围为-2048~2047。重构非零值采用下面的方程恢复，其结果要进行削波处理。

若 LEVEL=0, REC=0

若 QUANT=奇数：

$$|REC|=QUANT \times (2 \times |LEVEL| + 1)$$

若 QUANT=偶数：

$$|REC|=QUANT \times (2 \times |LEVEL| + 1) - 1$$

在计算完|REC|之后，将添加其符号以获得REC值。Sign(LEVEL)将由表10-20中TCOEF编码中的s位指定。

$$REC = sign(LEVEL) \times |REC|$$

对于intra-DC块，重构非零值为：

$$REC = 8 \times LEVEL$$

## 5. PLUSPTYPE图像层选项

PLUSPTYPE由PTYPE的6~8位决定是否出现，用来确定是否采用H.263的版本2可选项。当PLUSPTYPE出现时，PLUSPTYPE及其相关的字段紧跟在PTYPE之后，并且在PQUANT之前。

如果PLUSPTYPE出现，CPM将紧跟在PLUSPTYPE之后；否则，CPM紧跟在PQUANT之后。PSBI总是紧跟在CPM之后（前提是CPM=1）。

PLUSPTYPE是一个12位或者30位的码字，由三个子字段组成：UFEP、OPPTYPE和MPPTYPE。图10-11中给出了对PLUSPTYPE及其相关字段的说明。

### □ 更新全扩展的PTYPE (UFEP)

UFEP是一个3位的码字，它由PTYPE中的“扩展PTYPE”决定是否出现。

当UFEP值为000时，图像头中仅仅包含MPPTYPE。

当UFEP值为001时，图像头中同时包含OPPTYPE和MPPTYPE。如果图像类型为intra或者EI，UFEP字段必须设置为001。

此外，如果PLUSPTYPE出现在连续序列图像的每一帧中，那么每隔5秒钟或者每隔5帧（两者中较大的优先），UFEP字段必须设置为001。在易错环境中，UFEP字段应该更频繁地设置为001。

除000和001之外的其他值，目前暂时保留。

PLUSPTYPE的可选部分 (OPPTYPE)

OPPTYPE字段包含的特征不会一帧一帧地变化，也就是说这个字段在所有帧中都是一致的。如果UFEP设置为001，下面的18位将出现在OPPTYPE中。

位1~3为源格式

000 = 保留

001 = SQCIF

010 = QCIF

011 = CIF

100 = 4CIF

101 = 16CIF

110 = 定制源格式

111 = 保留

位4为定制图像时钟频率

0 = 标准，1 = 定制

位5表示非限制运动矢量模式 (UMV)

0 = 关，1 = 开

位6表示基于语法的算术编码模式 (SAC)

0 = 关，1 = 开

位7表示高级预测模式 (AP)

0 = 关，1 = 开

位8表示高级帧内编码模式 (AIC)

0 = 关，1 = 开

位9表示去块滤波器模式 (DF)

0 = 关，1 = 开

位10表示条结构化模式 (SS)

0 = 关，1 = 开

位11表示参考图像选择模式 (RPS)

0 = 关，1 = 开

位12表示独立段解码模式 (ISD)

0 = 关，1 = 开

位13表示交替Inter-VLC模式 (AIV)

0 = 关，1 = 开

位14表示改进的量化模式 (MQ)

0 = 关，1 = 开

位15为1

位16为0

位17为0

位18为0

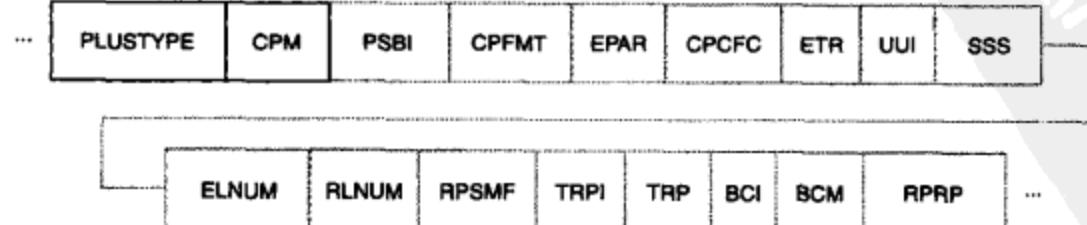


图10-11 H263的PLUSPTYPE及其相关字段

**□ PLUSPTYPE的必选部分 (MPPTYPE)**

无论UFEP的值为多少，下列9位总是出现在MPPTYPE中：

位1~3表示图像编码类型

000=I帧 (intra)

001=P帧 (inter)

010=改进的PB帧

011=B帧

100=EI帧

101=EP帧

110=保留

111=保留

位4表示参考图像重采样模式

0=关，1=开

位5表示降低分辨率更新模式 (RPR)

0=关，1=开

位6表示取整类型模式 (RTYPE)

0=关，1=开

位7为0

位8为0

位9为1

**□ 定制图像格式 (CPFMT)**

CPFMT (Custom Picture Format) 是一个23位的码字。如果UFEP设置为001，并且PLUSPTYPE的OPPTYPE中前3位指定的源格式为定制图像格式，CPFMT将出现。

位1~4 为像素宽高比 (PAR) 编码

0000=保留

0001=1 : 1

0010=12 : 11

0011=10 : 11

0100=16 : 11

0101=40 : 33

0110~1100=保留

1111=扩展的像素宽高比EPAR

位5~13为图像宽度指示 (PWI)

每行采样点数 = (PWI+1) × 4

位14为1

位15~23表示图像高度指示 (PHI)

每帧扫描线数 = (PHI+1) × 4

**□ 扩展的像素宽高比 (EPAR)**

EPAR (Extended Pixel Aspect Ratio) 是一个16位的值。如果CPFMT出现，且CPFMT的前4位指示采用扩展的像素宽高比，EPAR将出现。

位1~8表示PAR宽度

位9~16表示PAR高度

**□ 定制图像时钟频率编码 (CPCFC)**

CPCFC (Custom Picture Clock Frequency Code) 是一个8位的值。只有当PLUSPTYPE出现、

UFEP设置为001且在PLUSPTYPE的OPPTYPE中第4位指示采用定制图像时钟频率时，CPCFC才会出现。定制的图像时钟频率（单位为赫兹）为：

$1\ 800\ 000 / (\text{时钟因子} \times \text{时钟转换因子})$

位1表示时钟转换因子编码

0=1000, 1=1001

位2~8表示时钟因子

扩展时间参考 (ETR)

ETR (Extended Temporal Reference) 是一个2位的值。当采用定制图像时钟频率时，ETR将会出现。它是10位TR值中的2个最高有效位。

503

无约束非限制运动矢量指示器 (UUI)

UUI (Unlimited Unrestricted Motion Vectors Indicator) 是一个1位或者2位的可变长值，它指示运动矢量变化的有效范围。如果UFEP设置为001，且PLUSPTYPE的OPPTYPE中第5位指示采用非限制运动矢量模式，UUI将会出现。

UUI值为0时，表示运动矢量变化范围将根据表10-22和表10-23确定；而UUI值为01时，表示运动矢量在图像大小内是无约束的，即可以为图像大小内任意值。

表10-22 可选的水平运动矢量范围

图像宽度	水平运动矢量范围
4~352	-32, +31.5
356~704	-64, +63.5
708~1408	-128, +127.5
1412~2048	-256, +255.5

表10-23 可选的垂直运动矢量范围

图像高度	垂直运动矢量范围
4~288	-32, +31.5
292~576	-64, +63.5
580~1152	-128, +127.5

条带结构化子模式位 (SSS)

SSS (Slice Structured Submode Bit) 是一个2位的值。只有当UFEP设置为001，且在PLUSPTYPE的OPPTYPE中第10位指示采用条带结构化模式时，SSS才会出现。如果当前采用的是条带结构化模式，而UFEP不是001，那么最近的SSS值仍然保持有效。

位1表示矩形条带：0=不是，1=是。

位2表示任意条带顺序：0=顺序的，1=随机的。

增强层编号 (ELNUM)

ELNUM (Enhancement Layer Number) 是一个4位的值，它只有在时域、信噪比 (SNR) 和空域可分级模式中出现。ELNUM标识一个特定的增强层。基本层之上的第一个增强层指派为增强层编号为2，基本层编号为1。

参考层编号 (RLNUM)

RLNUM (Reference Layer Number) 是一个4位的值，它只有在时域、信噪比 (SNR) 和空域可分级模式下，且UFEP设置为001时出现。用作参考锚点 (reference anchor) 的帧的层编号由RLNUM确定。

参考图像选择模式标志 (RPSMF)

RPSMF (Reference Picture Selection Mode Flag) 是一个3位的码字，它只有在参考图像选择模式下且UFEP设置为001时才会出现。当RPSMF出现时，它表示解码器需要哪一个反向信道消息。如果当前采用的是参考图像选择模式，而RPSMF没有出现，那么RPSMF最近的值仍然保持有效。

000~011=保留

100=ACK和NACK都不需要

101=需要ACK

110=需要NACK

504 111=ACK和NACK都需要

预测时间参考指示 (TRPI)

TRPI (Temporal Reference for Prediction Indication) 是一个1位的值，它只有在参考图像选择模式中出现。当TRPI出现时，它表示接下来的TRP字段是否会出现。0=TRP，字段不会出现；1=TRP，字段出现。当图像头指示当前帧为I帧或者EI帧时，TRPI必须设置为0。

预测时间参考 (TRP)

TRP (Temporal Reference for Prediction) 是一个10位的值，对于除B帧之外的其他所有帧，TRP表示用于编码预测的时间参考。对于B帧，由TRP确定的时间参考将仅用于正向方向的编码预测。

如果没有采用定制时钟频率，那么参考图像的图像头中TRP的两位最高有效位将为0，而最低有效位包含8位TR值。如果采用定制时钟频率，TRP则为一个10位的值，由来自参考图像头中的ETR和TR串联组成。

如果TRP没有出现，前一个锚点图像将被用于预测，这与不是在参考图像选择模式中是一致的。TRP值将直到出现下一个PSC、GSC或者SSC时才有效。

反向信道消息指示 (BCI)

BCI (Back-Channel Message Indication) 是一个1位或者2位的可变长码字，它只有在可选的参考图像选择模式中出现。BCI值为1时表示有可选的反向信道消息 (BCM) 字段。而BCI值为01时，表示没有BCM字段或者BCM字段结束。当BCI出现时，BCM和BCI可以重复。

反向信道消息 (BCM)

如果前面的BCI字段设置为1，可变长的BCM (Back-Channel Message) 将会出现。

参考图像重采样参数 (RPRP)

RPRP (Reference Picture Resampling Parameters) 是一个可变长的字段，它只在可选的参考图像选择模式中出现。RPRP包含了参考图像重采样模式的相关参数。

### 10.2.3 可选的H.263模式

#### 1. 非限制运动矢量模式

通常运动矢量的范围被限制在参考帧内，而在非限制运动矢量模式中取消了这种限制，运动矢量可以指向图像之外。当某个运动矢量所指的参考像素位于图像之外时，可以用边缘像素代替这个“不存在的像素”。这个边界像素的确定方法是：将运动矢量限制在编码图像区域内的最后一个全像素位置。对运动矢量的限制是在像素的基础上实现的，并且是分别针对运动矢量的水平分量和垂直分量进行的。

此外，非限制运动矢量模式扩展了运动矢量范围，因此，在该模式下可以采用更大的运动矢量（表10-22和表10-23）。对于分辨率较高的图像格式（如4CIF或者16CIF），这些大范围的运动矢量将很好地改善编码效率。对于其他图像格式，如果在图像边缘处存在运动（摄像机运动或者背景运动），也将获得明显的图像质量改善。

当在H.263版本2中采用非限制运动矢量模式时，将采用一种新的可逆变长编码 (RVLC) 用于运动矢量的编码，如表10-24所示。这些编码是单值的 (single-valued)，与此相对的是基线双值VLC。由于受到其可扩展性及其高实现代价的限制，双值编码不是很普遍，而RVLC实现相对简单。

表10-24中的每一行表示一个运动矢量差分（单位为半像素）。“...X<sub>1</sub>X<sub>0</sub>”表示在运动矢量差分绝对值的二进制表示中第一位1之后的所有位。s位表示运动矢量差分的符号：0表示为正，1表示为负。运动矢量差分的二进制表示采用比特交织，并且指示编码是否继续或者结束。最后一位如果为0，表示编码结束。

RLVC也可以增强对信道传输误码的错误恢复能力。解码可以通过在前向和反向方向处理运动矢量来执行。如果在某个方向上解码时探测到一个误码，解码器可以在反方向上继续进行下去，从而改善比特流的错误恢复能力。此外，运动矢量范围被扩展到[-256, +256]之间，具体值取决于图像大小。

表10-24 H.263运动矢量的可逆变长编码

半像素精度运动矢量差分的绝对值	编 码
0	1
1	0s <sub>0</sub> 0
“x <sub>0</sub> ” + 2(2 - 3)	0x <sub>0</sub> 1s0
“x <sub>1</sub> x <sub>0</sub> ” + 4(4 - 7)	0x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 8(8 - 15)	0x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 16(16 - 31)	0x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 32(32 - 63)	0x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 64(64 - 127)	0x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>6</sub> x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 128(128 - 255)	0x <sub>6</sub> 1x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>7</sub> x <sub>6</sub> x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 256(256 - 511)	0x <sub>7</sub> 1x <sub>6</sub> 1x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>8</sub> x <sub>7</sub> x <sub>6</sub> x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 512(512 - 1023)	0x <sub>8</sub> 1x <sub>7</sub> 1x <sub>6</sub> 1x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>9</sub> x <sub>8</sub> x <sub>7</sub> x <sub>6</sub> x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 1024(1024 - 2047)	0x <sub>9</sub> 1x <sub>8</sub> 1x <sub>7</sub> 1x <sub>6</sub> 1x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0
“x <sub>10</sub> x <sub>9</sub> x <sub>8</sub> x <sub>7</sub> x <sub>6</sub> x <sub>5</sub> x <sub>4</sub> x <sub>3</sub> x <sub>2</sub> x <sub>1</sub> x <sub>0</sub> ” + 2048(2048 - 4095)	0x <sub>10</sub> 1x <sub>9</sub> 1x <sub>8</sub> 1x <sub>7</sub> 1x <sub>6</sub> 1x <sub>5</sub> 1x <sub>4</sub> 1x <sub>3</sub> 1x <sub>2</sub> 1x <sub>1</sub> 1x <sub>0</sub> 1s0

## 2. 基于语法的算术编码模式

在这个可选模式中，可变长编码被算术编码取代了。信噪比SNR和重构图像将会是一样的，但是比特率将会减少近5%，因为没有了每一个符号必须用固定长度比特数编码的需求。

在基于语法的算术编码模式中，图像层、块组层和宏块层的语法完全保持一致，块层的语法稍微改变了一点，即TCOEF实体数目可能会出现任意值。

值得注意的是，基于语法的算术编码模式并没有被普遍地采用。

## 3. 高级预测模式

通常16×16的宏块只有一个运动矢量，而在这种可选的高级预测模式下，每个宏块的4个Y块各有一个运动矢量。此外，在此模式下，对P帧中的Y块进行了交叠块运动补偿（overlapped block motion compensation, OBMC）处理。

如果对某一特定宏块只传送了一个运动矢量，那么这与传送四个值完全相同的矢量是等价的。如果四个运动矢量被用于一个宏块，那么第一个运动矢量将是MVD<sub>1</sub>码字，并且应用到图10-9中的Y<sub>1</sub>；第二个运动矢量为MVD<sub>2</sub>码字，应用到Y<sub>2</sub>；第三个运动矢量为MVD<sub>3</sub>码字，应用到Y<sub>3</sub>；第四个运动矢量为MVD<sub>4</sub>码字，应用到Y<sub>4</sub>。宏块中Cb和Cr的运动矢量将从这四个Y块运动矢量导出。

在高级预测模式下，编码器必须决定采用哪种类型的运动矢量。四个运动矢量将采用更多的位，但是它们将提供更好的预测。这种模式将改善图像间预测，从而通过减少方块效应(blocking artifacts)，在相同的比特率下将获得明显的图像质量改善。

## 4. PB帧模式

与MPEG一样，H.263选择性地支持PB帧。如图10-12所示，在H.263中一个PB帧单元由两帧

组成，其中的P帧是由前一个P帧预测所得的，而B帧则是由前一个P帧和本PB帧中的P帧通过双向预测所得的结果。

在该编码选项下，在不大幅提高比特率的情况下图像速率获得了明显的提高。然而，在Annex M<sup>①</sup>中，提供了对改进的PB帧模式的支持。因此，仅仅为了保持与在采用Annex M之前设计的系统兼容，原始PB帧模式得以保留。

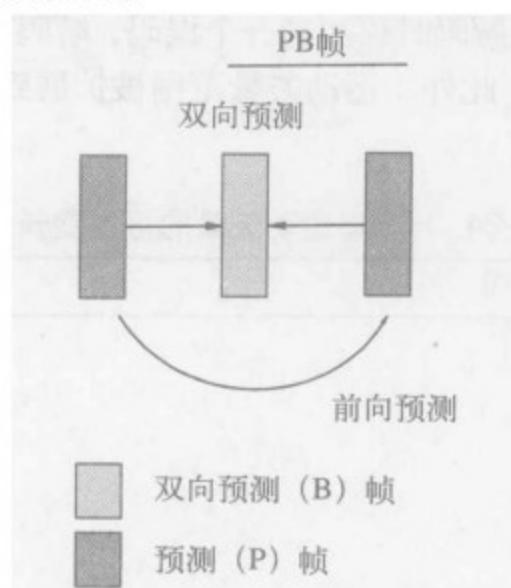


图10-12 基准H.263 PB帧

### 5. 连续出现多点和视频复用模式

在连续出现多点和视频复用模式下，一个单独的比特流最多可以复用4个独立的H.263子比特流。除非通过其他方法设置了不同的优先级约定，通常情况下，ID号最低的子比特流（通过SBI字段发送）被赋予了最高优先级。

上述特征被设计用于连续出现多点应用或分离逻辑信道无法获得的情形，这种情况下必须采用复用视频比特流。在H.324中将不采用连续出现多点和视频复用模式。

### 6. 前向纠错模式

前向纠错模式为传输H.263视频数据提供了前向纠错功能（编码和成帧）。在H.324中不采用该模式。

在前向纠错模式下，H.263的成帧和前向纠错编码与H.261中是完全一致的。

### 7. 高级帧内编码模式

高级帧内编码模式改善了宏块内的压缩效率。它采用来自相邻块内的块内预测、修改的intra-DCT系数反量化和单独的intra系数的VLC表。这种模式明显地改善了基准H.263帧内编码的压缩性能。

在高级帧内编码模式下，在宏块层中的MCBPC字段之后增加了一个附加的1位或者2位的变长码字（INTRA\_MODE），用来指示预测模式：

- 0=仅DC
- 10=垂直DC和AC
- 11=水平DC和AC

对于帧内编码块，如果预测模式仅为DC，采用图7-57中的锯齿形扫描顺序。如果预测模式为垂直DC和AC，采用图7-58中交替垂直扫描顺序。如果预测模式为水平DC和AC，采用图7-59中的交替水平扫描顺序。

<sup>①</sup>一种ADSL技术。——译者注

对于非帧内编码块，将采用图7-57中的锯齿形扫描顺序。

### 8. 去块滤波器模式

在该可选模式中，编码环路中引入了一个去块滤波器。这个去块滤波器被应用到 $8 \times 8$ 块的边界处，用来减少方块效应。

去块滤波器的系数取决于宏块量化器的步长，大的系数被用于粗粒度的量化器。与高级预测模式一样，去块滤波器模式也允许每个宏块采用四个运动矢量，并且与非限制运动矢量模式一样，也允许运动矢量指向图像区域之外。由于没有采用高级预测模式中高计算代价的交叠运动补偿运算，因此保持了该模式最小的复杂性。

采用去块滤波器模式获得的结果是更好的预测和方块效应减少。

### 9. 条带结构化模式

在该可选模式中，条带层取代了GOB层。这种模式提供了错误恢复能力，使得比特流容易被分组传输传递机制所使用，并且最小化视频时延。

条带层由一个条带头和连续的完整宏块组成。在条带结构化模式中，有两个附加的模式可以用来反映传输的顺序（顺序的或者随机的）和条带的形状（矩形的或者非矩形的）。这些附加的模式使得条带结构更加灵活，从而能够被用于不同的应用。

### 10. 补充的增强信息

在该可选模式下，比特流中可能会包含额外的补充信息，以获得增强的显示能力。

典型的增强信息可以为视频复合提供全部或者部分图像冻结<sup>①</sup>、图像冻结释放或者色度抠像的能力。508

在某些情况下，即便解码器无法使用这些补充的增强信息，它们仍然有可能会出现在比特流中。对此，解码器只是简单地丢弃这些补充信息，除非编解码器已经通过另外的方法协商以支持这种能力的需求。

### 11. 改进的PB帧模式

改进的PB帧模式是相对基准H.263的PB帧可选项的一种改进。这种模式允许在PB帧中对B帧进行前向、后向和双向预测。表10-25中给出了MODB字段变化的操作。

表10-25 改进的PB帧模式中MODB的H.263可变长编码表

CBPB	MVDB	编 码	编码模式
x		0	双向预测
	x	10	双向预测
x	x	110	前向预测
x	x	1110	前向预测
x	x	11110	后向预测
x	x	11111	后向预测

改进的PB帧模式中的双向预测方法与PB帧模式的双向预测方法是一样的，唯一不同之处就是在改进的PB帧模式中不发送 $\delta$ 矢量。

在前向预测中，B宏块由前一个的P宏块预测得到，并且发送一个单独的运动矢量。

在后向预测中，预测的宏块等同于后一个P宏块，因此没有运动矢量被发送。

改进的PB帧不易受到可能发生在帧之间的变化的影响，如前一个P帧和当前PB帧之间的场景切变。

<sup>①</sup> 图像冻结为在视频播放过程中整帧或者部分图像冻结，其他部分继续播放的效果。——译者注

## 12. 参考图像选择模式

在基准H.263中，某一帧可能由前一帧预测得到。如果参考帧的某一部分由于信道误码或者分组丢失而被丢失了，由这些参考帧推导出来的帧的质量将会下降。采用参考图像选择模式，编码器将有可能选择采用哪一个参考帧来进行预测，从而最小化误码传播。

编码器和解码器利用四种反向信道消息信号（NEITHER、ACK、NACK和ACK+NACK）来确定哪一个图像段将被用于预测。例如，解码器发送一个NACK消息给编码器，表示给定帧质量由于信道误码而下降了。因此，编码器可以选择不使用这一帧用于接下来的预测，转而采用另外的、质量没受影响的参考帧。这可以降低误码传播，从而在易错环境中保持较好的图像质量。

## 13. 时域、SNR和空域可分级模式

在这个可选模式中，分别支持三种不同的可分级模式：时域、SNR和空域。可分级使得解码器可以在多个质量层次上解码视频序列。可分级是通过采用将图像层次化且将增强图像分成一个或多个层而实现的。最低层被称为基本层。

基本层是一个可单独解码的比特流。增强层可以与基本层一起解码，从而增加图像速率，提高图像质量或者增大图像尺寸。

时域可分级是通过双向预测图像（或B帧）来达到的。在这种模式下，预测可以是基于基本层中的前一个图像或者后一个图像，或者两者都采用。其结果是获得比P帧更好的压缩比。这些B帧不同于PB帧或者改进的PB帧中的B图像部分，主要在于它们在比特流中是独立的实体。

**509** SNR可分级依靠增强信息来提高图像质量而不增加分辨率。由于压缩引入了块效应，因此，解码后图像与原始图像之间不可避免地存在差异，这种差异称为编码误差。正常情况下，编码误差在编码器中被丢弃，并且永远不会恢复。而在SNR可分级模式下，编码误差被发送到解码器，使得解码图像获得增强。这些额外的数据主要用于提高图像的SNR，这也就是术语“SNR可分级”的由来。

空域可分级与SNR可分级非常相关。唯一的不同就是在参考层中图像被用于预测空域增强层中图像之前，它被插入了1/2的垂直或者水平运动矢量（一维空域可分级），或者两者都有（二维空域可分级）。除了上采样过程之外，空域可分级图像的处理和语法与SNR可分级图像是一样的。

由于在采用SNR可分级的帧与采用空域可分级的帧之间有非常小的句法差异，因此，采用这两种可分级模式的帧分别被称为EI帧和EP帧。

在基本层中，用于在EI帧或者EP帧中向上预测的帧可能是I帧、P帧、PB帧中的P部分或者改进的PB帧的中P部分（但是不可能是B帧，PB帧中的B部分或者改进的PB帧中的B部分）。

该模式对于带宽容量变化的网络非常有用。

## 14. 参考图像重采样模式

在该可选模式中，在将参考帧用于预测之前，需要对参考帧进行重采样，以获得不同的图像大小。

参考图像重采样模式使得编码器有一个与被预测的帧不同的源参考格式。它也可以通过改变(warping)参考帧的形状、大小和位置来进行全局运动估计，或者旋转矢量估计。

## 15. 降低分辨率更新模式

该可选模式允许编码器发送某一帧以较低分辨率编码的更新信息，而对于参考帧，则仍然保持较高的分辨率，从而获得一个最终分辨率较高的帧。

**510** 降低分辨率更新模式最适合用于编码高动态场景，对于场景中运动部分，允许编码器增加其帧率，而对于场景中的静态区域则维持较高的分辨率。

降低分辨率更新模式的语法与基准H.263是一致的，但是语义的解释有所差别。在该模式下，

宏块的维度翻倍了，所以宏块数据大小将是不采用这种模式时的四分之一。因此，在所有维度上的运动矢量也必须翻倍。为了产生最终的图像，宏块将上采样到扩展的分辨率。在向上采样以后，将全分辨率帧与运动补偿帧相加，以生成全分辨率帧作为后续帧的参考。

#### 16. 独立段解码模式

在该可选模式中，将图像的段边界作为图像的边界进行处理，即允许在段边界处无数据依赖。

采用这种模式可以防止差错的传播，提供错误恢复和恢复的能力。这种模式最适合用于条带层，例如，可以调整条带大小以匹配特定的分组大小。

#### 17. 交替Inter-VLC模式

当采用交替Inter-VLC模式时，用于高级帧内编码模式中的intra-VLC表同样可以用于块间编码。

当采用小的量化器步长时，通常出现在块内部的大量化系数和小游戏程越来越频繁地出现在块之间。当可以比特节省，并且解码器可以探测到采用内部量化DCT系数表时，解码器将采用intra表。解码器将首先试图采用inter表解码量化的系数。如果导致寻址系数在 $8 \times 8$ 块的64个系数之外，解码器将采用intra表。

#### 18. 修改的量化模式

该可选模式将改善对编码的比特率控制，减少CbCr量化误差，扩展DCT系数的范围以及对系数值给予一定的限制。

在基准H.263中，可以在宏块层修改量化器值。然而，仅仅允许对最近的量化器的值进行小的调整（±1或±2）。修改的量化模式允许将量化器的值修改为任意值。

在基准H.263中，Y块和CbCr块量化器是一致的。在修改的量化模式中，通过对相对于Y块的Cb和Cr块采用一个较小的量化器步长，同样提高了Cb和Cr块的图像质量。

在基准H.263中，当采用一个小于8位的量化器时，超出[−127, +127]范围的量化系数将被削波。而在修改的量化模式下，量化系数被允许超出[−127, +127]范围。因此，当选择一个非常精细的量化步长时，Y块的质量将获得较大的提升。

#### 19. 增强的参考图像选择模式

增强的参考图像选择（ERPS）模式提供了增强的编码效率和错误恢复能力。它管理着一个存储图像的多图像缓存器。

#### 20. 数据分区的条带模式

数据分区的条带（DPS）模式提供了增强的错误恢复能力。它将头数据和运动矢量数据从DCT系数数据中分离出来，并且通过采用可逆的表示方法保护运动矢量数据。

511

#### 21. 额外的补充增强信息说明

额外的补充增强信息说明提供了向后兼容的增强，如下所示。

- (1) 采用一个特定固定点IDCT的指示。
- (2) 图像消息，包括的消息类型有：任意二进制数据，文本（任意、版权、标题、视频描述或URL），图像头副本（有可靠时间参考的当前、前一个、下一个图像或者无可靠时间参考的下一个图像），隔行扫描字段（顶部或者底部），备用参考图像标识。

### 10.2.4 档次

档次定义了与特定应用相关的常用配置语法（如算法）。

#### 1. 档次0

基本档次或者档次0，不采用任何可选模式。

## 2. 档次1

档次1（H.320编码效率版本2向后兼容档次）提供与H.242和H.320的兼容性。它由档次0与下面的模式共同组成：高级帧内编码，去块滤波器，补充的增强信息——全图像冻结，修改的量化。

## 3. 档次2

档次2（版本1向后兼容档次）为H.263的第一个版本提供增强的编码效率。它由档次0加下面的模式组成：

512

高级预测

## 4. 档次3

档次3（版本2交互式的和流化无线档次）为无线设备提供增强的编码效率性能和增强的错误恢复能力。它由档次0加上下列模式组成：高级帧内编码，去块滤波器，条带结构化，修改的量化。

## 5. 档次4

档次4（版本3交互式的和流化无线档次）为无线设备提供可增强的编码效率性能和增强的错误恢复能力。它由档次0和档次3加上下列模式组成：数据分区条带，补充的增强信息——前一个图像头副本。

## 6. 档次5

档次5（会话式高压缩档次）在不增加与采用B图像相关的时延和增加错误恢复特性的情况下，提供增强的编码效率。它由档次0、档次1和档次2加上下列模式组成：非限制运动矢量——UUI=1，增强的参考图像选择。

## 7. 档次6

档次6（会话式Internet档次）在不增加与采用B图像相关的时延，增加某些适合于在IP网络上使用的错误恢复能力的情况下，提供增强的编码效率性能。它由档次0和档次5加上下面的模式组成：随机条带顺序的条带结构化，补充的增强信息——前一个图像头副本。

## 8. 档次7

档次7（会话式隔行扫描档次）为低时延应用提供增强的编码效率性能，并且支持隔行扫描视频源。它由档次0和档次5加上下面的模式组成：补充的增强信息——隔行扫描字段指示（对于 $240 \times 288$ 图像）。

## 9. 档次8

档次8（高延迟档次）为没有紧要时延约束的应用提供增强的编码效率性能。它由档次0和档次6加上下列模式组成：参考图像重采样，时域可分级——B图像。

### 10.2.5 级别

级别定义了一个档次中的各种参数（分辨率、帧率和比特率等）。

#### 1. 级别10

最大支持 $176 \times 144$ 分辨率，最大比特率为 $64\text{kbit/s}$ 。

#### 2. 级别20

最大支持 $352 \times 288$ 分辨率，最大比特率为 $128\text{kbit/s}$ 。

#### 3. 级别30

最大支持 $352 \times 288$ 分辨率，最大比特率为 $384\text{kbit/s}$ 。

513

**4. 级别40**

最大支持 $352 \times 288$ 分辨率，最大比特率为2Mbit/s。

**5. 级别45**

最大支持 $176 \times 144$ 分辨率，最大比特率为128kbit/s。

**6. 级别50**

最大支持 $352 \times 288$ 分辨率，最大比特率为4Mbit/s。

**7. 级别60**

最大支持 $720 \times 288$ 分辨率，最大比特率为8Mbit/s。

**8. 级别70**

最大支持 $720 \times 576$ 分辨率，最大比特率为16Mbit/s。

## 参考文献

1. *Efficient Motion Vector Estimation and Coding for H.263-Based Very Low Bit-Rate Video Compression*, by Guy Cote, Michael Gallant, and Faouzi Kossentini, Department of Electrical and Computer Engineering, University of British Columbia.
2. *H.263 +: Video Coding at Low Bit-Rates*, by Guy Cote, Berna Erol, Michael Gallants, and Faouzi Kossentini, Department of Electrical and Computer Engineering, University of British Columbia.
3. ITU-T H.261, *Video Codec for Audiovisual Services at  $p \times 64$  kbits*, 3/93.
4. ITU-T H.263, *Video Coding for Low Bit-Rate Communication*, 01/2005

# 第11章 消费DV

DV (Digital Video, 数字视频) 格式由基于磁带的数字便携式摄像机使用，它基于IEC 61834 (25Mbit/s比特率) 和新近的SMPTE 314M和370M规范 (25、50或者100Mbit/s比特率)。尽管消费DV与MPEG的I帧有许多地方相似，但是它所使用的压缩算法既不是M-JPEG也不是MPEG。消费DV使用一个可以编辑的专有压缩算法，因为这种算法是一种帧内技术。

数字化的视频在压缩之前存储在内存中。然后测量缓冲器中存储的两个场之间的相关性。如果相关性低，表示场间存在运动，两个场单独进行压缩。通常情况下，压缩的是整个帧。不管是哪一种情况，都采用基于DCT的压缩。

为了实现恒定的25、50或者100Mbit/s的比特率，DV使用自适应量化，即对每个帧使用合适的DCT量化表。

图11-1描述了磁带上所写的一个磁道的内容。ITI区 (Insert and Track Information sector, 插入和磁道信息区) 包含磁道状态的信息，作为视频编辑期间的约定控制磁道。



图11-1 480i系统磁道分区的安排。每个磁道的总位数为134 975 (134 850) bit，包括重写裕度。每个视频帧有10 (12) 个磁道。576i系统参数 (如果不同) 示范在括号中

音频区 (audio sector) 包含音频数据和辅助音频数据 (auxiliary audio data, AAUX)，如图11-2所示。

视频区 (video sector) 包含视频数据和辅助视频数据 (auxiliary video data, VAUX)，如图11-3所示。VAUX数据包括录制日期和时间、透镜孔径、快门速度、色平衡和其他摄像机设置。

子码区 (subcode sector) 存储一系列的信息，包括时间码、图文电视、多语言中的隐藏字幕、多语言中的对白字幕和卡拉OK歌词、标题、目录、章节等。子码区、AAUX数据和VAUX数据使用5字节的数据块，这样的数据块称为包组。

## 11.1 音频

音频帧以第1行开头 (480i系统) 或者第623行中间 (576i系统) 前面50个样本中的一个音频样本为起始点。

同步块号	同步 2B	ID 3B	AAUX数 据5B	音频数据 72B	内部奇 偶校验 8B
0	同步 2B	ID 3B	AAUX数 据5B	音频数据 72B	内部奇 偶校验 8B
1	同步 2B	ID 3B	AAUX数 据5B	音频数据 72B	内部奇 偶校验 8B
8	同步 2B	ID 3B	AAUX数 据5B	音频数据 72B	内部奇 偶校验 8B
9	同步 2B	ID 3B		外部奇偶校验	内部奇 偶校验 8B
13	同步 2B	ID 3B		外部奇偶校验	内部奇 偶校验 8B

图11-2 音频区同步块的结构

同步块号	同步 2B	ID 3B	VAUX数据 77B	内部奇 偶校验 8B
0	同步 2B	ID 3B	VAUX数据 77B	内部奇 偶校验 8B
1	同步 2B	ID 3B	VAUX数据 77B	内部奇 偶校验 8B
2	同步 2B	ID 3B	视频数据 77B	内部奇 偶校验 8B
136	同步 2B	ID 3B	视频数据 77B	内部奇 偶校验 8B
137	同步 2B	ID 3B	VAUX数据 77B	内部奇 偶校验 8B
138	同步 2B	ID 3B	外部奇偶校验	内部奇 偶校验 8B
148	同步 2B	ID 3B	外部奇偶校验	内部奇 偶校验 8B

图11-3 视频区同步块的结构

每个磁道包含9个音频同步块，每个音频同步块包含5B的音频辅助数据（AAUX）和72B的音频数据，如图11-2所示。音频样本散乱地位于帧的各个磁道和数据同步块中。剩余的5个音频同步块用于错误校正。

2个44.1kHz、16位的通道要求大约1.64Mbit/s的数据传输率。4个32kHz、12位的通道要求大约1.536Mbit/s的数据传输率。2个48kHz、16位通道要求大约1.536Mbit/s的数据传输率。

### 11.1.1 IEC 61834

IEC 61834支持多种音频采样率：48kHz（16位，2通道）、44.1kHz（16位，2通道）、32kHz（16位，2通道）、32kHz（12位，4通道）。音频采样可以锁定到也可以不锁定到视频帧频上。

音频数据按帧处理。对于一个锁定的48kHz采样率，每个帧或者包含1600或1602个音频样本（480i系统），或者包含1920个音频样本（576i系统）。对于480i系统，每帧音频样本的数量遵循一个5帧序列：

1600, 1602, 1602, 1602, 1600, ……

对于一个锁定的32kHz采样率，每个帧或者包含1066或1068个音频样本（480i系统），或者包含1280个音频样本（576i系统）。对于480i系统，每帧音频样本的数量遵循一个15帧序列：

1066, 1068, 1068, 1068, 1068, 1068, 1068, 1066,

1068, 1068, 1068, 1068, 1068, 1068, 1068, ……

对于未锁定的音频采样，尽管可以指定每帧音频样本的最大值和最小值，但是每帧音频样本的数量不固定。

### 11.1.2 SMPTE 314M/370M

SMPTE 314M和370M支持更加有限的选项，音频采样要锁定到视频帧频上：

- 25Mbit/s, 48kHz (16位, 2通道)
- 50Mbit/s, 48kHz (16位, 4通道)
- 100Mbit/s, 48kHz (16位, 8通道)

音频数据按帧处理。对于一个锁定的48kHz采样率，每个帧或者包含1600或1602个音频样本（60场/帧制），或者包含1920个音频样本（50场/帧制）。对于60场/帧制，每帧音频样本的数量遵循一个5帧序列：

1600, 1602, 1602, 1602, 1600, ……

音频容量对于60场/帧制能够满足每帧1620个样本，或者对于50场/帧制能够满足每帧1944个样本。每帧末尾未使用的空间填充的是任意数据。

### 11.1.3 音频辅助数据 (AAUX)

AAUX信息添加到图11-2中所示的散乱的音频数据上。AAUX包组包含一个1字节的包组头和4字节的数据（载荷），结果生成一个5字节的AAUX包组。因为每个视频帧中包含9个AAUX包组，所以AAUX包组的编号为0~8。压缩流中必须包含一个AAUX源（AAUX source, AS）包组和一个AAUX源控制（AAUX source control, ASC）包组。尽管IEC 61834支持许多其他包组格式，但是当前，SMPTE只支持AS和ASC包组。

#### 1. AAUX源 (AS) 包组

该包组的格式如表11-1所示。

表11-1 AAUX源 (AS) 包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	0
PC1	LF	1				AF		
PC2	SM		CHN	PA			AM	
PC3	1	ML	50/60				STYPE	
PC4	EF	TC		SMP			QU	
SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	0
PC1	LF	1			AF			

(续)

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC2	0	CHN		1	AM			
PC3	1	1	50/60		STYPE			
PC4	1	1		SMP		QU		

LF为锁定的音频采样率：0=锁定到视频，1=不锁定到视频。

AF为音频帧大小。指定每帧音频样本的数量。

SM为立体声模式：0=多立体声音频，1=集总音频 (lumped audio)。518

PA指定记录在CH1 (CH3) 中的音频信号是否与记录在CH2 (CH4) 中的音频信号相关：0=成对通道中的一对，1=独立通道。

CHN为一个音频块中音频通道的数量：00=每块1个通道，01=每块2个通道，10=保留，11=保留。

AM指定每个通道上音频信号的内容。

ML为多语言标志：0=以多语言录制，1=没有以多语言录制。

50/60表示50场或60场制式：0=60场制式，1=50场制式。

STYPE表示对于SMPTE 314M/370M，指定每帧音频块的数量：00000=2个音频块，00001=保留，00010=4个音频块，00011=8个音频块，00100~11111=保留。对于IEC 61834，指定视频系统：00000=标准清晰度，00001=保留，00010=高清晰度，00011~11111=保留。

EF为音频加重标志：0=打开，1=关闭。

TC表示加重时间常量：1=50/15μs，0=保留。519

SMP表示音频采样频率：000=48kHz，001=44.1kHz，010=32kHz，011~111=保留。

QU表示音频量化：000=16位线性，001=12位非线性，010=20位线性，011~111=保留。

## 2. AAUX源控制 (ASC) 包组

该包组的格式如表11-2所示。

表11-2 AAUX源控制 (ASC) 包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	1
PC1	CGMS		ISR		CMP		SS	
PC2	REC S	REC E		REC M			ICH	
PC3	DRF				SPD			
PC4	1				GEN			

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	0	1	0	0	0	1
PC1	CGMS		1	1	1	1	EFC	
PC2	REC S	REC E	FADE S	FADE E	1	1	1	1
PC3	DRF				SPD			
PC4	1	1	1	1	1	1	1	1

CGMS为复制代次管理系统：00=允许无限制复制，01=保留，10=允许复制一次，11=不允许复制。

ISR为前一输入源：00=模拟输入，01=数字输入，10=保留，11=没有信息。

CMP为压缩倍数：00=1倍，01=2倍，10=3倍或更大，11=没有信息。

SS表示源和录制情况：00=观众受限的加密源并且没有解密录制，01=观众不受限的加密源并且没有解密录制，10=观众受限的源或者观众受限的解密源，11=没有信息。

EFC为音频加重标志：00=加重关闭，01=加重打开，10=保留，11=保留。

REC S为录制起始点：0=位于录制起始点，1=没有位于录制起始点。

REC E为录制结束点：0=位于录制结束点，1=没有位于录制结束点。

REC M为录制模式：001=原始，011=插入1个CH，100=插入4个CH，101=插入2个CH，111=无效录制。

**520** FADE S表示录制起始点的淡入：0=淡入关闭，1=淡入打开。

FADE E表示录制结束点的淡出：0=淡出关闭，1=淡出打开。

ICH为插入音频通道：001=CH1，001=CH2，010=CH3，011=CH4，100=CH1、CH2，101=CH3、CH4，110=CH1、CH2、CH3、CH4，111=没有信息。

DRF为方向标志：0=反向，1=正向。

SPD为回放速度。

GEN指示音频源的种类。

## 11.2 视频

如表11-3所示，IEC 61834对 $720 \times 480i$ 视频（图3-5）使用4:1:1 YCbCr，对 $720 \times 576i$ 视频

**521** （图3-11）使用4:2:0 YCbCr。

表11-3 IEC 61834、SMPTE 314M和SMPTE 370M YCbCr参数

参 数	480i系统	576i系统	720p系统	1080i系统
有效分辨率 (Y)	$720 \times 480i$	$720 \times 576i$	$1280 \times 720p$	$1920 \times 1080i$
帧频	29.97Hz	25Hz	50Hz 59.94Hz	25Hz 29.97Hz
YCbCr采样结构				
IEC 61834	4:1:1	4:2:0		
SMPTE 314M	4:1:1、4:2:2	4:1:1、4:2:2		
SMPTE 370M			4:2:2	4:2:2
YCbCr编码格式	均匀量化PCM，每样本8 bit		均匀量化PCM，每样本10 bit	
有效行号	23~262、285~524	23~310、335~622	26~745	21~560、584~1123

SMPTE 314M对25Mbit/s实现的两类视频标准都使用4:1:1 YCbCr（图3-5）。4:2:2 YCbCr（图3-3）适用于50Mbit/s和100Mbit/s的实现。

### 11.2.1 DCT块

每个帧的Y、Cb和Cr样本分成 $8 \times 8$ 个块，称为DCT块。除了4:1:1模式期间Cb和Cr最右边的DCT块，每个DCT块传送视频数据的8样本 $\times$ 8行。DCT块的第1、3、5和7排处理第1场，而DCT块的第0、2、4和6排处理第2场。

对于480i系统，每个视频帧有10 800 (4:2:2) 或者8100 (4:1:1) 个DCT块。

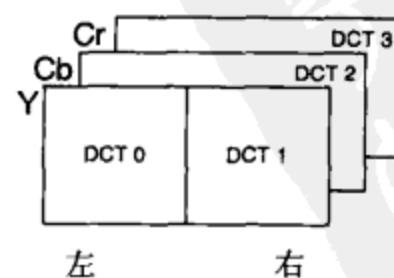


图11-4 4:2:2宏块排列

对于576i系统，每个视频帧有12 960 (4 : 2 : 2) 或者9720 (4 : 1 : 1、4 : 2 : 0) 个DCT块。

### 11.2.2 宏块

如图11-4所示，4 : 2 : 2模式中的每个宏块（macroblock）包含4个DCT块。如图11-5和图11-6所示，4 : 1 : 1模式和4 : 2 : 0模式中的每个宏块包含6个DCT块。

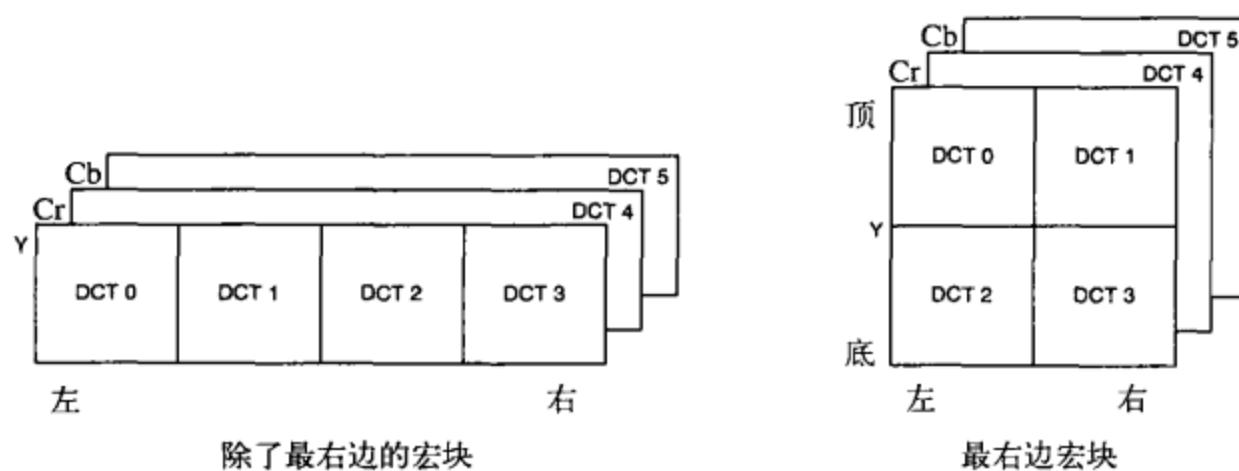


图11-5 4 : 1 : 1宏块排列

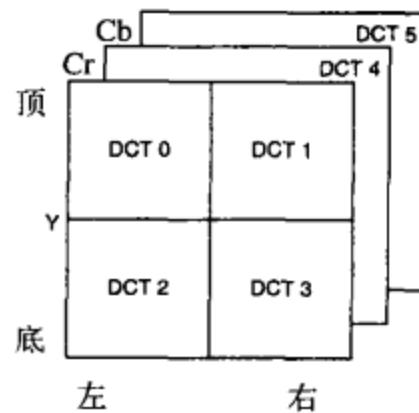


图11-6 4 : 2 : 0宏块排列

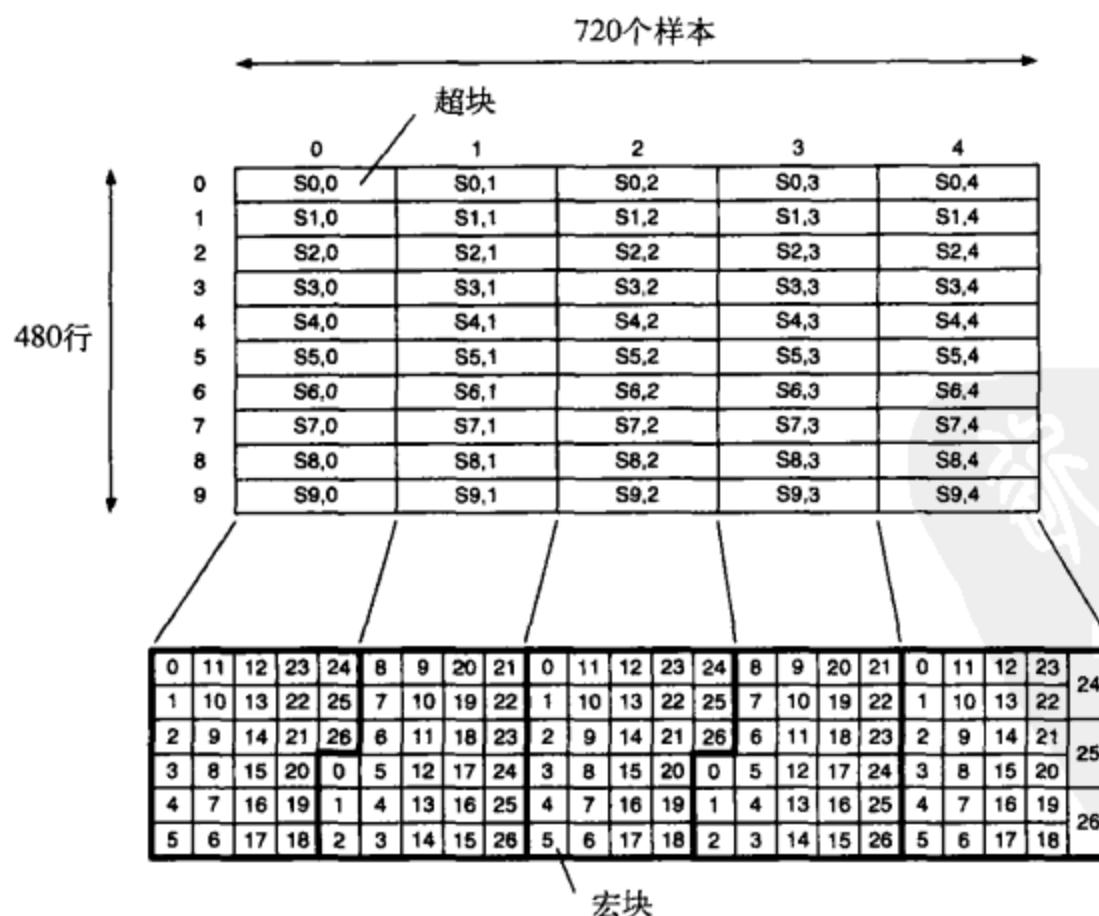


图11-7 超块和宏块之间的关系 (4 : 1 : 1 YCbCr, 720 × 480i)

对于480i系统，4：1：1和4：2：2 YCbCr数据的帧的宏块排列分别如图11-7和图11-8所示。每个视频帧有2700（4：2：2）或者1350（4：1：1）个宏块。

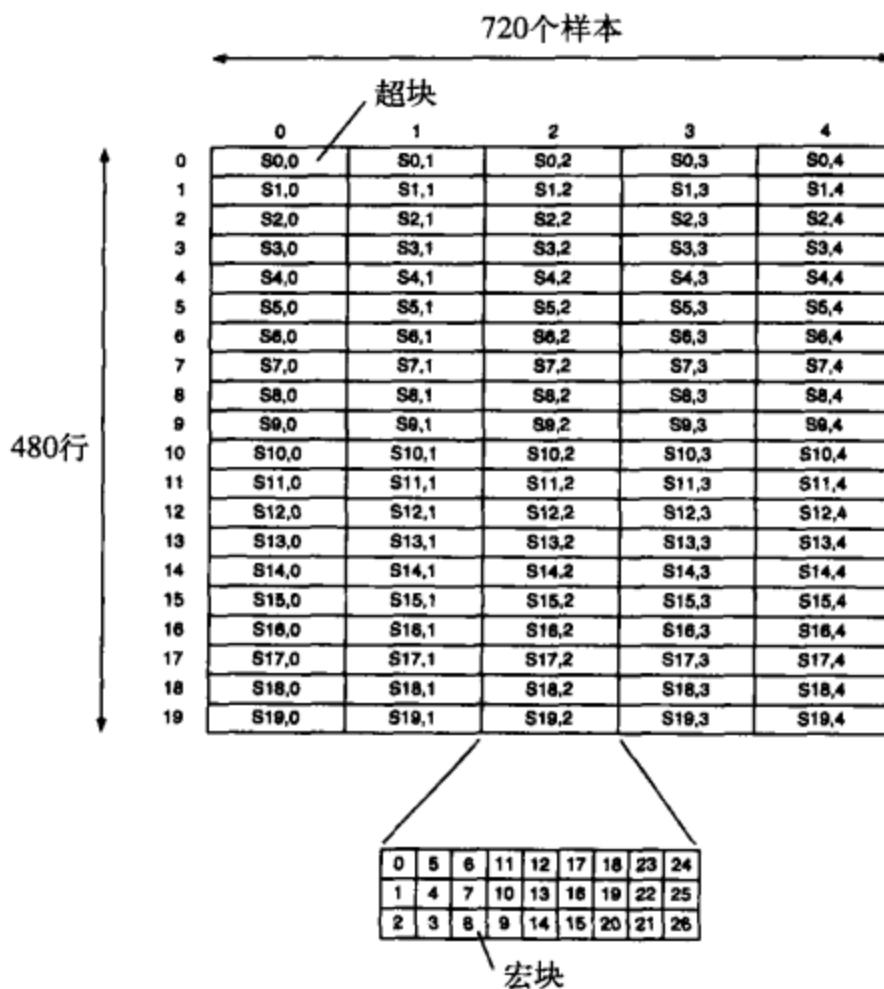


图11-8 超块和宏块之间的关系（4：2：2 YCbCr, 720×480i）

对于576i系统，4：2：0、4：1：1和4：2：2 YCbCr数据的帧的宏块排列分别如图11-9、图11-10和图11-11所示。每个视频帧有3240（4：2：2）或者1620（4：1：1、4：2：0）个宏块。

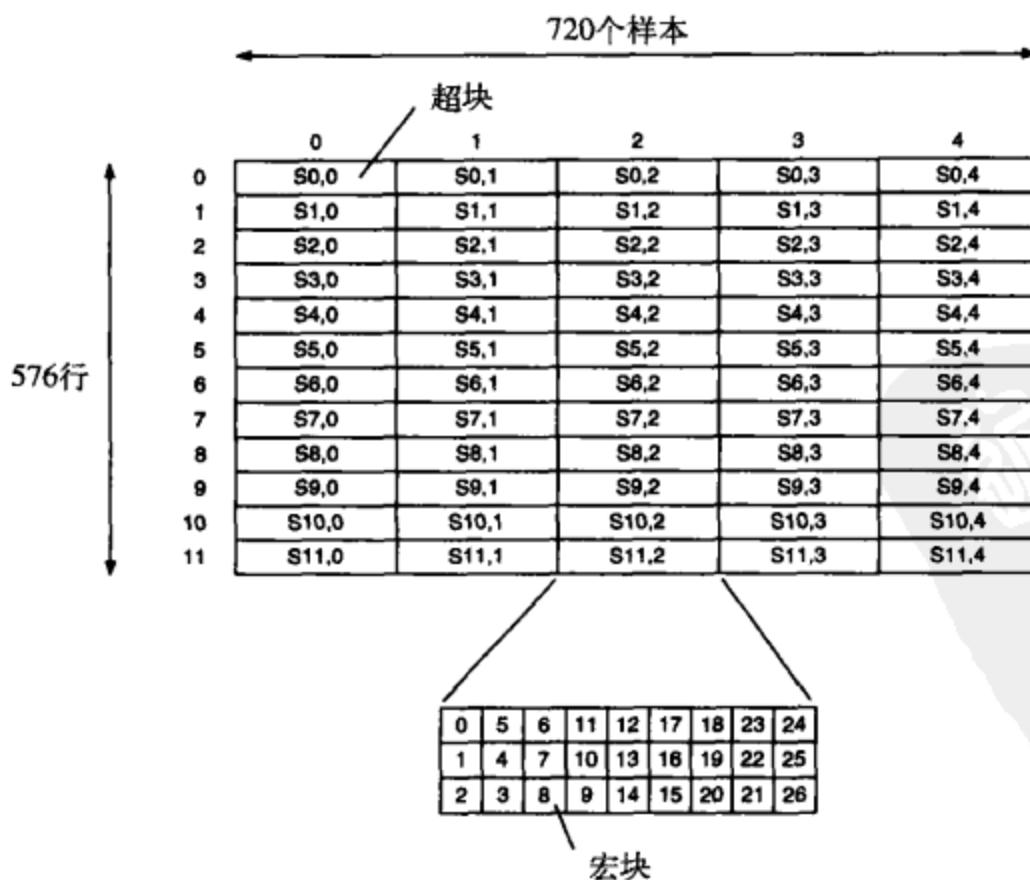


图11-9 超块和宏块之间的关系（4：2：0 YCbCr, 720×576i）

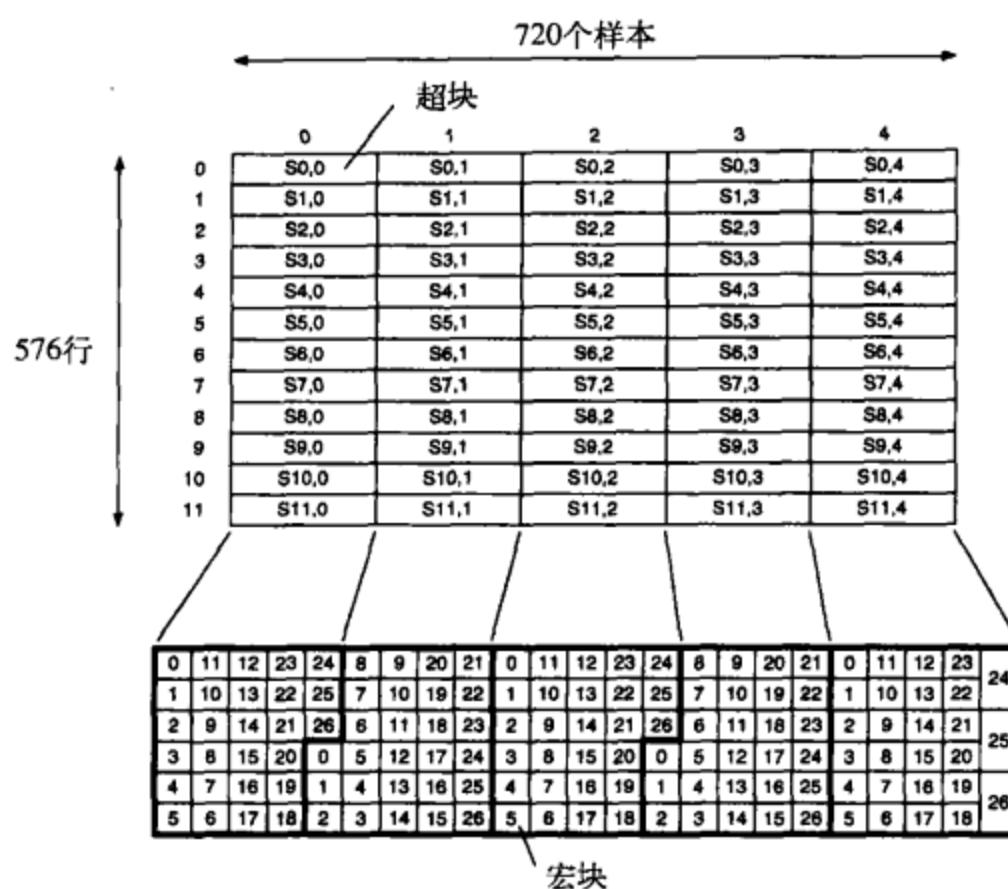


图11-10 超块和宏块之间的关系 (4 : 1 : 1 YCbCr, 720 × 576i)

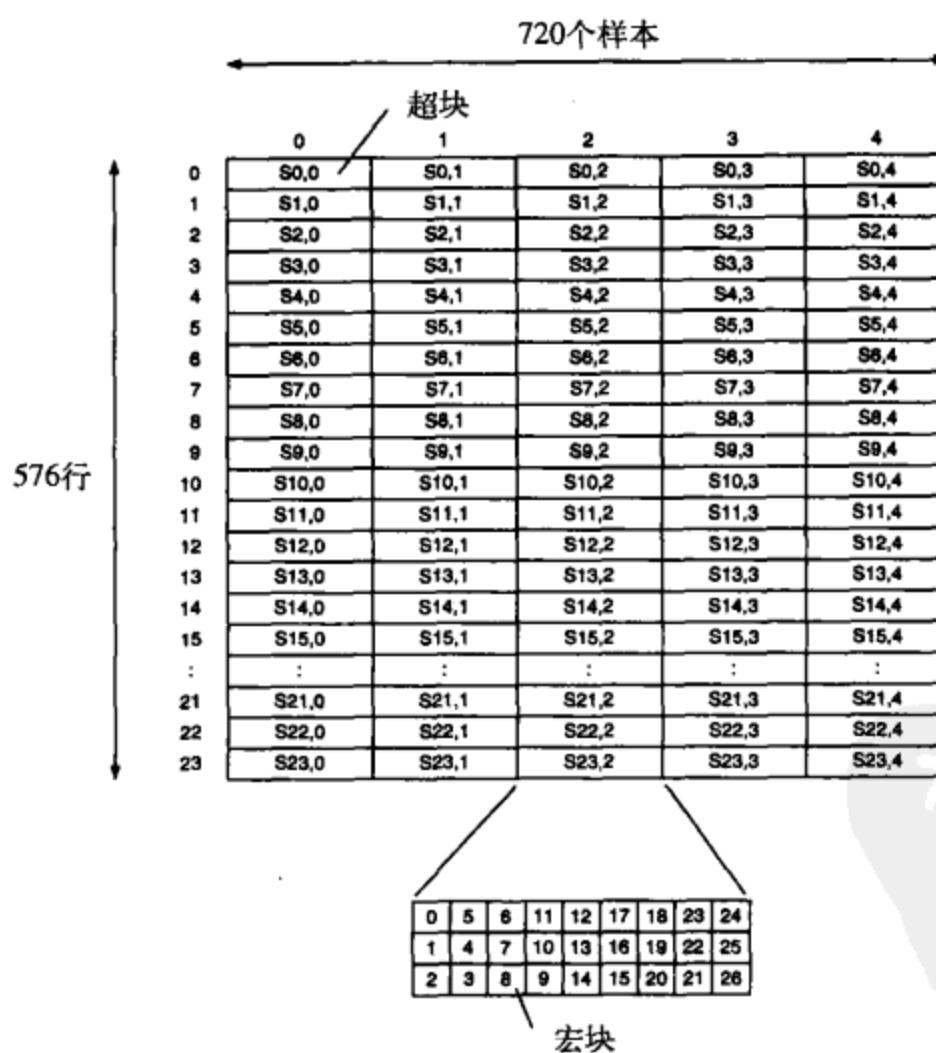


图11-11 超块和宏块之间的关系 (4 : 2 : 2 YCbCr, 720 × 576i)

### 11.2.3 超块

每个超块 (superblock) 包含27个宏块。

对于480i系统，4：1：1和4：2：2 YCbCr数据的帧的超块排列分别如图11-7和图11-8所示。每个视频帧有100（4：2：2）或者50（4：1：1）个超块。

对于576i系统，4：2：0、4：1：1和4：2：2 YCbCr数据的帧的超块排列分别如图11-9、图11-10和图11-11所示。每个视频帧有120（4：2：2）或者60（4：1：1、4：2：0）个宏块。

#### 11.2.4 压缩

像MPEG和H.263一样，DV使用基于DCT的视频压缩。但是，在这种情况下，DCT块由2个场构成，每个场提供4个扫描行中的样本以及8个水平样本。

DCT模式有2种，分别称为8-8-DCT和2-4-8-DCT，根据视频帧两场之间内容的变化程度，压缩转换过程可以选择使用这两种模式。8-8-DCT是普通的 $8 \times 8$  DCT，当两个场之间的相关程度较高（运动很少）时使用。2-4-8-DCT使用2个 $4 \times 8$  DCT（每个场一个），当两个场之间的相关程度较低（运动较多）时使用。关于使用哪个DCT的信息使用一个单独位存储在DC系数中。

**522** DCT系数量化为9位，然后除以一个量化数，从而将一个视频段中的数据量限制到5个压缩宏观中。

每个DCT块基于量化噪声（quantization noise）和AC系数的最大绝对值分成4类。2bit的分类号存储在DC系数区域中。

区域号用于量化阶的选择。区域号有4个，它们的选择基于行频和场频。

量化阶由分类号、区域号和量化号（QNO）决定。量化信息传递到视频块的DIF头中。

可变长度的编码将量化的AC系数转换为可变长度的代码。

图11-12和图11-13描述了压缩宏块的排列。

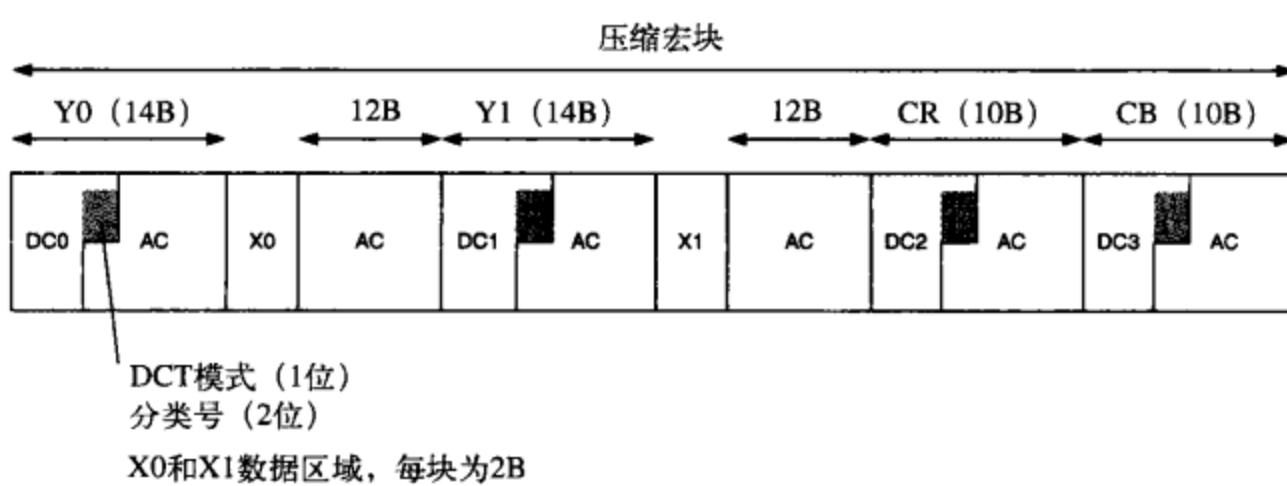


图11-12 4：2：2压缩宏块排列

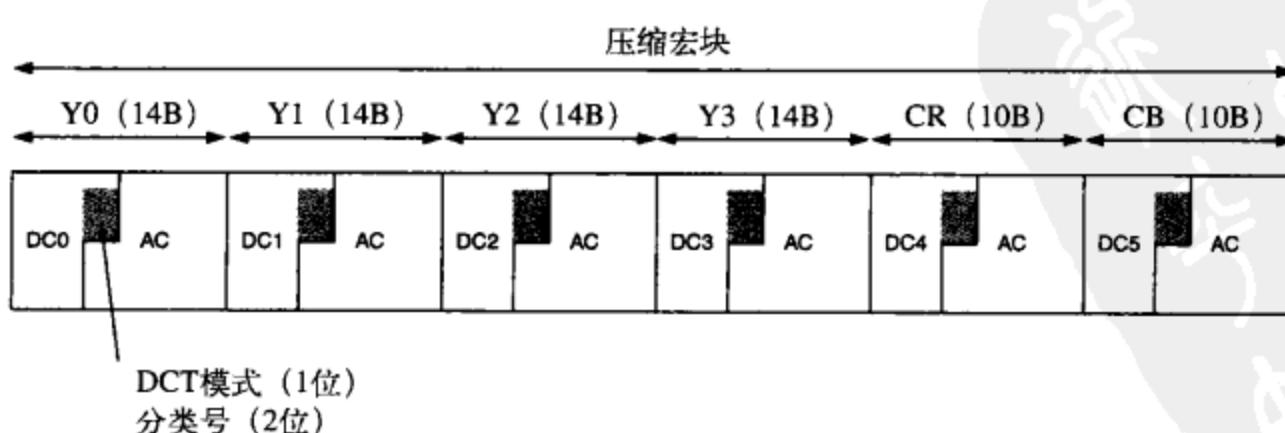


图11-13 4：2：0和4：1：1压缩宏块排列

### 11.2.5 视频辅助数据 (VAUX)

VAUX信息添加到图11-3中所示的散乱的视频数据上。VAUX包组包含一个1字节的包组头和4字节的数据（载荷），结果生成一个5字节的VAUX包组。因为每个视频帧中包含45个VAUX包组，所以VAUX包组的编号从0至44。压缩流中必须包含一个VAUX源 (VAUX source, VS) 包组和一个VAUX源控制 (VAUX source control, VSC) 包组。尽管IEC 61834支持许多其他包组格式，但是当前，SMPTE 314M只支持VS和VSC包组。

#### 1. VAUX源 (VS) 包组

该包组的格式如表11-4所示。

表11-4 VAUX源 (VS) 包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	TVCH (十位, 0~9)				TVCH (个位, 0~9)			
PC2	B/W	EN	CLF		TVCH (百位, 0~9)			
PC3	SRC		50/60		STYPE			
PC4	TUN							
SMPTE 314M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	1	1	1	1	1	1	1	1
PC2	B/W	EN	CLF		1	1	1	1
PC3	1	1	50/60		STYPE			
PC4	VISC							
SMPTE 370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	0
PC1	1	1	1	1	1	1	1	1
PC2	1	1	1	1	1	1	1	1
PC3	1	1	50/60		STYPE			
PC4	0	1	1	1	1	1	1	1

TVCH为电视频道号，0~999。值0xEEE为预录制磁带或者线输入保留。值0xFFFF为“没有信息”保留。

B/W为黑白标志：0=黑白视频，1=彩色视频。

EN为CLF有效标志：0=CLF有效，1=CLF无效。

CLF为彩色帧标识码。对于480i系统：00=彩色帧A，01=彩色帧B，10=保留，11=保留。对于576i系统：00=第1、2场，01=第3、4场，10=第5、6场，11=第7、8场。

SRC定义视频信号的输入源。

50/60同AAUX。

STYPE表示对于SMPTE 314M，指定视频信号类型：00000=4:1:1压缩，00001~00011=保留，00100=4:2:2压缩，00101~11111=保留。对于SMPTE 370M，指定视频信号类型：00000~10011=保留，10100=1080i30或1080i25，10101=1035i30，10110=保留，10111=保留，11000=720p60或720p50，11001~11111=保留。对于IEC 61834，指定视频制式：00000=标准清晰度，00001=保留，00010=高清晰度，00011~11111=保留。

TUN为调谐器类别，包含3位区域号和5位卫星号。11111111表示没有信息可用。

VISC:  $10001000 = -180 \dots 00000000 = 0 \dots 01111000 = 180$ ,  $01111111 = \text{没有信息}$ , 其他值 = 保留。

## 2. VAUX源控制 (VSC) 包组

该包组的格式如表11-5所示。

表11-5 VAUX源控制 (VSC) 包组

IEC 61834	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	1
PC1	CGMS		ISR		CMP		SS	
PC2	REC S	1		REC M	1		DISP	
PC3	FF	FS	FC	IL	SF	SC	BCS	
PC4	1				GEN			

SMPTE 314M/370M	D7	D6	D5	D4	D3	D2	D1	D0
PC0	0	1	1	0	0	0	0	1
PC1	CGMS		1	1	1	1	1	1
PC2	1	1	0	0	1		DISP	
PC3	FF	FS	FC	IL	1	1	0	0
PC4	1	1	1	1	1	1	1	1

CGMS同AAUX。

ISR同AAUX。

CMP同AAUX。

SS同AAUX。

REC S同AAUX。

REC M表示录制模式: 00 = 原始, 01 = 保留, 10 = 插入, 11 = 无效录制。

BCS为广播制式, 指示带有DISP的显示格式的信息类型: 00 = 类型0 (IEC 61880、CEA-608), 01 = 类型1 (ETS 300 294), 10 = 保留, 11 = 保留。

DISP为宽高比信息。

FF为帧/场标志。指示两个场/帧是按顺序输出, 还是在一/两个帧周期内输出一个场两次: 0 = 一个场/帧输出两次, 1 = 两个场/帧按顺序输出。

FS为第一/第二标志。指示在场/帧1周期内输出哪个场/帧: 0 = 场/帧2, 1 = 场/帧1。

FC为帧变化标志。指示当前帧的图像是否与刚才前一帧的图像相同: 0 = 相同图像, 1 = 不同图像。

IL为隔行标志, 指示构成一帧的两场的数据是隔行的还是逐行的: 0 = 逐行, 1 = 隔行或者无法识别。

SF为静止场图像标志, 指示一帧内两场之间的时间差: 0 = 0秒, 1 = 1001/60或者50秒。

SC为照相机图像标志: 0 = 照相机图像, 1 = 非照相机图像。

GEN指示视频源的类别。

## 11.3 数字接口

IEC 61834和SMPTE 314M都指定了通用数字接口的数据格式。例如, 该数据格式可以通过IEEE 1394或者SDTI发送。图11-14描述了帧数据结构。

每个 $720 \times 480i$  4 : 1 : 1 YCbCr帧压缩为103 950B。包含开销和音频时数据量增加到120 000B。

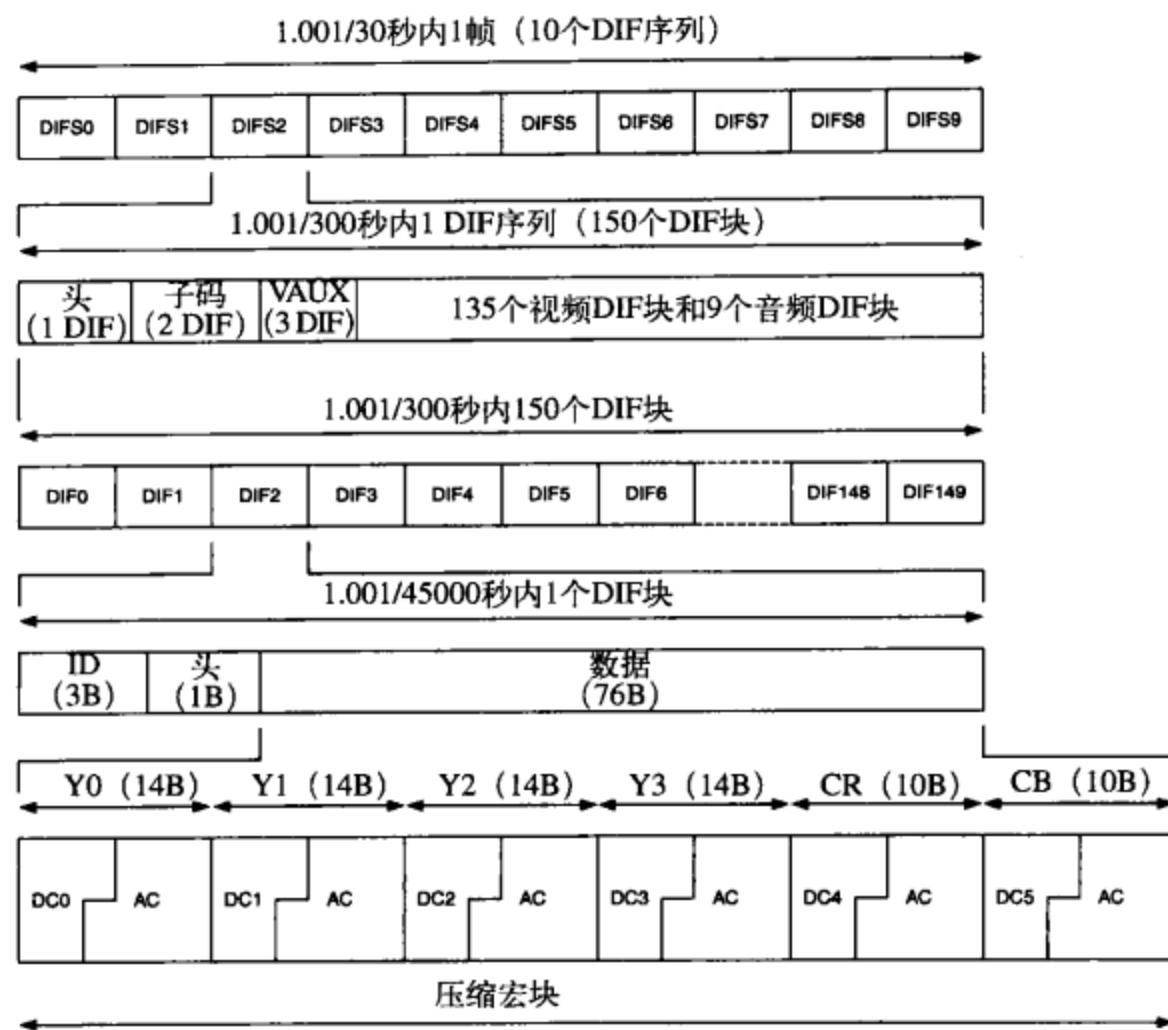


图11-14 25Mbit/s 4 : 1 : 1 YCbCr 720 × 480i系统的包格式

压缩 $720 \times 480i$ 帧分成10个DIF (data in frame, 帧中数据) 序列。每个DIF序列包含150个DIF块，每个DIF块80字节，DIF序列的分配情况如下所示：

- 135个DIF块用于视频；
- 9个DIF块用于音频；
- 6个DIF块用于头、子码和视频辅助 (VAUX) 信息。

图11-14详细描述了DIF序列的结构。每个视频DIF块包含80字节压缩宏块数据：

- 3字节用于DIF块ID信息；
- 1字节用于包含量化号 (QNO) 和块状态 (STA) 的头；
- Y0、Y1、Y2和Y3各使用14B；
- Cb和Cr各使用10B。

$720 \times 576i$ 帧可以使用4 : 2 : 0 YCbCr格式 (IEC 61834) 或者4 : 1 : 1 YCbCr格式 (SMPTE 314M)，而且需要12个DIF序列。每个 $720 \times 576i$ 帧压缩为124 740字节。包含开销和音频时数据量增加到144 000B，要求300个包进行传送。

注意通过接口传送的数据组织与实际DV录制格式有所不同，这是因为数字传输不需要错误校正。另外，尽管视频块按图11-15中的顺序编号，但是该序列并不对应于视频数据块从左至右、从上到下的传输顺序。压缩宏块散乱传送是为了使错误的影响最小，有助于错误隐蔽。音频数据也是散乱传送的。数据按与录制相同的散乱顺序传送。

为了描述视频的散乱性，DV视频帧被组织成超块，每个超块由27个压缩宏块构成，如图11-7~图11-11所示。5个超块为一组 (从每个超块列中取一个超块)，构成一个DIF序列。表11-6和表11-7描述了DIF块的传输顺序。

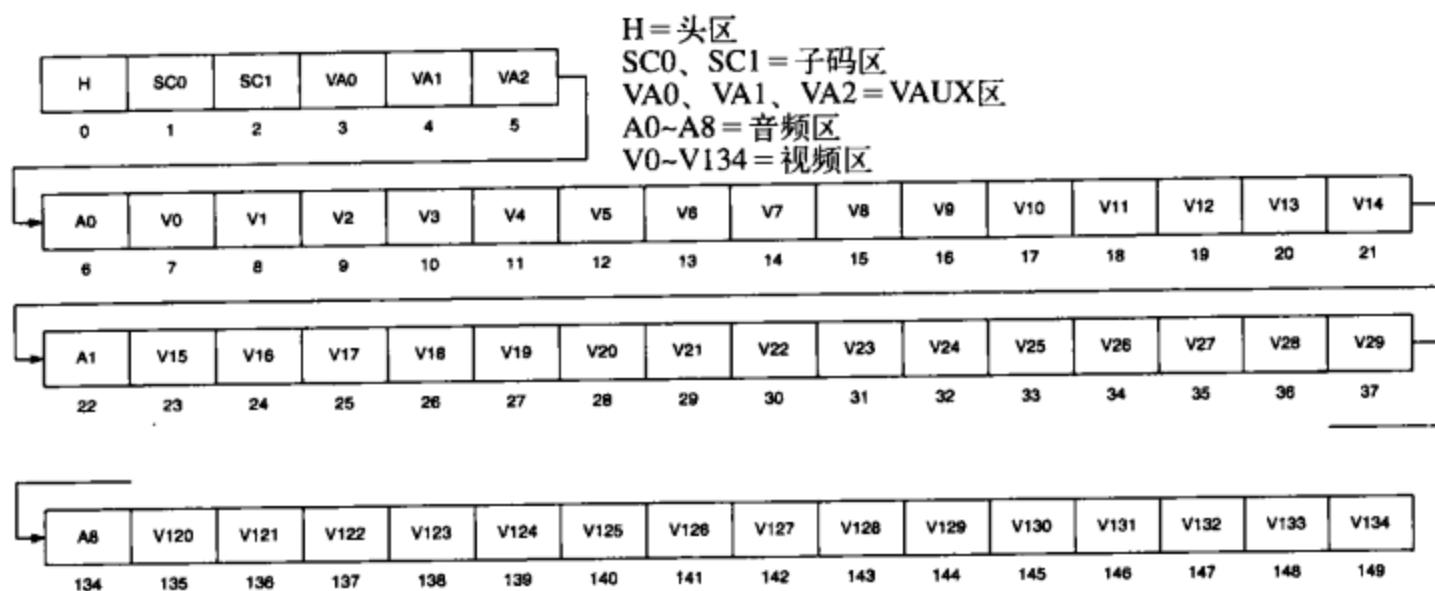


图11-15 DIF序列细节

表11-6 25Mbit/s (4:1:1或4:2:0 YCbCr) 的视频DIF块和压缩宏块

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
0	0	2, 2	0	n-1	0	1, 2	0
	1	6, 1	0		1	5, 1	0
	2	8, 3	0		2	7, 3	0
	3	0, 0	0		3	n-1, 0	0
	4	4, 4	0		4	3, 4	0
	:				133	n-1, 0	26
	133	0, 0	26		134	3, 4	26
	134	4, 4	26				
	0	3, 2	0				
	1	7, 1	0				
1	2	9, 3	0				
	3	1, 0	0				
	4	5, 4	0				
	:						
	133	1, 0	26				
	134	5, 4	26				

注：对于480i系统n=10，对于576i系统n=12。

表11-7 50Mbit/s (4:2:2 YCbCr) 的视频DIF块和压缩宏块

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
0	0, 0	4, 2	0	n-1	0, 0	2, 2	0
	0, 1	5, 2	0		0, 1	3, 2	0
	1, 0	12, 1	0		1, 0	10, 1	0
	1, 1	13, 1	0		1, 1	11, 1	0
	2, 0	16, 3	0		2, 0	14, 3	0
	:				134, 0	6, 4	26
	134, 0	8, 4	26				
	134, 1	9, 4	26				

(续)

DIF序列号	视频DIF块号	压缩宏块		DIF序列号	视频DIF块号	压缩宏块	
		超块号	宏块号			超块号	宏块号
1	0, 0	6, 2	0	$n-1$	134, 1	7, 4	26
	0, 1	7, 2	0				
	1, 0	14, 1	0				
	1, 1	15, 1	0				
	2, 0	18, 3	0				
	:						
	134, 0	10, 4	26				
	134, 1	11, 4	26				

注：对于480i系统 $n=10$ ，对于576i系统 $n=12$ 。

对于50Mbit/s的SMPTE 314M格式，每个压缩 $720 \times 480i$ 帧或 $720 \times 576i$ 帧分成2个通道。每个通道使用10个（480i系统）或者12个（576i系统）DIF序列。

### 11.3.1 IEEE 1394

使用IEEE 1394接口传送DV信息的内容在第6章中讨论。

### 11.3.2 SDTI

SDTI的一般概念在第6章中讨论。

#### 1. SMPTE 314M数据

SMPTE 221M详细描述了如何通过SDTI传送SMPTE 314M DV数据。

#### 2. IEC 61834数据

SMPTE 222M详细描述了如何通过SDTI传送IEC 61834 DV数据。

535

## 11.4 100Mbit/s DV不同之处

100Mbit/s SMPTE 370M格式支持 $1920 \times 1080i$ 和 $1280 \times 720p$ 源。

$1920 \times 1080i$ 源缩放为 $1280 \times 1080i$ 。 $1280 \times 720p$ 源缩放为 $960 \times 720p$ 。所用采样为4:2:2 YCbCr采样。

每个压缩帧分成4个通道。每个通道使用10个（ $1080i30$ 或者 $720p60$ 系统）或者12个（ $1080i25$ 或者 $720p50$ 系统）DIF序列。

## 11.5 HDV格式

HDV格式由佳能、夏普、索尼和JVC共同开发，HDV支持录制 $1920 \times 1080$ 和 $1280 \times 720$ 内容，使用的是标准的DV磁带。HDV格式基于25Mbit/s MPEG-2分组基本流和19Mbit/s MPEG-2传输流，视频压缩使用MPEG-2，音频压缩使用MPEG-1 Layer II。

## 11.6 AVCHD格式

AVCHD格式由松下和索尼公司共同开发，AVCHD支持录制 $1920 \times 1080$ 、 $1280 \times 720$ 、 $720 \times 576$ 和 $720 \times 480$ 内容，使用8cm DVD-RW、8cm BD-R/RE、SD存储卡或者HDD来代替磁带。AVCHD格式基于24Mbit/s MPEG-2传输流，视频压缩使用MPEG-4.10（H.264），音频压缩使用

536 Dolby® Digital或者LPCM。

## 参考文献

1. IEC 61834-1, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 1: General specifications.*
2. IEC 61834-2, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 2: SD format for 525-60 and 625-50 systems.*
3. IEC 61834-4, *Recording—Helical-scan digital video cassette recording system using 6.35mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)—Part 4: Pack header table and contents.*
4. SMPTE 314M-2005, *Television—Data Structure for DV-Based Audio, Data and Compressed Video-25 and 50 Mbps.*
5. SMPTE 321M-2002, *Television—Data Stream Format for the Exchange of DVBased Audio, Data and Compressed Video Over a Serial Data Transport Interface.*
6. SMPTE 322M-2004, *Television—Format for Transmission of DV Compressed Video, Audio and Data Over a Serial Data Transport Interface.*
7. SMPTE 370M-2006, *Television—Data Structure for DV-Based Audio, Data and Compressed Video at 100 Mb/s 1080/60i, 1080/50i, 720-60p, 720/50p.*

538



# 第12章 MPEG-1

MPEG-1视音频压缩标准是为存储和分发视频和音频数据而开发的，其特征包括随机存取、快进和快退播放等功能，是早期视频CD（即VCD）的基础标准。

在CD时代，人们根据可获得的媒体来定义信道带宽和图像分辨率，其目标是采用比特率为1.416Mbit/s（其中1.15Mbit/s用于视频）的标准光盘来播放数字视频和音频。

作为一个ISO标准，MPEG-1（ISO/IEC 11172）由以下6部分组成：

系统	ISO/IEC 11172-1
视频	ISO/IEC 11172-2
音频	ISO/IEC 11172-3
低比特率音频	ISO/IEC 13818-3
一致性测试	ISO/IEC 11172-4
仿真软件	ISO/IEC 11172-5

MPEG-1的比特流隐式地定义了解压缩算法。而压缩算法则由各厂商独立开发，并允许在国际标准范围内获得所属权益。

## 12.1 MPEG和JPEG

JPEG（ISO/IEC 10918）是为压缩静态连续色调（continuous-tone）的灰度图像或彩色图像而设计的。JPEG不善于表现只有黑白两种颜色的二值图，此外，如果要处理伪彩色图像，必须先将图像转换成未上色（unmapped color）的形式。JPEG图像可以针对任意分辨率和色彩空间，处理方法包括无损或者有损两种类型。

作为一种通用的标准，JPEG特色丰富，功能强大。通过调整各种各样的参数，压缩后图像大小与重建图像质量之间可以在一个较大的范围内获得折中。图像质量可以从“浏览级”（100：1的压缩比）到“无失真级”（约3：1的压缩比）之间变化。一般来说，源图像按10：1到20：1的比率压缩，才能看出重建图像与源图像之间存在视觉差异。

JPEG并不是采用单一算法，而是一系列的压缩算法，包含4种方式，分别应用于不同领域。其中最为人们熟知的有损压缩算法是顺序DCT，顺序DCT可能采用Huffman编码（基本JPEG）或算术编码。图像解码的时候，顺序从左到右，自上而下。

另一种有损压缩算法是累进DCT，需要对一幅图像进行多次扫描。图像解码的时候，首先得到整幅图像的粗糙效果，随后逐渐清晰，直至解码完成。这样做有利于图像数据库浏览等应用场合。累进DCT可能采用光谱选择或者连续逼近两个选项中的一个，或者两个选项都使用。为快速得到一幅图像，光谱选择选项首先对低频DCT系数编码，再逐渐增加细节，接着对高频系数编码。连续逼近选项首先对重要的DCT系数编码，再对次要的DCT系数编码。

层次化模式将图像表示为多种分辨率，如可能有 $512 \times 512$ 、 $1024 \times 1024$ 和 $2048 \times 2048$ 等几种版本的图像。对较高分辨率图像与其相邻最近的低分辨率图像之间的差异进行编码，用编码结果来表征较高分辨率图像，从而使得编码所需的比特数比两者独立存储时要少很多。当然，这样做也使得总的比特数比仅仅存储最高分辨率图像多了不少。此外，层次序列中的单个图像可能会根

据需要按照累进模式编码。

与变换域不同的是，在像素域还有一种无失真的空间压缩算法。该算法利用一个像素点邻近的三个像素点值估算该点的值，然后再减去该点的实际值，将差值采用Huffman或者算术编码方法进行无失真编码。无失真编码操作能达到约2：1的压缩比。

考虑到视频实际上是一系列静态图像，且基本JPEG编解码容易实现，因此有人提出利用基本JPEG压缩实时视频（这种方法也被称为运动JPEG或MJPEG）。但是，这种方法没有利用帧与帧之间的冗余来提高压缩能力，而MPEG利用了。

除此之外，也许最重要的一个问题，JPEG是对称的，即编码器和解码器的开销大致相似。与此不同的是，MPEG是以对视频进行一次压缩、而在各种平台多次播放为主要设计目的，因此，为最小化MPEG硬件解码器开销，MPEG被设计成为不对称的，即编码过程所需计算量是解码过程的100倍。

因为MPEG面向特定应用，所以其硬件仅支持一些特定的分辨率。同样，MPEG仅支持一种8位采样的颜色空间YCbCr。此外，MPEG仅在有限的压缩率范围中优化应用。

如果是将捕获视频用于编辑，可以采用基本JPEG或者单纯I帧（内码帧）MPEG将内容实时压缩到盘片。采用JPEG方法，需要系统为SIF（标准输入格式）分辨率提供4Mbit/s的传输和硬盘访问能力。一旦编辑完毕，结果可转换成MPEG以达到最大程度的压缩。

## 12.2 质量问题

在比特率约3~4Mbit/s的时候，视频传输采用MPEG-1可达到“广播级质量”。然而，对于复杂的时空场景变化（如体育运动节目）序列，由于MPEG-1基于帧的处理，导致传输比特率需求提高至5~6Mbit/s。540 MPEG-2支持基于场的处理，因而允许“广播级质量”的比特传输率达到4~6Mbit/s。

影响MPEG压缩视频质量的因素包括原始视频源的分辨率、压缩后信道带宽容许的比特率和运动估计效果。

影响压缩视频质量的一个重要因素是原始视频源的分辨率，如果原始分辨率过低，细节将普遍缺失。

运动估计效果决定运动失真情况，也就是当运动开始或者运动量超过一定限度时产生的视频质量下降状况。差的运动估计会导致视频质量效果整体下降。

最重要的是，比特率（信道带宽）越高，可以传输的信息量就会越大，此时运动失真减少，或者能够显示更高分辨率的图像。一般来说，降低比特率不会导致解码视频质量下降得太厉害。而一旦比特率下降到一定限度，视频质量将迅速下降， $8 \times 8$ 的方块将变得清晰可见。

## 12.3 音频概述

MPEG-1采用一系列共三个层次的音频编码方案，称为层I、层II和层III，计算复杂度和声音质量逐层提高。这三个层是分等级的：层III解码器处理层I、层II和层III；层II解码器只处理层I、层II；层I仅处理层I。所有层都支持采用16kHz、22.05kHz、24kHz、32kHz、44.1kHz或48kHz采样频率的16位音频。

对于每一层，其比特流格式和解码器是指定的。而编码器没有指定，允许后续改进。所有层工作在相似的比特率下。

- 层I：32~448kbit/s。
- 层II：8~384kbit/s。

### □ 层III：8~320kbit/s。

两个声道支持四种操作模式：正常立体声，联合（增强以及中间）立体声，双通道单声道，单通道单声道。在正常立体声模式下，一个信道传送左声道音频信号，另一个信道传送右声道音频信号。在增强立体声模式（所有层都支持）下，2kHz以上的高频子带将合成一个信号。在这种模式下，立体声像得以保持，但是仅传送时域包络。在中间立体声编码（仅层3支持）模式下，一个信道负责传送和信号（左声道+右声道），另一个传送差信号（左声道-右声道）。此外，还支持传送预加重（pre-emphasis）、版权标记以及原创/复制说明等信息。

#### 12.3.1 声音质量

为确定对于特定应用，具体应采用哪一层方案，需要考察各层可达到的比特率（因为各层都支持在一定比特率下，声音质量下降最小）。

层I，作为层II的简化版本，目标比特率是每信道192kbit/s或更高。

层II等同于MUSICAM，目标比特率为每信道128kbit/s，用于在声音质量与编码复杂度之间的折中。层II特别适合于比特率为每信道96~128kbit/s。

层III（即大家熟知的mp3）融合MUSICAM和ASPEC的思想，速率能达到每信道64kbit/s。层III格式定义了一系列高级特性，只为完成一个共同的目标：在相对低的码率下尽可能保证声音质量。

#### 12.3.2 背景理论

所有层均采用基于心理声学原则（特别是掩蔽效应）的编码策略。例如，根据掩蔽效应，收听到的响度较高的频率成分会掩盖响度较低的频率成分。

假设有一个强音调频率为1000Hz，另有一个1100Hz的音调，强度比前者低18dB。则1100Hz的音调将不会被听到，它被1000Hz的音调掩蔽了。然而，一个频率为2000Hz、强度比1000Hz的音调低18dB的音调则可以被听到。2000Hz的音调要想被1000Hz的音调掩蔽，则需要降低约45dB。任何相对低的频率都会被高频掩蔽；另外，距离声源越远，掩蔽效应越不明显。

因此，可以做出相对能量与掩蔽频率（当前掩蔽）之间的曲线图。如果响度变化显著（30~40dB），掩蔽效应可以在一个强音之前发生（预掩蔽）或之后发生（后掩蔽）。其原因是大脑需要反应时间。预掩蔽时间大概是2~5ms，而后掩蔽可以持续到100ms。

调整噪声基底可以减少需要的数据量，使得压缩效果更好。CD采用16位分辨率，可达到近96dB的信噪比（SNR），这正好符合人耳听力域的动态范围（也就是大多数人在播放静音时不会听到噪声）。如果采用8位分辨率，在音乐或者语声的间隙会出现明显的噪声。然而，由于隐蔽效应，高声带中的噪声不会被引起注意，也就是在声音较响的时候，可以提高噪声基底，因为噪声会被掩蔽。

对于立体声信号，信道间经常有冗余，所有层都可以采用联合立体声模式减少这些立体声效果，联合立体声是层III中采用的最灵活的方法。

## 12.4 视频编码层

MPEG-1允许在每秒60帧的逐行扫描方式下，分辨率最高可达到 $4095 \times 4095$ 。正如很多人认为的，MPEG-1是约束参数比特流（Constrained Parameters Bitstream，CPB）的子集。CPB是包括采样和比特率参数的有限集合，这些参数用来标准化缓存大小和存储带宽，使得编解码器有名义上确保的互操作，同时解决最大可能的应用范围。不能够处理这些参数的设备不能认为是真正的

MPEG-1。表12-1列出了一些约束参数。

CPB限制视频为396个宏块（共101 376像素）。因此，MPEG-1视频通常采用分辨率为 $352 \times 240p$ 或者 $352 \times 288p$ 的SIF格式编码。在编码中，分辨率为 $704 \times 480i$ 或 $704 \times 576i$ 的原始BT.601按比例缩减成SIF格式分辨率。这通常是通过忽略场2，在水平方向按比例缩减场1来实现的。在解码中，SIF分辨率被按比例提高成 $704 \times 480i$ 或者 $704 \times 576i$ 分辨率。注意，为保证Y分量数值是16的倍数，某些完整有效扫描线和扫描线上的像素会被抛弃。表12-2列出了比较常见的MPEG-1分辨率。  
[542]

表12-1 MPEG-1中的一些约束参数

线分辨率	$\leq 768$ 个样本点
场分辨率	$\leq 576$ 条扫描线
图像面积	$\leq 396$ 个宏块
像素速率	$\leq$ 每秒 $396 \times 25$ 个宏块
图像速率	$\leq$ 每秒30帧
比特率	$\leq 1.856\text{Mbit/s}$

表12-2 常见MPEG-1分辨率

分辨率	帧/秒
$352 \times 240p$	29.97
$352 \times 240p$	23.976
$352 \times 288p$	25
$320 \times 240p^1$	29.97
$384 \times 288p^1$	25

注：1. 方块像素格式

编码视频速率被控制在 $1.856\text{Mbit/s}$ 上，但是，在某些带宽为 $6\text{Mbit/s}$ 或更高的应用中，比特率是个常常被舍弃的参数。

MPEG-1视频数据采用 $4:2:0$ 的YCbCr格式，如图3-7所示。

#### 12.4.1 隔行视频

MPEG-1曾被用作处理逐行（即非隔行）视频。在早期，为尽力提高视频质量，人们设计了几种方法来充分利用隔行图像的两场数据。

例如，将两场合为一帧 $704 \times 480p$ 或 $704 \times 576p$ 分辨率的图像，并对此图像进行编码。在解码的时候，再将两场分开。然而，这样将导致运动对象产生运动效应，在两场中有些错位。而对两场独立编码虽然避免了运动效应，但是没利用两场中的冗余信息，降低了压缩率。

除此之外，还有其他一些处理隔行视频的方法，因此，MPEG-2定义了一个处理隔行视频的标准（这一部分将在第13章中详细介绍）。

#### 12.4.2 编码预处理

在进行MPEG编码之前，通过对视频流预处理，可以得到更好的图像。

在对特定图像编码的时候，为避免严重失真，可以对整幅图像（或者指定的区域）采取预滤波。在压缩处理前的预滤波与模/数转换前的抗混叠（anti-alias）滤波类似。预滤波考虑纹理样式、运动和边缘特点，可应用到图像、条带、宏块或者块等不同层次上。

对于快速或随机运动较少以及光照条件较好的场景，MPEG编码效果最好。此时，前景光应该清晰，背景光模糊。前景对比度和细节应当正常，而背景对比度最好能比较低，且边缘较为模糊。编辑工具通常会允许你预处理潜在问题区域。

MPEG-1规范中有将BT.601按比例缩减到SIF分辨率的滤波器样例。在实例中，场2被忽略，一半场分辨率被抛弃，并且采用抽取滤波器（抽取因子为2）减少余下扫描线的线分辨率。同样，对于Cb、Cr分量也必须进行适当的抽取。

按比例缩减成SIF分辨率前，对原始图像进行去隔行（deinterlacing）操作，可以得到更好的

视频质量。当处理宏块（这个名词稍后定义）的时候，如果两场间的宏块差别较小，可将两宏块求平均，产生一个新的宏块。另外，采用来自具有同奇偶性（same parity）场的宏块区域，可以避免运动效应的产生。

### 12.4.3 编码帧类型

编码帧分为4类。I (intra) 帧（约1比特/像素）是作为独立静态图像编码的帧。I帧提供在视频流中随机存取的随机存取点。通常，在1秒钟视频中约出现2次I帧。此外，在场景发生突变的地方也应该采用I帧。

P (predicted) 帧（约0.1比特/像素）是由与其相邻的前一个I帧或P帧预测而得到的帧（即所谓的前向预测处理），如图12-1所示。通过运动补偿技术，P帧比I帧压缩得更小，而且能够作为B帧和接下来的P帧的参考帧。

B (bi-directional) 帧（约0.015比特/像素）是采用最邻近的前后的两个I帧或P帧作为参考帧、进行双向预测得到的帧，如图12-1所示。B帧有最高的压缩效率，并利用前后两帧的均值减少了噪声。通常，I帧或P帧之间有2个B帧。

D (DC) 帧是仅包括DCT变换中的直流系数成分、按照独立静态图像编码的帧。D帧不能混在其他帧当中，而且很少使用。

图像组 (Group Of Picture, GOP) 是用于辅助随机存取和编辑的一帧或多帧编码帧序列。在编码过程中GOP值是可以设置的，GOP值越小，对运动的响应效果越好（因为I帧靠得越近），但压缩效率却下降了。

543

544

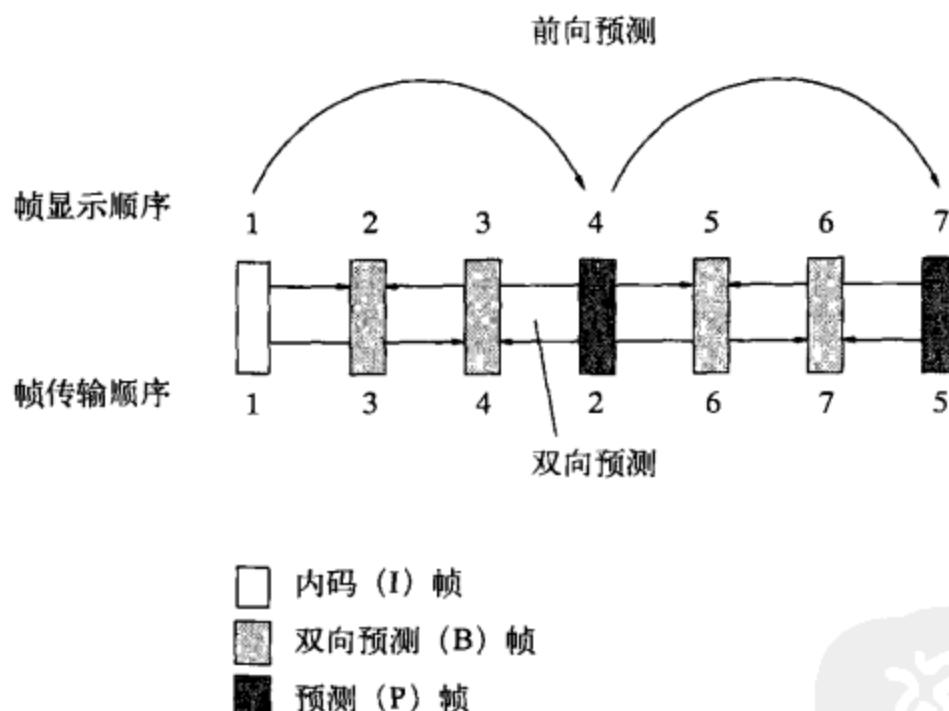


图12-1 MPEG-1中的I帧、P帧和B帧。不按照显示顺序传输的帧序列会使插值过程变得复杂，因此需要MPEG解码器对其进行重新排序。箭头表示帧之间的依赖关系

在编码比特流中，一个图像组必须以I帧作为开始，随后可以任意顺序接任意个I帧、P帧或B帧。在显示顺序中，图像组必须以I帧或B帧开始，以I帧或P帧结束。因此，最小的图像组是单个I帧，而最大图像组的大小是无限的。

最初，每个图像组不依赖其他图像组进行编码和显示。然而，这种情况只有没有B帧在I帧前才有可能，否则只能进行逆向运动补偿。于是形成开放和封闭两种图像组格式。封闭图像组不需要之前的图像组做运动补偿就能完成解码，而开放图像组是需要的。

#### 12.4.4 运动补偿

运动补偿用于消除帧间的时间冗余，提高对P帧和B帧的压缩效果。运动补偿主要在宏块层（稍后介绍）发挥作用。

运动补偿技术原理基于以下事实：在几幅大致相同图像的序列中，多数对象基本处在相同的位置，其余对象移动了一小段距离。这些运动被描述成一个二维运动矢量，这个二维运动矢量用来指定从前一个解码帧中的哪个位置获得一个宏块，该宏块用来预测当前宏块的像素值。

545

在采用运动补偿压缩宏块之后，待编码宏块和参考宏块之间存在空间差异（运动矢量）和内容差异（误差项）。

注意，有可能出现一个场景的信息不能从前一个场景预测出来的情况，例如打开一个门，先前的场景是没有包含门后区域的细节信息的。在这些情况下，当P帧中的宏块不能表征运动补偿的时候，将采用与I帧中的宏块编码方式相同的方式进行编码（采用帧间编码模式）。

B帧中的宏块采用最邻近的前后I帧或P帧作为参照，于是有四种编码方式：

- 帧内编码，没有运动补偿；
- 前向预测，最近的前一个I帧或P帧作为参考；
- 后向预测，最近的后一个I帧或P帧作为参考；
- 双向预测，两帧都用作参考——最近的前一个I帧或P帧，以及后一个I帧或P帧。

后向预测用于预测前一帧没有覆盖的区域。

#### 12.4.5 I帧

图像块和预测误差块有很高的空间冗余。为提高压缩率，去除冗余，采用了以下几个步骤，在解码的时候，采用相反的步骤可以恢复数据。

##### 1. 宏块

宏块（如图7-55所示）由16像素×16行的Y分量和对应的8像素×8行的Cb、Cr分量组成。

块是8像素×8行的Y、Cb、Cr值的集合。注意Y块大小为图像（即相应的Cb、Cr块）大小的四分之一。因此，一个宏块包含4个Y块、1个Cb块和一个Cr块，如图12-2所示。

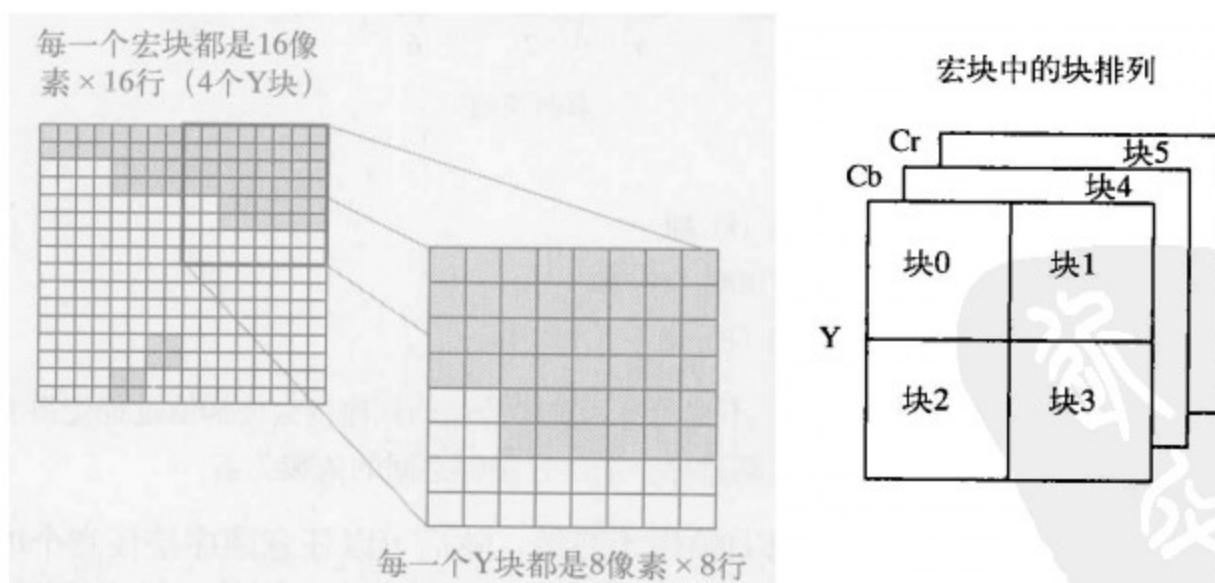


图12-2 MPEG-1宏块和块

I帧内有两种宏块类型，两者都采用帧内编码的，如表12-9所示。一种（称为intra-d）采用当前的量化器刻度，另一种（称为intra-q）采用新的量化器刻度。

如果宏块类型为intra-q，宏块首部指定一个量化器刻度因子，长度为5位。解码器用它从传送

的量化系数中计算DCT系数。量化器刻度因子可以从1到31，但不允许为0。

如果宏块类型为intra-d，将不传送量化器刻度，解码器使用当前的刻度。

## 2. DCT

每个输入像素或预测误差项的 $8 \times 8$ 块采用 $8 \times 8$ 的DCT（离散余弦变换）处理，得到 $8 \times 8$ 的水平和垂直方向的频率系数块，如图7-56所示。

输入像素值范围是0~255，得到范围为0~2040的DC系数和-1000~+1000的AC系数。

## 3. 量化

546

频率系数的 $8 \times 8$ 块是均匀量化的，取值有限。量化器步长刻度由量化矩阵和量化器刻度决定，不同的系数对应不同的量化器步长刻度，宏块与宏块间的量化器步长也有可能发生变化。

DC系数的量化器步长大小固定为8。DC量化系数由DC系数除以8取整得到。AC系数由帧内量化矩阵量化得到。

## 4. Z形扫描

如图7-57所示，Z形扫描从DC分量开始，按照频率递增的顺序将量化频率系数排列线性流，这将产生长游程的零系数。

## 5. 量化DC系数编码

DC系数量化后，将采用无失真编码方式对其进行编码。

宏块中的Y块编码按照图12-2所示的顺序进行。块4的DC值为下一个宏块中块1的DC值做预测。在每个条带的开头，DC预测值被置为1024。

每个Cb、Cr块的值采用先前宏块的对应块的DC值进行预测编码，在每个条带开头，DC预测值被置为1024。

DCT变换后的DC差分值按照绝对值形式组织，如表12-16所示。[size]项指定定义Level所需的附加比特数目，[size]用一个可变长编码传送，由于统计结果的不同，这个可变长编码值在Y和CbCr块中也是不同的。例如，size为4表示需要为4个附加比特。

解码器按照相反的过程反量化DC系数。

547

## 6. 量化AC系数编码

AC系数量化后，将按照图7-57所示的Z形顺序扫描，采用游程（run-length）和非零值（level）编码。如图7-57所示，扫描从位置1开始，因为在位置0的DC系数单独编码。

游程和非零值采用如表12-18所示方式编码。s位表示该非零值的符号，0为正，1为负。

对于表12-18中没有给出的游程-非零值组合的情况，引入了一个转义序列，由一个转义码（ESC）后面接着如表12-19所示的游程和非零值编码。

对最后一个DCT系数编码完成后，添加一个EOB符号，指示解码器在这个 $8 \times 8$ 块中已经没有其他量化系数了。

## 12.4.6 P帧

### 1. 宏块

根据运动补偿产生的复杂程度不同，P帧分为8种类型的宏块，如表12-10所示。

不编码宏块（skipped macroblock）为零运动矢量的预测宏块。因此，不需要修正；解码器复制先前帧的不编码宏块放在当前帧。不编码宏块的优点是它们几乎不需要传输，它们不需要编码，而是通过指明宏块地址增量把它们跳过。

如果表12-10中的[macroblock quant]项为1，那么需要传送量化器刻度。对剩下的宏块类型，采用先前的量化器刻度进行DCT修正编码。

如果表12-10中的[motion forward]项为1，水平和垂直前向运动矢量应当相继传输。

如果表12-10中的[coded pattern]项为1，6位编码块模式（coded block pattern）将作为变长编码传输。它会指示解码器宏块的6个块中哪些已经编码（设为1），哪些没有编码（设为0）。表12-14列出63种可能组合对应的码字。如果没有块被编码，就没有码字。编码块模式由宏块类型决定。对于I帧中的宏块以及P帧和B帧中的帧内编码宏块，编码块模式将不需要传输，且假定为63（表示所有块都已被编码）。

为确定采用何种类型的宏块，编码器通常需要做出一系列决策，如图12-3所示。

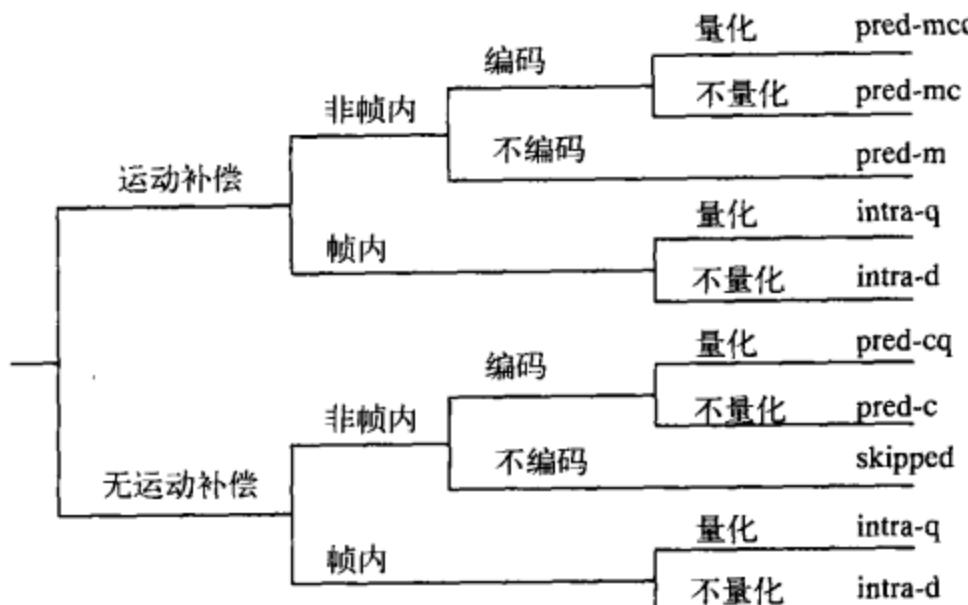


图12-3 MPEG-1 P帧宏块类型选择

## 2. DCT

帧内块AC系数采用与I帧相同方法的变换得到。帧内块DC系数变换方法有所不同，其预测值设为1024，除非先前的块是采用帧内编码。

非帧内块系数表示预测像素值之间的差异，而不是实际像素值。它们由当前宏块的值减去前一帧的运动补偿值得到。不预测DC值。

**548** 输入像素值为 $-255 \sim +255$ ，输出结果为范围从 $-2000 \sim +2000$ 的AC系数。

## 3. 量化

帧内块与I帧中的量化方法相同。

非帧内块采用量化器刻度和非帧内量化矩阵进行量化。AC和DC系数采用相同方式进行量化。

## 4. 帧内块编码

帧内块与I帧帧内块编码方式相同。在处理DC系数的时候有一点差异，其预测值为128，除非先前块采用帧内编码。

## 5. 非帧内块编码

编码块模式（Coded Block Pattern, CBP）用于指示哪些块有系数数据。除了DC系数与AC系数采用相同的编码方式以外，其编码方式与帧内块编码方式类似。

## 12.4.7 B帧

### 1. 宏块

根据后向运动补偿附加的复杂程度不同，B帧有12种宏块类型，如表12-11所示。

不编码宏块（skipped）为与前一个宏块有相同运动矢量和宏块类型的宏块，它不能采用帧内编码方式编码。不编码宏块的优点是它们几乎不占用传输带宽，不用编码，只需对宏块地址增量

进行编码，从而跳过对该宏块的编码。

如果表12-11中的[macroblock quant]项是1，将传输量化器刻度。而对于剩下的宏块类型，则采用前一个量化器刻度对DCT修正进行编码。549

如果表12-11中的[motion forward]项是1，将要相继传输水平和垂直前向运动矢量。如果表12-11中的[motion backward]项是1，将要相继传输水平和垂直后向运动矢量。如果前向后向运动类型都出现了，运动矢量按照以下顺序传输：

水平前向→垂直前向→水平后向→垂直后向

如果表12-11中的[coded pattern]项是1，6位长的编码块模式将采用变长编码方式传输，并指示解码器宏块的6个块中哪一块已经被编码（设为1），哪一块还没有被编码。表12-14列出63种可能组合对应的码字。如果还没有块被编码，就没有码字。编码块模式由宏块类型决定。对于I帧中的宏块和P帧、B帧中的帧内编码宏块，编码块模式不需要传输，且假定为63（表示所有块都已经被编码）。

为确定采用何种类型的宏块，编码器需要做出一系列决策，如图12-4所示。

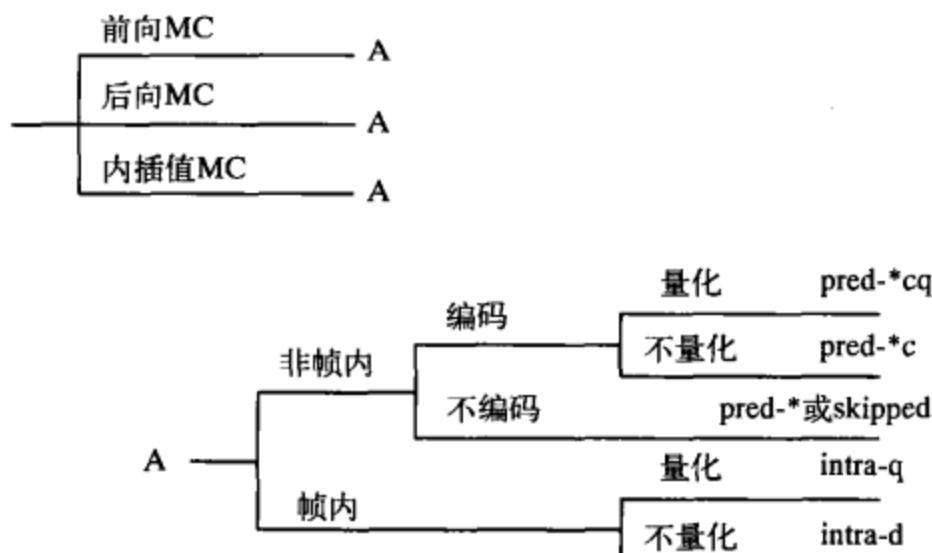


图12-4 MPEG-1 B帧宏块类型选择

## 2. 编码

块中的DCT系数将转换成量化系数，并且其编码方式与P帧的编码方式相同。

### 12.4.8 D帧

D帧只包含DC频率数据，一般用于快速可视化搜索应用。D帧包含的数据应当足以确保帮助用户定位所需的视频。550

## 12.5 视频比特流

图12-5对视频比特流进行了举例说明，从上到下总共7个层次，各层分别是：

- 视频序列
- 序列头
- 图像组 (GOP)
- 图像
- 条带
- 宏块 (MB)
- 块

注意，起始码（0x000001xx）必须是字节对齐的，其实现方式是在高位之前插入0~7个0达到。

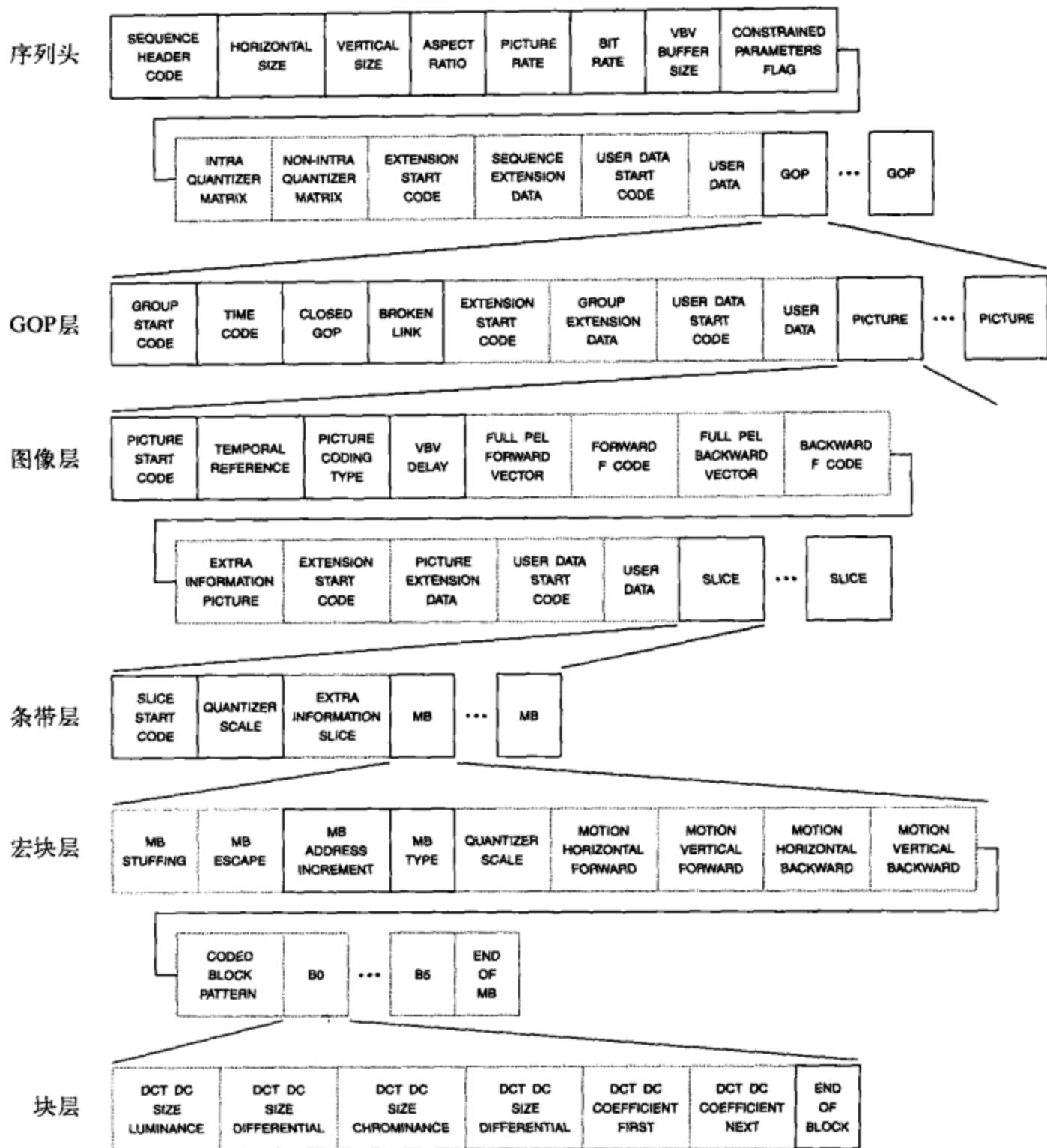


图12-5 MPEG-1视频比特流层次结构。标记位和保留位没有给出

### 12.5.1 视频序列

#### sequence\_end\_code

长度为32位的字段，值为0x000001B7，标志一个视频序列的结束。

### 12.5.2 序列头

每个序列的数据由序列头和随后的图像组（GOP）组成。结构如图12-5所示。

#### 1. sequence\_header\_code

长度为32位的字段，值为0x000001B3，表示序列头的开始。

#### 2. horizontal\_size

长度为12位的二进制值，表示Y分量可视部分的宽度。该宽度在宏块中定义为(horizontal\_size

$+15)/16$ 。

### 3. vertical\_size

长度为12位的二进制值，表示Y分量可视部分的高度。该高度在宏块中定义为 $(vertical\_size + 15)/16$ 。

### 4. pel\_aspect\_ratio

长度为4位的码字，表示像素的宽高比，如表12-3所示。

表12-3 MPEG-1 pel\_aspect\_ratio码字

高度/宽度	示例	宽高比编码
禁止		0000
1.0000	方形像素	0001
0.6735		0010
0.7031	576行16：9	0011
0.7615		0100
0.8055		0101
0.8437	480行16：9	0110
0.8935		0111
0.9157	576行4：3	1000
0.9815		1001
1.0255		1010
1.0695		1011
1.0950	480行4：3	1100
1.1575		1101
1.2015		1110
保留		1111

### 5. picture\_rate

长度为4位的码字，表示帧率，如表12-4所示。

表12-4 MPEG-1 picture\_rate码字

帧/秒	图像速率编码	帧/秒	图像速率编码
禁止	0000	60	1000
24/1.001	0001	保留	1001
24	0010	保留	1010
25	0011	保留	1011
30/1.001	0100	保留	1100
30	0101	保留	1101
50	0110	保留	1110
60/1.001	0111	保留	1111

### 6. bit\_rate

长度为18位的二进制值，表示比特流的比特率，以400bit/s为单位度量，不足时取上界。该值不允许为0。当值为0xFFFF时，表示此时为可变比特率操作。如果constrained\_parameters\_flag为1，那么比特率必须小于等于1.856Mbit/s。

### 7. marker\_bit

恒为1。

**8. vbv\_buffer\_size**

长度为10位的二进制数，表示解码器对视频解码时，校验缓冲需要的最小存储容量。按照如下定义：

$$B = 16 \times 1024 \times \text{vbv\_buffer\_size}$$

如果constrained\_parameters\_flag为1，vbv\_buffer\_size必须小于等于40KB。

**9. constrained\_parameters\_flag**

如果满足如下条件，此位设为1：

horizontal\_size  $\leq$  768像素点

vertical\_size  $\leq$  576行

$$[(\text{horizontal\_size} + 15)/16] \times [(\text{vertical\_size} + 15)/16] \leq 396$$

$$[(\text{horizontal\_size} + 15)/16] \times [(\text{vertical\_size} + 15)/16] \times \text{picture\_rate} \leq 396 * 25$$

picture\_rate  $\leq$  30帧/秒

forward\_f\_code  $\leq$  4

backward\_f\_code  $\leq$  4

**10. load\_intra\_quantizer\_matrix**

如果其后紧接intra\_quantizer\_matrix，此位设为1。如果设为0，采用如下默认值，直到下一个序列头出现：

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

**11. intra\_quantizer\_matrix**

可选的64个8位值列表，用来替换当前的帧内量化器值。不允许为0。intra\_quant[0,0]通常为8。这些数值一直有效，直到下一个序列头出现。

**12. load\_non\_intra\_quantizer\_matrix**

如果其后紧跟non\_intra\_quantizer\_matrix，此位设为1。如果设为0，采用如下默认值，直到下一个序列头出现：

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

**13. non\_intra\_quantizer\_matrix**

可选的64个8位值列表，用来替换当前的非帧内量化器值。不允许为0。这些数值一直有效，直到下一个序列头出现。

**14. extension\_start\_code**

可选的32位字符串，值为0x000001B5，表示sequence\_extension\_data（序列扩展数据）的开始。

`sequence_extension_data`一直持续，直到探测到下一个起始码。

#### 15. `sequence_extension_data`

长度为 $n \times 8$ 位，当`extension_start_code`出现时才出现。

#### 16. `user_data_start_code`

可选的32位字符串，值为0x000001B2，表示`user_data`（用户数据）的开始，`user_data`一直持续，直到探测到下一个起始码。

#### 17. `user_data`

长度为有 $n \times 8$ 位，当`user_data_start_code`出现时才出现。`user_data`不能包含超过连续23个0的字符串。

554

### 12.5.3 图像组层

每个图像组由图像组头及后面跟着的图像数据组成。结构如图12-5所示。

#### 1. `group_start_code`

长度为32位，值为0x000001B8，表示图像组的开始。

#### 2. `time_code`

长度为25位，表示时间编码信息，如表12-5所示。当且仅当图像帧率为30/1.001（29.97）Hz时，`[drop_frame_flag]`设为1。

#### 3. `closed_gop`

1位标志位，如果图像组没有参考前一个的图像组运动矢量而进行编码，此位设为1。此位允许支持编辑压缩比特流。

#### 4. `broken_link`

1位标志位，编码时此位设为0。在编辑过程中，当图像组第一个I帧后面的B帧不能正确地解码时，此位设为1。

#### 5. `extension_start_code`

可选的长度为32位的字符串，值为0x000001B5，表示`group_extension_data`（组扩展数据）。`group_extension_data`一直持续，直到下一个起始码出现。

#### 6. `group_extension_data`

当`extension_start_code`出现时才出现，长度为 $n \times 8$ 位。

#### 7. `user_data_start_code`

可选的长度为32位的字符串，值为0x000001B2，表示`user_data`的开始，`user_data`一直持续，直到下一个起始码出现。

#### 8. `user_data`

当`user_data_start_code`出现时才出现，长度为 $n \times 8$ 位。`user_data`不能包含23个或更多个连续0的字符串。

555

表12-5 MPEG-1 `time_code`字段

时间编码	取值范围	位数
<code>drop_frame_flag</code>		1
<code>time_code_hours</code>	0~23	5
<code>time_code_minutes</code>	0~59	6
<code>market_bit</code>	1	1
<code>time_code_seconds</code>	0~59	6
<code>time_code_pictures</code>	0~59	6

### 12.5.4 图像层

每个图像层的数据由图像头和紧跟其后的条带数据组成。其结构如图12-5所示。

#### 1. `picture_start_code`

长度为32位，值为0x00000100，表示图像的开始。

**2. temporal\_reference**

长度为10位的二进制，对于每个图像组显示顺序中的第一帧，temporal\_reference值为0。对于显示顺序中的其他每一帧，该值以1024为模，每次递增1。

**3. picture\_coding\_type**

长度为3位的码字，表示帧类型（I帧、P帧、B帧或D帧），如表12-6所示。D帧不与其他帧在同一视频序列使用。

**4. vbv\_delay**

长度为16位的二进制值。对于恒定比特率，vbv\_delay设定了解码图像开始时候缓存的占用空间，以免出现溢出或者下溢情况。对可变比特率，vbv\_delay值恒为0xFFFF。

**5. full\_per\_forward\_vector**

1位标志位，当picture\_coding\_type为010（P帧）或011（B帧）时，该位出现。若值为1，则前向运动矢量基于整像素，而不是半像素。

**6. forward\_f\_code**

长度为3位的码字，在picture\_coding\_type为010（P帧）或011（B帧）时出现，取值范围为001到111，不允许取值为000。

解码器采用这个字段派生出的两个参数解码前向运动矢量：forward\_r\_size和forward\_f。forward\_r\_size比forward\_f\_code小1，表12-7中给出了forward\_f的定义。

表12-6 MPEG-1 picture\_coding\_type编码

编码类型	编码
禁用	000
I帧	001
P帧	010
B帧	011
D帧	100
保留	101
保留	110
保留	111

表12-7 MPEG-1 forward\_f\_code值

forward_f_code	forward_f_code值
001	1
010	2
011	4
100	8
101	16
110	32
111	64

**7. full\_pel\_backward\_vector**

1位标志位，在picture\_coding\_type为011（B帧）时出现。值为1时，后向运动矢量基于整像素，而不是半像素。

**8. backward\_f\_code**

长度为3位的二进制数，在picture\_coding\_type为011（B帧）时出现。取值范围为001~111，000禁止使用。

解码器采用这个字段派生出的两个参数解码后向运动矢量：backward\_r\_size和backward\_f。backward\_r\_size比backward\_f\_code小1，backward\_f定义与forward\_f相同。

**9. extra\_bit\_picture**

1位标志位，设为1时，表示后面跟有extra\_information\_picture（附加信息图像）。

**10. extra\_information\_picture**

如果extra\_bit\_picture设为1，这9位将出现其后，其中包括8位数据（extra\_information\_picture）和1位extra\_bit\_picture，表明是否还跟着一个9位的extra\_information\_picture，如此继续。

**11. extension\_start\_code**

可选的长度为32位的字符串，值为0x000001B5，表示picture\_extension\_data（图像扩展数据）的开始，picture\_extension\_data一直持续，直到下一个起始码出现。

**12. picture\_extension\_data**

在extension\_start\_code出现时才出现，长度为 $n \times 8$ 位。

**13. user\_data\_start\_code**

可选的长度为32位的字符串，值为0x000001B2，表示user\_data的开始，user\_data一直持续，直到下一个起始码出现。

**14. user\_data**

在user\_data\_start\_code出现时才出现，长度为 $n \times 8$ 位。用户数据不能包含23个以上连续0的字符串。

### 12.5.5 条带层

每个条带层数据由条带头和紧跟其后的宏块数据组成。其结构如图12-5所示。

**1. slice\_start\_code**

长度为32位的字段，前24位值为0x000001。后8位为slice\_vertical\_position，值为0x01~0xAF。

slice\_vertical\_position表示条带中第一个宏块在宏块组（macroblock unit）中的垂直位置。宏块的第一行值为1。

**2. quantizer\_scale**

长度为5位的二进制数，取值范围为1~31（不允许为0）。指定DCT系数重构非零值的比例因子。解码器一直使用这个值，直到在条带层或者宏块层中接收到另一个quantizer\_scale。

**3. extra\_bit\_slice**

1位标志位，设为1时表示后面跟着extra\_information\_slice（附加信息条带）。

**4. extra\_information\_slice**

如果extra\_bit\_slice为1，这9位将出现其后，其中包括8位的数据（extra\_information\_slice）和1位的extra\_bit\_slice，表明是否在其后还跟着另一个9位的extra\_information\_slice。如此反复。

### 12.5.6 宏块层

每个宏块层的数据由宏块头和紧跟其后的运动矢量和块数据组成。其结构如图12-5所示。

**1. macroblock\_stuffing**

可选的长度为11位的固定比特字符串，值为0000 0001 111，可用于提高比特率以满足存储和传输的需要。可以采用任意个数的连续macroblock\_stuffing。

**2. macroblock\_escape**

可选的长度为11位的固定比特字符串，值为0000 0001 000，如果当前宏块地址和前一个宏块地址的差值超过33，将采用此项。它强制macroblock\_address\_increment的值变为33。可以采用任意个数的连续macroblock\_escape。

**3. macroblock\_address\_increment**

可变长码字，指明当前宏块和前一个宏块的地址差值。最大值为33。超过33的值采用macroblock\_escape编码。表12-8列出了这些可变长编码。

表12-8 MPEG-1 macroblock\_address\_increment可变长编码表

增量值	编 码	增量值	编 码
1	1	17	0000 0101 10
2	011	18	0000 0101 01
3	010	19	0000 0101 00
4	0011	20	0000 0100 11
5	0010	21	0000 0100 10
6	0001 1	22	0000 0100 011
7	0001 0	23	0000 0100 010
8	0000 111	24	0000 0100 001
9	0000 110	25	0000 0100 000
10	0000 1011	26	0000 0011 111
11	0000 1010	27	0000 0011 110
12	0000 1001	28	0000 0011 101
13	0000 1000	29	0000 0011 100
14	0000 0111	30	0000 0011 011
15	0000 0110	31	0000 0011 010
16	0000 0101 11	32	0000 0011 001
		33	0000 0011 000

#### 4. macroblock\_type

可变长码字，指明编码方法和宏块类型。如表12-9~表12-12所示。

表12-9 MPEG-1 I帧macroblock\_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
intra-d	0	0	0	0	1	1
intra-q	1	0	0	0	1	01

表12-10 MPEG-1 P帧macroblock\_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-mc	0	1	0	1	0	1
pred-c	0	0	0	1	0	01
pred-m	0	1	0	0	0	001
intra-d	0	0	0	0	1	0001 1
pred-mcq	1	1	0	1	0	0001 0
pred-cq	1	0	0	1	0	0000 1
intra-q	1	0	0	0	1	0000 01
skipped						

表12-11 MPEG-1 B帧macroblock\_type可变长编码表

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-i	0	1	1	0	0	10
pred-ic	0	1	1	1	0	11
pred-b	0	0	1	0	0	010
intra-bc	0	0	1	1	0	011

(续)

宏块类型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编码
pred-f	0	1	0	0	0	0010
pred-fc	0	1	0	1	0	0011
intra-d	0	0	0	0	1	0001 1
pred-icq	1	1	1	1	0	0001 0
pred-fcq	1	1	0	1	0	0000 11
pred-bcq	1	0	1	1	0	0000 10
intra-q	1	0	0	0	1	0000 01
skipped						

表12-12 MPEG-1 D帧macroblock\_type可变长编码表

Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macroblock	编 码
0	0	0	0	1	1

**5. quantizer\_scale**

可选的长度为5位的二进制数，值为1~31（不允许为0），指示接收的DCT系数重构非零值的比例因子。解码器在接收下一个来自条带层或宏块层的quantizer\_scale前始终采用当前值。quantizer\_scale仅在[macroblock quant]为1时出现，如表12-9~表12-12所示。

**6. motion\_horizontal\_forward\_code**

可选的变长码字，包含前向运动矢量信息，如表12-13所示。当[motion forward]为1时出现，如表12-9~表12-12所示。

558

**7. motion\_horizontal\_forward\_r**

可选的二进制数（属于forward\_r\_size位），用于帮助解码前向运动矢量。当[motion forward]为1时出现，如表12-9~表12-12所示，forward\_f\_code ≠ 001且motion\_horizontal\_forward\_code ≠ 0。

**8. motion\_vertical\_forward\_code**

可选的变长码字，包含如表12-13所示的前向运动矢量信息，当[motion forward]为1时出现，如表12-9~表12-12所示。

表12-13 MPEG-1 motion\_horizontal\_forward\_code、motion\_vertical\_forward\_code、motion\_horizontal\_backward\_code和motion\_vertical\_backward\_code可变长编码表

运动矢量差分	编 码	运动矢量差分	编 码
-16	0000 0011 001	1	010
-15	0000 0011 011	2	0010
-14	0000 0011 101	3	000 10
-13	0000 0011 111	4	0000 110
-12	0000 0100 001	5	0000 1010
-11	0000 0100 011	6	0000 1000
-10	0000 0100 11	7	0000 0110
-9	0000 0101 01	8	0000 0101 10
-8	0000 0101 11	9	0000 0101 00
-7	0000 0111	10	0000 0100 10
-6	0000 1001	11	0000 0100 010
-5	0000 1011	12	0000 0100 000

(续)

运动矢量差分	编 码	运动矢量差分	编 码
-4	0000 111	13	0000 0011 110
-3	0001 1	14	0000 0011 100
-2	0011	15	0000 0011 010
-1	011	16	0000 0011 000
0	1		

**9. motion\_vertical\_forward\_r**

可选的二进制数（属于forward\_r\_size位），用于帮助解码前向运动矢量。当表12-9~表12-12中的[motion forward]为1时出现。forward\_f\_code ≠ 001且motion\_vertical\_forward\_code ≠ 0。

**10. motion\_horizontal\_backward\_code**

可选的变长码字，包含表12-13中定义的后向运动矢量信息。当表12-9~表12-12中的[motion backward]为1时出现。

**11. motion\_horizontal\_backward\_r**

可选的二进制数（属于backward\_r\_size位），用于帮助解码后向运动矢量。当表12-9~表12-12中的[motion backward]为1时出现。backward\_f\_code ≠ 001且motion\_horizontal\_backward\_code ≠ 0。

**12. motion\_vertical\_backward\_code**

可选的变长码字，包含如表12-13中定义的后向运动矢量信息，解码值帮助决定motion\_vertical\_backward\_r是否在比特流出现。当表12-9~表12-12中的[motion backward]为1时出现。

**13. motion\_vertical\_backward\_r**

可选的二进制数（属于backward\_r\_size位），用于帮助解码后向运动矢量。当表12-9~表12-12中的[motion backward]为1时出现。backward\_f\_code ≠ 001且motion\_vertical\_backward\_code ≠ 0。

**14. coded\_block\_pattern**

可选的变长码字，用于表示如表12-14所示的编码块模式。当表12-9~表12-12中的[coded pattern]为1时候出现。指示宏块中哪些（至少一个）变换系数需要传输。编码块模式的二进制编码采用如下形式表示：

P<sub>0</sub>P<sub>1</sub>P<sub>2</sub>P<sub>3</sub>P<sub>4</sub>P<sub>5</sub>

当块第[n]个系数出现时，对应的P<sub>n</sub>为1，否则为0。图12-2给出块编号（十进制格式）方法。

表12-14 MPEG-1 coded\_block\_pattern可变长编码表

编码块模式	编 码	编码块模式	编 码	编码块模式	编 码
60	111	9	0010 110	43	0001 0000
4	1101	17	0010 101	25	0000 1111
8	1100	33	0010 100	37	0000 1110
16	1011	6	0010 011	26	0000 1101
32	1010	10	0010 010	38	0000 1100
12	1001 1	18	0010 001	29	0000 1011
48	1001 0	34	0010 000	45	0000 1010
20	1000 1	7	0001 1111	53	0000 1001
40	1000 0	11	0001 1110	57	0000 1000
28	0111 1	19	0001 1101	30	0000 0111
44	0111 0	35	0001 1100	46	0000 0110

(续)

编码块模式	编 码	编码块模式	编 码	编码块模式	编 码
52	0110 1	13	0001 1011	54	0000 0101
56	0110 0	49	0001 1010	58	0000 0100
1	0101 1	21	0001 1001	31	0000 0011 1
61	0101 0	41	0001 1000	47	0000 0011 0
2	0100 1	14	0001 0111	55	0000 0010 1
62	0100 0	50	0001 0110	59	0000 0010 0
24	0011 11	22	0001 0101	27	0000 0001 1
36	0011 10	42	0001 0100	39	0000 0001 0
3	0011 01	15	0001 0011		
63	0011 00	51	0001 0010		
5	0010 111	23	0001 0001		

**15. end\_of\_macroblock**

可选的1位标志位，值为1。仅在D帧中出现。

### 12.5.7 块层

每个块层数据都由系数数据组成，结构如图12-5所示。

**1. dct\_dc\_size\_luminance**可选的可变长码字，用于Y块帧内编码。指示用于dct\_dc\_differential的位数目。表12-15列出了这些变长码字。559  
562

表12-15 MPEG-1 dct\_dc\_size\_luminance可变长编码表

dct_dc_size_luminance	编 码	dct_dc_size_luminance	编 码
0	100	5	1110
1	00	6	1111 0
2	01	7	1111 10
3	101	8	1111 110
4	110		

563

**2. dct\_dc\_differential**

可选的可变长码字，当dct\_dc\_size\_luminance不为0时，在dct\_dc\_size\_luminance之后出现，其值如表12-16所示。

表12-16 MPEG-1 dct\_dc\_differential可变长编码表

dct_dc_differential	Size	编码 (Y)	编码 (CbCr)	附加编码
-255~-128	8	1111110	11111110	00000000~01111111
-127~-64	7	111110	1111110	0000000~0111111
-63~-32	6	11110	111110	000000~011111
-31~-16	5	1110	11110	00000~01111
-15~-8	4	110	1110	0000~0111
-7~-4	3	101	110	000~011
-3~-2	2	01	10	00~01
-1	1	00	01	0

(续)

dct_dc_differential	Size	编码 (Y)	编码 (CbCr)	附加编码
0	0	100	00	
1	1	00	01	1
2~3	2	01	10	10~11
4~7	3	101	110	100~111
8~15	4	110	1110	1000~1111
16~31	5	1110	11110	10000~11111
32~63	6	11110	111110	100000~111111
64~127	7	111110	1111110	1000000~1111111
128~255	8	1111110	11111110	10000000~11111111

### 3. dct\_dc\_size\_chrominance

可选的变长码字，用于Cb、Cr块的帧内编码。指示用于dct\_dc\_differential的位数目。表12-17列出了这些可变长码字。

表12-17 MPEG-1 dct\_dc\_size\_chrominance可变长编码表

dct_dc_size_chrominance	编 码	dct_dc_size_chrominance	编 码
0	00	5	1111 0
1	01	6	1111 10
2	10	7	1111 110
3	110	8	1111 1110
4	1110		

### 4. dct\_dc\_differential

可选的可变长码字，当dct\_dc\_size\_chrominance不为0时，在dct\_dc\_size\_chrominance之后出现。其值如表12-16所示。

### 5. dct\_coefficient\_first

可选的可变长码字，用于非帧内编码块的第一个DCT系数，其定义在表12-18和表12-19中给出。

表12-18 MPEG-1 dct\_coefficient\_first和dct\_coefficient\_next可变长编码表

RUN	LEVEL	编 码 <sup>1</sup>	RUN	LEVEL	编 码
end_of_block		10	escape		0000 01
0 <sup>2</sup>	1	1 s	0	5	0010 0110 s
0 <sup>3</sup>	1	11 s	0	6	0010 0001 s
1	1	011 s	1	3	0010 0101 s
0	2	0100 s	3	2	0010 0100 s
2	1	0101 s	10	1	0010 0111 s
0	3	0010 1 s	11	1	0010 0011 s
3	1	0011 1 s	12	1	0010 0010 s
4	1	0011 0 s	13	1	0010 0000 s
1	2	0001 10 s	0	7	0000 0010 10 s
5	1	0001 11 s	1	4	0000 0011 00 s
6	1	0001 01 s	2	3	0000 0010 11 s
7	1	0001 00 s	4	2	0000 0011 11 s
0	4	0000 110 s	5	2	0000 0010 01 s

(续)

RUN	LEVEL	编 码 <sup>1</sup>	RUN	LEVEL	编 码
end_of_block		10	escape		0000 01
2	2	0000 100 s	14	1	0000 0011 10 s
8	1	0000 111 s	15	1	0000 0011 01 s
9	1	0000 101 s	16	1	0000 0010 00 s
0	8	0000 0001 1101 s	0	12	0000 0000 1101 0 s
0	9	0000 0001 1000 s	0	13	0000 0000 1100 1 s
0	10	0000 0001 0011 s	0	14	0000 0000 1100 0 s
0	11	0000 0001 0000 s	0	15	0000 0000 1011 1 s
1	5	0000 0001 1011 s	1	6	0000 0000 1101 0 s
2	4	0000 0001 0100 s	1	7	0000 0000 1010 1 s
3	3	0000 0001 1100 s	2	5	0000 0000 1010 0 s
4	3	0000 0001 0010 s	3	4	0000 0000 1001 1 s
6	2	0000 0001 1110 s	5	3	0000 0000 1001 0 s
7	2	0000 0001 0101 s	9	2	0000 0000 1000 1 s
8	2	0000 0001 0001 s	10	2	0000 0000 1000 0 s
17	1	0000 0001 1111 s	22	1	0000 0000 1111 1 s
18	1	0000 0001 1010 s	23	1	0000 0000 1111 0 s
19	1	0000 0001 1001 s	24	1	0000 0000 1110 1 s
20	1	0000 0001 0111 s	25	1	0000 0000 1110 0 s
21	1	0000 0001 0110 s	26	1	0000 0000 1101 1 s
0	16	0000 0000 0111 11 s	0	40	0000 0000 0010 000 s
0	17	0000 0000 0111 10 s	1	8	0000 0000 0011 111 s
0	18	0000 0000 0111 01 s	1	9	0000 0000 0011 110 s
0	19	0000 0000 0111 00 s	1	10	0000 0000 0011 101 s
0	20	0000 0000 0110 11 s	1	11	0000 0000 0011 100 s
0	21	0000 0000 0110 10 s	1	12	0000 0000 0011 011 s
0	22	0000 0000 0110 01 s	1	13	0000 0000 0011 010 s
0	23	0000 0000 0110 00 s	1	14	0000 0000 0011 001 s
0	24	0000 0000 0101 11 s	1	15	0000 0000 0001 0011 s
0	25	0000 0000 0101 10 s	1	16	0000 0000 0001 0010 s
0	26	0000 0000 0101 01 s	1	17	0000 0000 0001 0001 s
0	27	0000 0000 0101 00 s	1	18	0000 0000 0001 0000 s
0	28	0000 0000 0100 11 s	6	3	0000 0000 0001 0100 s
0	29	0000 0000 0100 10 s	11	2	0000 0000 0001 1010 s
0	30	0000 0000 0100 01 s	12	2	0000 0000 0001 1001 s
0	31	0000 0000 0100 00 s	13	2	0000 0000 0001 1000 s
0	32	0000 0000 0011 000 s	14	2	0000 0000 0001 0111 s
0	33	0000 0000 0011 111 s	15	2	0000 0000 0001 0110 s
0	34	0000 0000 0011 110 s	16	2	0000 0000 0001 0101 s
0	35	0000 0000 0010 101 s	27	1	0000 0000 0001 1111 s
0	36	0000 0000 0010 100 s	28	1	0000 0000 0001 1110 s
0	37	0000 0000 0010 011 s	29	1	0000 0000 0001 1101 s
0	38	0000 0000 0010 010 s	30	1	0000 0000 0001 1100 s
0	39	0000 0000 0010 001 s	31	1	0000 0000 0001 1011 s

注：1. s表示非零系数（level）的符号，0表示正，1表示负。

2. 用于dct\_coefficient\_first。

3. 用于dct\_coefficient\_next。

表12-19 dct\_coefficient\_first和dct\_coefficient\_next中转义码后的游程、非零系数编码

RUN	LEVEL	固定长度编码
0		0000 00
1		0000 01
2		0000 10
:		:
63		1111 11
	-256	禁止
	-255	1000 0000 0000 0001
	-254	1000 0000 0000 0010
	:	:
	-129	1000 0000 0111 1111
	-128	1000 0000 1000 0000
	-127	1000 0001
	-126	1000 0010
	:	:
	-2	1111 1110
	-1	1111 1111
	0	禁止
	1	0000 0001
	:	:
	127	0111 1111
	128	0000 0000 1000 0000
	129	0000 0000 1000 0001
	:	:
	255	0000 0000 1111 1111

#### 6. dct\_coefficient\_next

最大为63的可选的可变长码字，仅用于I帧、P帧和B帧。表示在第一个系数之后的DCT系数，其定义在表12-18和表12-19中给出。

#### 7. end\_of\_block

长度为2位，仅在I帧、P帧和B帧中出现，表示其后面没有其他非零系数。该参数值为10。

## 12.6 系统比特流

系统比特流将视频音频比特流复用到一个单独的比特流，并且与控制信息一起格式化成MPEG-1定义的指定协议。

包数据可能包含视频或音频信息。至多可复用16路视频和32路音频，并支持两种类型的私有数据流。一种是完全私有的，另一种用于支持同步和缓冲管理。

尽管数据包长度可以非常大，但通常最大的长度为2048字节。在CD-ROM存储时，数据包长度与扇区大小一致。典型情况是6个或7个视频数据包附带一个音频数据包。

图12-6举例说明了系统比特流的3层等级结构，自顶向下分别是：

- ISO/IEC 11172层
- 包组 (Pack)
- 包 (Packet)

注意，起始码（0x000001xx）必须为字节对齐，实现方式是在起始码之前插入0~7个0位。

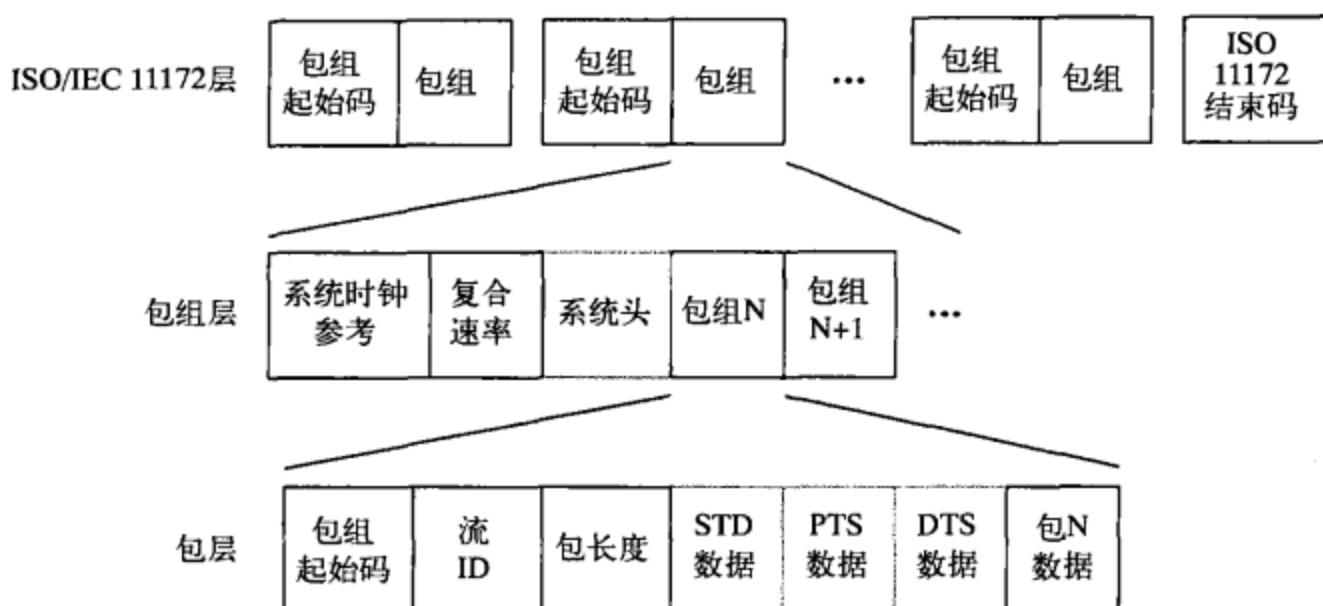


图12-6 MPEG-1系统比特流层结构，标记位和保留位没有给出

### 12.6.1 ISO/IEC 11172层

**ISO\_11172\_end\_code**

长度为32位的字段，值为0x000001B9，用于终止一段系统比特流。

### 12.6.2 包组层

每个包组由包组头和紧跟其后的一个系统头（可选的）和包数据组成。结构如图12-6所示。

1. **pack\_start\_code**

长度为32位的字段，值为0x000001BA，表示包组的开始。

2. **fixed\_bits**

长度为4位的二进制数，值恒为0010。

3. **system\_clock\_reference\_32~30**

system\_clock\_reference (SCR) 为长度为33位的数，该数由标记位分成的三部分编码而成。

system\_clock\_reference 表示 system\_clock\_reference 字段的最后一个字节到达解码器输入处的预计时间。system\_clock\_reference 的值为 90kHz 时钟周期倍数。

4. **marker\_bit**

标记位，值恒为1。

5. **system\_clock\_reference\_29~15**

6. **marker\_bit**

标记位，值恒为1。

7. **system\_clock\_reference\_14~0**

8. **marker\_bit**

标记位，值恒为1。

9. **marker\_bit**

标记位，值恒为1。

10. **mux\_rate**

长度为22位的二进制数，指示解码器接收比特流的速率。以50字节/秒为单位，不足的取上

界。不允许为0。

#### 11. marker\_bit

标记位，值恒为1。

### 12.6.3 系统头

#### 1. system\_header\_start\_code

长度为32位的字段，值为0x000001BB，表示系统头的开始。

#### 2. header\_length

长度为16位的二进制数，指示跟在系统头之后的字节数。

#### 3. marker\_bit

标记位，值恒为1。

#### 4. rate\_bound

长度为22位的二进制数，指定一个大于等于mux\_rate最大值的整数值，解码器可以据此确定是否有能力对整个比特流解码。

#### 5. marker\_bit

标记位，值恒为1。

#### 6. audio\_bound

长度为6位的二进制数，取值范围为0~32，指定一个大于等于同时活动的音频流最大数的整数值。

#### 7. fixed\_flag

571 1位标志位，指示操作是固定比特率（为1）还是可变比特率（为0）。

#### 8. CSPS\_flag

1位标志位，指示比特流是约束系统参数流（为1），否则为0。

#### 9. system\_audio\_lock\_flag

如果音频采样率和解码器系统时钟频率保持固定关系，该位值为1。

#### 10. system\_video\_lock\_flag

如果视频图像速率和解码器系统时钟频率保持固定关系，该位值为1。

#### 11. marker\_bit

标记位，值恒为1。

#### 12. video\_bound

长度为5位的二进制数，取值范围为0~16，指定一个大于等于同时活动的视频流最大数目的整数值。

#### 13. reserved\_byte

长度为8位，通常为1111 1111。

#### 14. stream\_ID

可选的8位的字段，如表12-20定义。表示接下来的与STD\_buffer\_bound\_scal和STD\_buffer\_size\_bound相关的类型和流编号。每个系统比特流中出现的视频和音频流在系统头中只出现一次。

表12-20 MPEG-1 stream\_ID码

流类型	流ID
所有音频流	1011 1000
所有视频流	1011 1001
保留流	1011 1100
私有流1	1011 1101
填充流	1011 1110
私有流2	1011 1111
音频流编号xxxx	110x xxxx
视频流编号xxxx	1110 xxxx
保留数据流编号xxxx	1111 xxxx

**15. fixed\_bits**

可选的长度为2位的字段，值为11。当stream\_ID出现时才出现。

**16. STD\_buffer\_bound\_scale**

可选的1位字段，指示用于解释STD\_buffer\_size\_bound的缩放因子。对于音频流，该字段值为0；对于视频流，该字段值为1。对其他流类型，1和0都可以。当stream\_ID出现时该字段才出现。 572

**17. STD\_buffer\_size\_bound**

可选的长度为13位的二进制数，指定一个大于或等于解码器输入缓冲最大值的整数值。当STD\_buffer\_bound\_scale为0时，STD\_buffer\_size\_bound以128字节为单位。当STD\_buffer\_bound\_scale为1时，以1024字节为单位。该字段仅当stream\_ID出现时才出现。

#### 12.6.4 包层

**1. packet\_start\_code\_prefix**

长度为24位的字段，值为0x000001，跟在stream\_ID相应的参数项后面，表示包的开始。

**2. stream\_ID**

长度为8位的二进制数，指定当前比特流的类型和数量。如表12-20所示。

**3. packet\_length**

长度为16位二进制数，指定包的字节数。

**4. stuffing\_byte**

可选参数，值为1111 1111。根据存储媒介需要，至多可使用16个连续的stuffing\_bytes。当stream\_ID不为私有流2时出现。

**5. STD\_bits**

可选的长度为2位二进制数，值为01，表示接下来跟着STD\_buffer\_scale和STD\_buffer\_size。当stream\_ID不为私有流2时出现。

**6. STD\_buffer\_scale**

可选的1位字段，指示解释STD\_buffer\_size的比例因子，值为0时表示音频流，为1时表示视频流，对其他类型流，0和1都可以。当stream\_ID不为私有流2时出现。

**7. STD\_buffer\_size**

可选的长度为13位的二进制数，指定解码器输入缓冲的大小。当STD\_buffer\_scale为0时，STD\_buffer\_size以128字节为单位，当STD\_buffer\_scale为1，以1024字节为单位。当stream\_ID不为私有流2时出现。

**8. PTS\_bits**

可选的4位二进制数，值为0010，表示接下来将出现展示时间戳选项，当stream\_ID不为私有流2时出现。

**9. presentation\_time\_stamp\_32~30**

可选的33位二进制数，presentation\_time\_stamp (PTS) 由标记位分成3段来编码，表示解码器显示的计划时间。PTS取值为90kHz系统时钟的倍数。当PTS\_bits出现且stream\_ID不为私有流2时出现。

**10. marker\_bit**

可选标记位，值恒为1，当PTS\_bits出现且stream\_ID不为私有流2的时候出现。

**11. presentation\_time\_stamp\_29~15**

可选，当PTS\_bits出现且stream\_ID不为私有流2的时候出现。

**12. marker\_bit**

**573** 可选的标记位，值恒为1。当PTS\_bits出现且stream\_ID不为私有流2时出现。

**13. presentation\_time\_stamp\_14~0**

可选字段，当PTS\_bits出现且stream\_ID不为私有流2的时候出现。

**14. marker\_bit**

可选标记位，值恒为1。当PTS\_bits出现且stream\_ID不为私有流2时出现。

**15. DTS\_bits**

可选的长度为4位的二进制数，值为0011，表示接下来将出现展示和解码时间戳选项。当stream\_ID不为私有流2时出现。

**16. presentation\_time\_stamp\_32~30**

可选的长度为33位数，presentation\_time\_stamp (PTS) 由标记位分成3段来编码，表示解码显示的计划时间。PTS取值为90kHz系统时钟的倍数。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**17. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**18. presentation\_time\_stamp\_29~15**

可选字段，当DTS\_bits出现且stream\_ID不为私有流2时出现。

**19. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**20. presentation\_time\_stamp\_14~0**

可选字段，当DTS\_bits出现且stream\_ID不为私有流2时出现。

**21. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**22. fixed\_bits**

可选长度为4位的二进制数，值为0001，当DTS\_bits出现且stream\_ID不为私有流2时出现。

**23. decoding\_time\_stamp\_32~30**

可选的长度为33位的二进制数，decoding\_time\_stamp (DTS) 由标记位分成3段来编码，表示解码显示的计划时间。DTS取值为90kHz系统时钟的倍数。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**24. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**25. decoding\_time\_stamp\_29~15**

可选字段，当DTS\_bits出现且stream\_ID不为私有流2时出现。

**26. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**27. decoding\_time\_stamp\_14~0**

可选字段，当DTS\_bits出现且stream\_ID不为私有流2时出现。

**28. marker\_bit**

可选标记位，值恒为1。当DTS\_bits出现且stream\_ID不为私有流2时出现。

**29. nonPTS\_nonDTS\_bits**

可选的长度为8位的二进制数，值为00001111，当STD\_bits、PTS\_bits或DTS\_bits（及其相应

后续字段)不出现时出现。

### 30. packet\_data\_byte

由包层中的stream\_ID指定的比特流中 $n$ 字节数据,字节数由参数packet\_length决定。

## 12.7 视频解码

系统解复用器分析系统比特流,解复用视频音频比特流。

视频解码器本质上执行与编码器相反的工作过程。它从编码视频比特流中重构I帧。利用I帧和附加编码数据、运动矢量等构建P帧和B帧。最后,把各帧按照合适的顺序输出。

### 12.7.1 快进播放注意事项

快进播放操作可采用D帧或仅对I帧解码来实现。然而,在高速率下仅对I帧解码会给传输媒质和解码器带来很大的负担。

另一种方法是,发送源可以仅挑选快放所需的I帧传输,使得比特率仍然保持稳定。

### 12.7.2 暂停模式注意事项

解码器必须能够控制输入的比特流,否则,在播放恢复的时候会出现延迟或跳帧现象。

### 12.7.3 快退播放注意事项

解码器必须能够在视频播放方向解码每一个图片组,并存储这些解码的图片组,这样才能实现快退播放。为最小化解码器的存储空间,图片组必须比较小,或者可以对帧进行重排序。重排序可以采用另一个顺序传输帧来实现,也可以通过在解码器缓存里重新调整图片组的顺序来实现。

### 12.7.4 解码后期处理

SIF数据通常转换为 $720 \times 480i$ 或 $720 \times 576i$ 。MPEG-1规范中考虑过隐式上采样过滤。原始解码线与场1对应。场2采用内插线。575

## 12.8 现实问题

### 12.8.1 系统比特流终止

系统比特流中的sequence\_end\_code位置错误是一个常见的问题。如果这种情况发生了,一些解码器将不知道视频结束的位置,而输出乱七八糟的东西。

消除拖动帧使得系统比特流缩短,同时会一并去掉sequence\_end\_code关键字,此时会产生另外一个问题:解码器无法确定在什么时候停止播放。

### 12.8.2 时间码

由于一些解码器依赖时间码信息,所以必须有时间码措施。为让问题最小化,视频流开始时将时间码设为0,每一帧时间码加1。

### 12.8.3 可变比特率

虽然MPEG-1支持可变比特率,但如果可能还是尽量采用恒定比特率。由于vbv\_delay对可变

比特率没意义，对可变比特率情况，MPEG-1标准指定其设为最大值。

然而，某些解码器针对可变比特率也使用vbv\_delay参数，这样导致起始视频延迟2~3秒，头60~90帧数据都被跳过去了。

#### 12.8.4 约束比特流

多数MPEG-1解码器只能处理一部分MPEG-1约束参数。为保证效果最好，只允许采用那部分约束参数。

#### 12.8.5 源采样时钟

要得到高压缩、低失真的效果，需要视频源产生或使用非常稳定的采样时钟。采用稳定时钟保证整幅图像样本的垂直对齐。质量差的时钟使得图像的失真更加剧烈。

### 参考文献

1. Digital Video Magazine, "Not All MPEGs Are Created Equal", by John Toebe, Doug Walker, and Paul Kaiser, August 1995.
2. Digital Video Magazine, "Squeeze the Most From MPEG", by Mark Magel, August 1995.
3. ISO/IEC 11172-1, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 1: Systems*.
4. ISO/IEC 11172-2, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 2: Video*.
5. ISO/IEC 11172-3, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 3: Audio*.
6. ISO/IEC 11172-4, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 4: Compliance testing*.
7. ISO/IEC 11172-5, *Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, Part 5: Software simulation*.
8. Watkinson, John, *The Engineer's Guide to Compression*, Snell and Wilcox Handbook Series.

# 第13章 MPEG-2

MPEG-2扩展了MPEG-1的相关功能和特性，从而覆盖了更加广泛的应用范围。在阅读本章之前，请重新回顾一下第12章关于MPEG-1的有关内容，以熟悉MPEG的相关基础知识。

MPEG-2制定过程中所确定的主要目标应用是在4~9Mbit/s比特率下的全数字广播级视频传输。然而，MPEG-2对于许多其他应用也是非常有用的，如HDTV，并且MPEG-2现在可支持的比特率在1.5~60Mbit/s之间。

MPEG-2是一个ISO标准（ISO/IEC 13818），它由下列11部分组成：

系统	ISO/IEC 13818-1
视频	ISO/IEC 13818-2
音频	ISO/IEC 13818-3
一致性测试	ISO/IEC 13818-4
软件仿真	ISO/IEC 13818-5
DSM-CC扩展	ISO/IEC 13818-6
高级音频编码	ISO/IEC 13818-7
系统解码实时接口	ISO/IEC 13818-8
RTI扩展	ISO/IEC 13818-9
DSM-CC一致性	ISO/IEC 13818-10
IPMP	ISO/IEC 13818-11

与MPEG-1一样，MPEG-2中的压缩比特流隐式地定义了解压缩算法。而压缩算法则由各厂商独立开发，并允许在国际标准范围内获得所属权益。

数字存储媒体命令与控制（Digital Storage Media Command and Control, DSM-CC）扩展（ISO/IEC 13818-6）是一个用于开发与MPEG-2流相关的控制信道的工具集。除了提供VCR风格特征（如快进、快退、暂停等），DSM-CC还可用于各种其他目的的应用，如分组数据传输。DSM-CC将工作在下一代分组网络上，并与相关的因特网协议如RSVP、RSTP、RTP和SCP等一起工作。

实时接口（Real Time Interface, RTI）扩展（ISO/IEC 13818-9）定义了一个终端设备制造商和网络运营商都能设计的公共接口点。RTI在解码器的输入处为MPEG-2系统流的字节定义了一个传输模型，尽管MPEG-2系统定义了一个理想的字节传输时间表。

知识产权管理和保护（Intellectual Property Management and Protection, IPMP）是一个数字版权管理（Digital Right Management, DRM）标准，是根据MPEG-4的IPMP扩展规范改编过来的。IPMP并不是一个完整的系统，而是一个提供各种函数的框架。

## 13.1 音频概述

除了非向后兼容（non-backwards-compatible）音频扩展（ISO/IEC 13818-7）之外，MPEG-2还最多可支持5个与MPEG-1音频编码兼容的全带宽信道。MPEG-2同样也将MPEG-1的音频编码扩展到了半采样率（16kHz、22.05kHz和24kHz），从而改善在比特率小于等于每信道64kHz情况下的音频质量。

MPEG-2.5是一个对MPEG-2音频能力非官方的通用扩展，它增加了8kHz、11.025kHz和12kHz的采样率。

## 13.2 视频概述

在MPEG-2中，档次定义了语法（即算法），级别定义了各种参数（如分辨率、帧率和比特率等）。位于主级别的主要档次的目标应用是SDTV，而位于高级别的主要档次的目标应用为HDTV。

### 13.2.1 级别

MPEG-2支持4种级别，这4种级别分别为给定的档次定义了不同的分辨率、帧率、编码速率等。

#### 1. 低级别 (LL)

低级别 (Low Level) 中，MPEG-2约束参数比特流 (CPB) 支持的最大帧率为每秒30帧，最大分辨率为 $352 \times 288$ ，最大比特率为4Mbit/s。

#### 2. 主级别 (ML)

主级别 (Main Level) 中，MPEG-2约束参数比特流 (CPB) 支持的最大帧率为每秒30帧，最大分辨率为 $720 \times 288$ ，目标应用为SDTV。最大比特率为15~20Mbit/s。

#### 3. 高1440级别

高1440级别 (High 1440 Level) 支持的最大帧率为每秒60帧，最大分辨率为 $1440 \times 1080$ ，其目标应用为HDTV。最大比特率为60~80Mbit/s。

#### 4. 高级别 (HL)

高级别 (High Level) 支持的最大帧率为每秒60帧，最大分辨率为 $1920 \times 1088$ ，其目标应用为HDTV。最大比特率为80~100Mbit/s。

### 13.2.2 档次

MPEG-2支持6个档次，这6个档次定义了各自所采用的编码语法（算法）。表13-1~表13-8说明了MPEG-2所允许的档次与级别的各种不同组合。

表13-1 MPEG-2可接受的级别与档次的组合

级 别	档 次						
	非可分级的				可分级的		
	简 单	主 要	多 视 角	4 : 2 : 2	SNR	空 域	高
高	-	是	-	是	-	-	是
高1440	-	是	-	-	-	是	是
主	是	是	是	是	是	-	是
低	-	是	-	-	是	-	-

表13-2 部分MPEG-2档次约束

约 束	档 次						
	非可分级的				可分级的		
	简 单	主 要	多 视 角	4 : 2 : 2	SNR	空 域	高
色度格式	4 : 2 : 0	4 : 2 : 0	4 : 2 : 0	4 : 2 : 0或 者4 : 2 : 2	4 : 2 : 0	4 : 2 : 0	4 : 2 : 0或 者4 : 2 : 2

(续)

约 束	档 次						
	非可分级的				可分级的		
	简 单	主 要	多 视 角	4 : 2 : 2	SNR	空 域	高
图像类型	I, P	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B	I, P, B
可分级性模式	-	-	时域	-	SNR	SNR或者空域	SNR或者空域
intra dc精度(位)	8, 9, 10	8, 9, 10	8, 9, 10	8, 9, 10, 11	8, 9, 10	8, 9, 10	8, 9, 10, 11
序列可分级性扩展	否	否	是	否	是	是	是
图像空域可 分级性扩展	否	否	否	否	否	是	是
图像时域可分 级性扩展	否	否	是	否	否	否	否
重复的第一个字段	受约束的		不受约束的	受约束的	不受约束的		

表13-3 MPEG-2可分级性档次允许的层次编号

级 别	层次最大编号	档 次			
		SNR	空 域	高	多 视 角
高	所有层(基本层+增强层)	-	-	3	2
	空域增强层			1	0
	SNR增强层			1	0
	时域备用层			0	1
高1440	所有层(基本层+增强层)	-	3	3	2
	空域增强层		1	1	0
	SNR增强层		1	1	0
	时域备用层		0	0	1
主	所有层(基本层+增强层)	2	-	3	2
	空域增强层	0		1	0
	SNR增强层	1		1	0
	时域备用层	0		0	1
低	所有层(基本层+增强层)	2	-	-	2
	空域增强层	0			0
	SNR增强层	1			0
	时域备用层	0			1

表13-4 各个档次的若干MPEG-2视频解码器需求

档 次	档 次			基本解码器的各级别对应档次
	基 本 层	增 强 层 1	增 强 层 2	
SNR	4 : 2 : 0	SNR, 4 : 2 : 0	-	位于同级别的主要档次
	4 : 2 : 0	SNR, 4 : 2 : 0	-	位于同级别的主要档次
空域	4 : 2 : 0	空域, 4 : 2 : 0	-	位于下级别的主要档次
	4 : 2 : 0	SNR, 4 : 2 : 0	空域, 4 : 2 : 0	
	4 : 2 : 0	SNR, 4 : 2 : 0	SNR, 4 : 2 : 0	
高	4 : 2 : 0或4 : 2 : 2	-	-	位于同级别的主要档次
	4 : 2 : 0	SNR, 4 : 2 : 0	-	
	4 : 2 : 0或4 : 2 : 2	SNR, 4 : 2 : 2	-	

(续)

档次	档次			基本解码器的各级别对应档次
	基本层	增强层1	增强层2	
高	4:2:0	空域, 4:2:0	-	位于下一级别的高档次
	4:2:0或4:2:2	空域, 4:2:2	-	
	4:2:0	SNR, 4:2:0 或4:2:2	空域, 4:2:0 或4:2:2	
	4:2:0或4:2:2	SNR, 4:2:2	空域, 4:2:2	
	4:2:0	空域, 4:2:0	SNR, 4:2:0 或4:2:2	
	4:2:0	空域, 4:2:2	SNR, 4:2:2	
	4:2:2	空域, 4:2:2	SNR, 4:2:2	
多视角	4:2:0	时域, 4:2:0	-	位于同级别的主要档次

表13-5 MPEG-2分辨率上限和时域参数。“增强层”参数适用于单层或者SNR可分级性编码情况

级别	空域分辨率层	参数	档次					
			简单	主要	多视角	4:2:2	SNR/空域	高
高	增强层	每行像素数	-	1920	1920	1920	-	1920
		每帧行数		1088	1088	1088		1088
		每秒帧数		60	60	60		60
	低层	每行像素数	-	1920	-	-	-	960
		每帧行数		1088				576
		每秒帧数		60				30
高1440	增强层	每行像素数	-	1440	1440	-	1440	1440
		每帧行数		1088	1088		1088	1088
		每秒帧数		60	60		60	60
	低层	每行像素数	-	1440	-	-	720	720
		每帧行数		1088			576	576
		每秒帧数		60			30	30
主	增强层	每行像素数	720	720	720	720	720	720
		每帧行数	576	576	576	608	576	576
		每秒帧数	30	30	30	30	30	30
	低层	每行像素数	-	720	-	-	-	352
		每帧行数		576				288
		每秒帧数		30				30
低	增强层	每行像素数	-	352	352	-	352	-
		每帧行数		288	288		288	
		每秒帧数		30	30		30	
	低层	每行像素数	-	352	-	-	-	-
	每帧行数	288						
	每秒帧数	30						

注：上述级别和档次中原来指定为每帧最大行数为1152改为每帧1088行。

表13-6 MPEG-2主要档次的级别和分辨率示例

级 别	最大比特率 (Mbit/s)	典型有效 分辨率	帧率 (Hz) <sup>2</sup>										
			23.976p	24p	25p	29.97p	30p	50p	59.94p	60p	25i	29.97i	30i
高	80 (HP为100) (4:2:2档次 为300)	1920×1080 <sup>1</sup>	×	×	×	×	×				×	×	×
高1440	60 (HP为80)	1280×720	×	×	×	×	×	×	×	×			
		960×1080 <sup>1</sup>	×	×	×	×	×				×	×	×
		1280×1080 <sup>1</sup>	×	×	×	×	×				×	×	×
		1440×1080 <sup>1</sup>	×	×	×	×	×				×	×	×
主	15 (HP为20) (4:2:2档次 为50)	352×480	×	×		×	×		×	×		×	×
		352×576		×	×				×			×	
		480×480	×	×		×	×		×	×		×	×
		544×480	×	×		×	×		×	×		×	×
		544×576		×	×	×			×			×	
		640×480	×	×		×	×		×	×		×	×
		704×480, 720×480	×	×		×	×		×	×		×	×
		704×576, 720×576		×	×				×			×	
		320×240	×	×		×	×		×	×		×	×
		352×240	×	×		×	×		×	×		×	×
低	4	352×288		×	×				×			×	

注：1. 视频编码系统要求对于隔行图像有效扫描线数必须为32的倍数，对于逐行图像有效扫描线数必须是16的倍数。因此，对于1080线的隔行格式视频，视频编码器和解码器实际上必须采用1088行扫描线。多出的8线为没有实质内容的“虚拟”线，设计者之所以选择虚拟数据，是为了简化实现。多出的8线总是编码图像的最后8线。这些虚拟线并不携带有用信息，只是所需要传输的数据稍微有所增加。

2. p = 逐行；i = 隔行。

表13-7 MPEG-2 Y采样率上限 (M采样点/秒)。“增强层”参数适用于单层或者SNR可分级编码的情况

级 别	空间分辨率层	档 次						4 : 2 : 2
		简 单	主 要	多 视 角	SNR/空域	高	4 : 2 : 2	
高	增强层	—	62.668800	62.668800	—	62.668800 (4:2:2) 83.558400 (4:2:0)	—	62.668800
		—	—	62.668800	—	14.745600 (4:2:2) 19.660800 (4:2:0)	—	
高1440	增强层	—	47.001600	47.001600	47.001600	47.001600 (4:2:2) 62.668800 (4:2:0)	—	—
		—	—	—	—	—	—	

(续)

级 别	空间分辨率层	档 次					
		简 单	主 要	多 视 角	SNR/空域	高	4 : 2 : 2
高1440	低层	—	—	47.001600	10.368000	11.059200 (4 : 2 : 2) 14.745600 (4 : 2 : 0)	—
主	增强层	10.368000	10.368000	10.368000	10.368000	11.059200 (4 : 2 : 2) 14.745600 (4 : 2 : 0)	11.059200
	低层	—	—	10.368000	—	3.041280 (4 : 2 : 0)	—
低	增强层	—	3.041280	3.041280	3.041280	—	—
	低层	—	—	3.041280	—	—	—

表13-8 MPEG-2比特率上限 (Mbit/s)

级 别	档 次					
	非可分级的				可分级的	
	简 单	主 要	多 视 角	4 : 2 : 2	SNR/时域	高
高	—	80	130 (所有层) 80 (基本层)	300	—	100 (所有层) 80 (中间层+基本层) 25 (基本层)
高1440	—	60	100 (所有层) 60 (基本层)	—	60 (所有层) 40 (中间层+基本层) 15 (基本层)	80 (所有层) 60 (中间层+基本层) 20 (基本层)
主	15	15	25 (所有层) 5 (基本层)	50	15 (所有层) 10 (基本层)	20 (所有层) 15 (中间层+基本层) 4 (基本层)
低	—	4	8 (所有层) 4 (基本层)	—	4 (所有层) 3 (基本层)	—

### 1. 简单档次

无B帧的主要档次即为简单档次 (Simple Profile, SP)，目标应用为软件应用和可能的有线电视。

### 2. 主要档次

大多数MPEG-2解码器芯片都支持主要档次 (Main Profile, MP)，主要档次应该满足90%的消费类SDTV和HDTV应用。表13-6列出了主要档次支持的典型分辨率。

### 3. 多视角档次

通过采用已有的MPEG-2工具，多视角档次 (Multiview Profile, MVP) 可能能够解码同一场景、角度有稍微差别的两个摄像机拍摄的视频。

### 4. 4 : 2 : 2档次

4 : 2 : 2档次 (4 : 2 : 2 Profile, 422P) 即为人们熟知的“演播室档次 (studio profile)”，422P采用4 : 2 : 2的YCbCr替代4 : 2 : 0的YCbCr，此外，相对主要档次，422P增加了最大比特率到

50Mbit/s（对于高档次，这个值为300Mbit/s），从而支持专业级视频SDTV和HDTV需求。

### 5. SNR和空域档次

SNR和空域档次（SNR and Spatial Profile）增加了对SNR可分级性和空域可分级性的支持。

### 6. 高档次

高档次（High Profile, HP）目标应用为专业级视频HDTV应用。

578

## 13.2.3 可分级性

MPEG-2 SNR、空域和高档次支持四种可分级的操作模式。这些模式将MPEG-2视频分成不同的层次，从而能够区分视频数据的优先次序。采用分级性将导致效率下降近2dB（或者需要多出约30%比特），因此可分级性并不被普遍采用。

### 1. SNR可分级性

SNR可分级模式的目标应用是那些需要提供多个质量层次的应用。在该模式中，所有层都有相同的空间分辨率。基本层提供基本的视频质量，增强层通过为基本层DCT系数提供精细数据而提升视频质量。

### 2. 空域可分级性

空域可分级性对于同播（simulcasting）是非常有用的，在该模式中每一个层都有一个不同的空间分辨率。基本层提供基本的空间分辨率和时域速率。增强层采用对基本层进行空间插值来增加空域分辨率。例如，基本层可能实现 $352 \times 240$ 分辨率视频，利用增强层可生成 $704 \times 480$ 分辨率的视频。

### 3. 时域可分级性

时域可分级模式使得可以从低时域速率迁移到更高的时域速率系统。基本层提供基本的时域速率。增强层采用相对于基本层的时域预测。基本层和增强层可以组合在一起从而产生一个全时域速率输出。所有层都有相同的空域分辨率和色度格式。在增强层出现差错的情况下，基本层可以用来进行差错掩蔽（concealment）。

### 4. 数据分区

数据分区模式的目标是在ATM网络中进行信元丢失恢复（resilience）。在该模式中，将64个量化的变换系数分成两个比特流。高优先级比特流包含关键的低频DCT系数和伴随信息，如头和运动矢量等。低优先级比特流携带可以增加细节的高频DCT系数。

## 13.2.4 传输和节目流

MPEG-2系统标准定义了两种方法来复用音频、视频和其他数据到一个适合于传输和存储的格式。

节目流是为很少出现差错的应用而设计的。它包含了音频、视频和数据比特流（也称作基本比特流），并将这些比特流合并到一个单独的比特流中。节目流，也就是每一个基本比特流，可能是固定或者可变比特率。DVD和SVCD采用节目流，在与视频和音频流交织的私有数据流中携带DVD和SVCD相关的数据。

传输流是为较容易出现数据丢失的应用而设计的，它采用长度为188字节的固定大小包。传输流同样包含了音频、视频和数据比特流，并且将这些比特流都合并到一个单独的比特流，传输流可以携带多个节目。ARIB、ATSC、DVB和OpenCable™标准都采用传输流。

传输流和节目流都是基于通用的包结构，便于通用的解码器实现和转换。这两种流都是设计支持大量已知和未来可以预见的应用，且保持灵活性。

579

584

### 13.3 视频编码层

#### 13.3.1 YCbCr颜色空间

MPEG-2采用YCbCr颜色空间，支持4：2：0、4：2：2和4：4：4采样。4：2：2和4：4：4采样选项在4：2：0采样上提高了色度分辨率，从而获得更好的图像质量。

图3-8~图3-10给出了MPEG-2的4：2：0采样结构，图3-2和图3-3分别给出了MPEG-2的4：2：2和4：4：4采样结构。

#### 13.3.2 编码图像类型

MPEG-2中有三种类型的编码图像。I (intra) 图像是作为独立静态图像编码的场或帧。I图像在视频流中提供可以随机存取的随机存取点。通常，在1秒钟的视频中约出现2次I图像。此外，在场景发生突变的地方也应该采用I图像。

P (predicted) 图像是由与其最近邻的前一个I图像或P图像预测而得到的场或帧（即所谓的前向预测处理），如图13-1所示。通过采用运动补偿技术，P图像提供比I图像更好的压缩效率，而且也能够作为B图像和接下来的P图像的参考图像。

B (bi-directional) 图像是采用最邻近的前后两个I图像或P图像作为参考，进行双向预测得到的场或帧，如图13-1所示。B图像提供最高的压缩效率，并且通过利用前后两个图像的均值减少了噪声。通常，I图像或P图像之间有2个B图像。

MPEG-2中不支持D (DC) 图像，除了在解码过程中支持与MPEG-1的向后兼容性之外。

GOP是用于辅助随机存取和编辑的一个或多个编码图像序列。在编码过程中GOP值是可以配置的，GOP值越小，对运动的响应效果越好（因为与I图像靠得越近），但压缩效率却下降了。

在编码流中，一个图像组必须以I图像作为开始，随后可以按任意顺序连接任意个I图像、P图像或B图像。在显示顺序中，图像组必须以I图像或B图像开始，并以I图像或P图像结束。因此，最小的图像组是单个I图像，而最大图像组的大小是无限的。

每个图像组应该独立于其他任何图像组进行编码。然而，这种情况只有在第一个I图像前没有B图像才有可能，否则，只能进行逆向运动补偿。于是形成了开放和封闭两种图像组格式。封闭图像组不需要之前的图像组做运动补偿就能完成解码。而采用broken\_link标志标识的开放图像组[意味着在GOP头之后的第一个B图像（如果有）是紧跟在第一个I图像之后]可能无法正确地解码（因此也就不会显示），因为用于预测的参考图像由于编辑原因而不能获得。

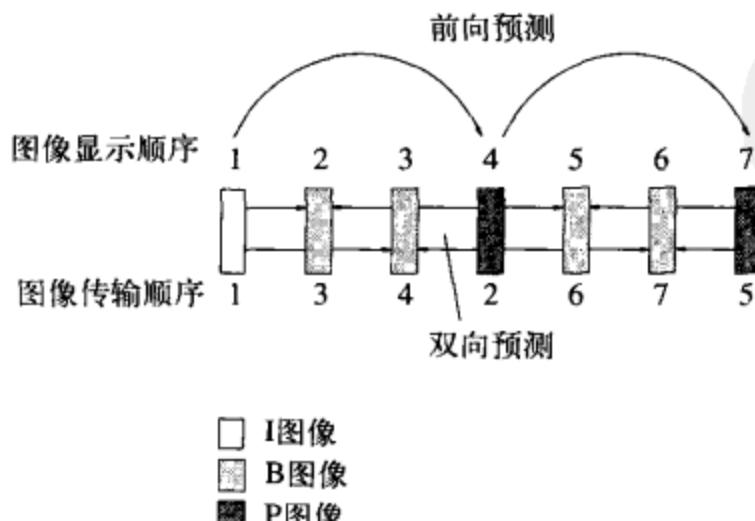


图13-1 MPEG-2中的I、P和B图像。某些图像可能不按顺序传输，从而使插值过程变得复杂，因此需要MPEG解码器对图像进行重新排序。箭头表示帧之间的依赖关系

### 13.3.3 运动补偿

由于引入了场 (field), MPEG-2的运动补偿变得更为复杂。采用运动补偿对一个宏块进行解压之后, 它既包含了参考宏块和解码宏块之间的空域差值 (运动矢量), 也包含了内容差值 (误差项)。

两个主要的预测类别分别是场和帧。在场图像中, 仅采用场预测。在帧图像中, 既可以采用场预测, 也可以采用帧预测 (在宏块层决定具体采用何种预测)。

MPEG-2的运动矢量总是采用半像素精度进行编码, 而MPEG-1既支持半像素也支持全像素精度。

#### 1. $16 \times 8$ 运动补偿选项

在该选项中, 采用每个宏块两个运动矢量 (对于B图像则为4个), 其中一个用于宏块上部的 $16 \times 8$ 区域, 另一个用于宏块下部的 $16 \times 8$ 区域。该模式仅用于场图像。

#### 2. 对偶运动补偿选项

该模式仅用于在预测和参考场或帧之间没有B图像的P图像。采用一个运动矢量, 同时还有一个小的差分运动矢量。所有必要的预测都是由这个矢量推导出来的。586

### 13.3.4 宏块

MPEG-2中有三种类型的宏块。

$4:2:0$ 宏块 (图13-2) 由4个Y块、1个Cb块和1个Cr块组成。图中给出了块的排列次序。

$4:2:2$ 宏块 (图13-3) 由4个Y块、2个Cb块和2个Cr块组成。图中给出了块排列次序。

$4:4:4$ 宏块 (图13-4) 由4个Y块、4个Cb块和4个Cr块组成。图中给出了块的排列次序。

P图像中的宏块采用最近的前一个I或P图像作为参考编码而得到, 因此导致下面两种可能的编码模式:

内部编码 (intra-coding), 没有运动补偿;

前向预测 (forward prediction), 最近的前一个I或P图像作为参考。

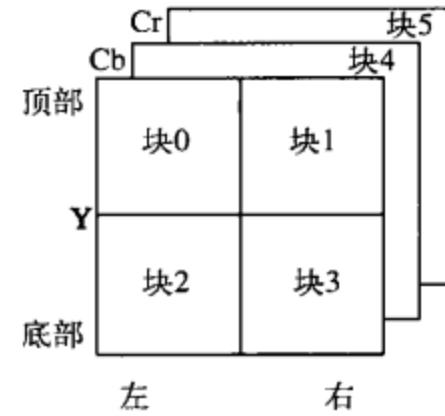


图13-2 MPEG-2  $4:2:0$ 宏块结构

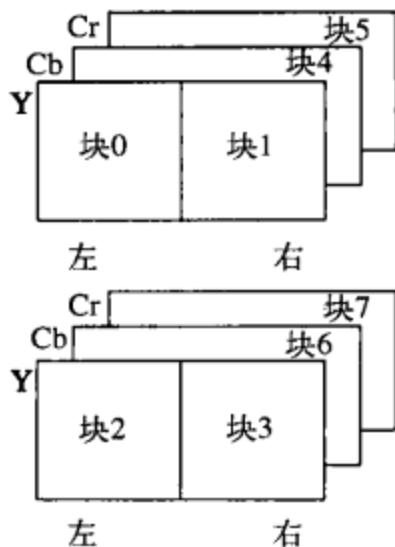


图13-3 MPEG-2  $4:2:2$ 宏块结构

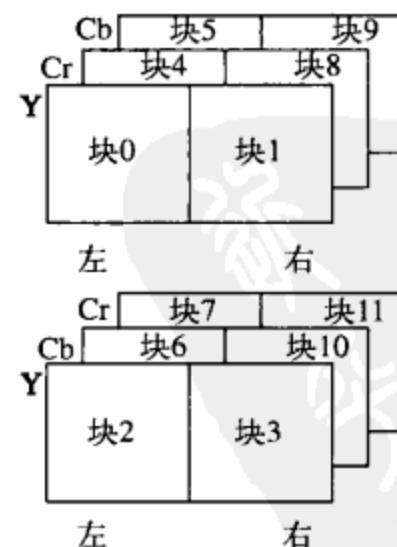


图13-4 MPEG-2  $4:4:4$ 宏块结构

B图像中的宏块采用最近的前一个和后一个I或P图像作为参考而编码得到, 因此导致下面四

种可能的编码模式：

- 内部编码 (intra-coding)，没有运动补偿；
- 前向预测 (forward prediction)，最近的前一个I或P图像作为参考；
- 后向预测 (backward prediction)，最近的后一个I或P图像作为参考；
- 双向预测 (bi-directional prediction)，采用两个图像作为参考——最近的前一个I或P图像和最近的后一个I或P图像。

### 13.3.5 I图像

#### 1. 宏块

I图像中有10种类型的宏块，如表13-27所示。

如果表13-27中的[macroblock quant]列值为1，量化器刻度将会传送。对于其他宏块类型，DCT修正将采用前一个值来编码，编码结果作为该宏块的量化器刻度。

如果表13-27中的[coded pattern]列值为1，长度为6位的编码块模式将被作为可变长编码传送。编码块模式将指示解码器在4：2：0宏块的6个块中哪一个块已经编码（标示为1），哪一个没有编码（标示为0）。表13-32列出了分配给63种可能组合的码字。当所有6个块都没有编码时，将没有码字，这由宏块类型来指示。对于4：2：2和4：4：4宏块，分别有附加的2位或者6位用来扩展编码块模式。

#### 2. DCT

每一个 $8 \times 8$ 块（输入采样点或者预测误差项）由一个 $8 \times 8$  DCT处理，得到一个 $8 \times 8$ 的垂直和水平频率系数的块。

输入采样点值为0~255，DCT变换之后，获得的DC系数范围为0~2040，AC系数为-2040~2047。

由于空域和SNR可分级性，非内部块（在非内部宏块内的块）的出现同样是可能的。非内部块系数表示的是采样值之间的差值，而不是实际采样值。它们是通过从当前宏块的运动补偿值中抽出前一个图像中的运动补偿值而获得的。  
587

#### 3. 量化

频率系数 $8 \times 8$ 块被均匀地量化，从而限制了量化值的数目。量化器步长刻度由量化矩阵和量化器刻度推导而得到，并且对于不同的系数可能有不同的值，在宏块之间也可能发生变化。

由于人的眼睛对于大亮度区域较为敏感，因此，DC系数的量化器步长刻度可选择8、9、10或11位的精度。量化的DC系数由DC系数除以8、4、4或1之后取整到最近的整数而确定。

AC系数采用两个量化矩阵量化：其中一个用于内部宏块，另一个用于非内部宏块。当采用4：2：2或者4：4：4数据时，对于Y和CbCr数据可能采用不同的矩阵。每一个量化矩阵都有一个默认的值集，该默认的值集可以被覆盖。

如果表13-27中的[macroblock quant]列值为1，量化器刻度将被传送。对于其他宏块类型，DCT修正将采用前一个值来编码，编码结果作为该宏块的量化器刻度。

#### 4. Z形扫描

如图7-57和图7-58所示，Z形扫描从DC分量开始，按照频率递增的顺序将量化频率系数排列成线性流。这将产生长游程的非零系数。

#### 5. 量化DC系数编码

DC系数量化后，将采用无失真编码方式对其进行编码。

宏块中的Y块编码按照图13-2所示的顺序进行。块4的DC值将作为下一个宏块中块1的DC值的预测值。在每个条带的开头，当一个宏块被跳过或者当一个非内部宏块被编码时，DC预测值

被置为128 (DC精度为8位)、256 (DC精度为9位)、512 (DC精度为10位) 或者1024 (DC精度为11位)。

每个Cb、Cr块的值采用前一个宏块的对应块的DC值作为预测值进行编码，在每个条带开头，当一个宏块被跳过或者当一个非内部宏块被编码时，DC预测值被置为128 (DC精度为8位)、256 (DC精度为9位)、512 (DC精度为10位) 或者1024 (DC精度为11位)。

然而，一个普遍的实现是将DC预测值设置为0，并且将内部块DC项中心值设置为0附近，而取代50%灰度值。因此，解码器仅需要处理量化器（有一个乘法器可以用来重建正确值），而不是剖析器（通常不接触数据，且没有乘法器）内的不同内部DC精度。

### 6. 量化AC系数编码

AC系数量化后，将按照图7-57或图7-58所示的顺序扫描，采用游程 (run-length) 和非零值 (level) 编码。如图7-57和图7-58所示，扫描从位置1开始，因为在位置0的DC系数单独编码。

游程和非零值采用如表13-36和表13-37所示方式编码。s位表示该非零值的符号，0为正，1为负。对于内部块，既可以采用表13-36，也可以采用表13-37，具体采用哪个表将由比特流中的 `intra_vlc_format` 来决定。对于非内部块，只能采用表13-36。

对于表13-36和表13-37中没有列出的游程-非零值组合的情况，引入了一个转义序列，由一个转义码 (ESC) 后面接着如表13-38和表13-39所示的游程和非零值编码。

当最后一个DCT系数编码完成后，添上一个EOB符号，以指示解码器在这个 $8 \times 8$ 块中已经没有其他量化系数了。

## 13.3.6 P图像

### 1. 宏块

如表13-28所示，由于运动补偿附加的复杂性，P图像内有一共26种类型的宏块。

如果比特流中的 `macroblock_address_increment` 参数大于1，不编码宏块 (skipped macroblock) 将会出现。对于P场图像，解码器将从与被预测场同奇偶性的场进行预测，运动矢量预测值被设置为0。对于P帧图像，解码器设置运动矢量预测值和运动矢量都为0。

如果表13-28中的 [macroblock quant] 项为1，需要传送量化器刻度。对于其他宏块类型，采用先前的量化器刻度对DCT修正进行编码。

如果表13-28中的 [motion forward] 列值为1，水平和垂直前向运动矢量应当相继传输。

如果表13-28中的 [coded pattern] 列值为1，6位编码块模式 (coded block pattern) 将作为变长编码传输。它会指示解码器当前宏块的6个块中哪些已经编码 (设为1)，哪些没有编码 (设为0)。表13-32列出63种可能组合对应的码字。如果没有块被编码，就没有码字。编码块模式由宏块类型决定。对于P帧、B帧中的帧内编码宏块，编码块模式将不需要传输，且假定为63 (表示所有块都已被编码)。对于4:2:2和4:4:4宏块，分别有附加的2位或者6位用来扩展编码块模式。

### 2. DCT

内部块AC系数采用与I图像中的内部块AC系数相同的方式进行变换。内部块DC系数变换方式则不同，预测值设置为1024，除非前一个块是采用帧内编码方式进行编码的。

非内部块系数表示的是采样值之间的差值，而不是实际采样值，它们是通过从当前宏块的运动补偿值中抽取出前一个图像中的运动补偿值而获得的。DC值是没有预测的。

输入采样点值范围为 -255~+255，DCT变换之后，获得的AC系数范围为 -2000~+2000。

### 3. 量化

P图像中的内部块采用与I图像中内部块相同的方式进行量化。

非内部块采用量化器刻度和非内部量化矩阵进行量化。AC和DC系数采用相同的方式进行量化。

#### 4. 内部块编码

P图像中的内部块采用与I图像中内部块相同的方式进行编码。在处理DC系数时稍微有点不同，即在P图像中预测值设为128，除非前一个块是内部编码块。

#### 5. 非内部块编码

590 编码块模式（CBP）用来指定哪一个块有系数数据。除了在非内部块中DC系数与AC系数采用系统的编码方式之外，非内部块编码方式与内部块编码方式类似。

### 13.3.7 B图像

#### 1. 宏块

根据后向运动补偿附加的复杂程度不同，B图像中有34种宏块类型，如表13-29所示。

对于B场图像，解码器将从与被预测场同奇偶性的场进行预测。预测方向（前向、后向或者双向）与前一个宏块的预测方向相同，运动矢量预测值不受影响，且运动矢量将从适当的运动矢量预测值中得到。对于B帧图像，预测方向（前向、后向或者双向）与前一个宏块的预测方向相同，运动矢量预测值不受影响，且运动矢量将从适当的运动矢量预测值中得到。

如果表13-29中的[macroblock quant]列值为1，将传输量化器刻度。而对于剩下的宏块类型，则采用前一个量化器刻度对DCT修正进行编码。

如果表13-29中的[motion forward]列值为1，将要相继传输水平和垂直前向运动矢量。如果表13-29中的[motion backward]列值为1，将要相继传输水平和垂直后向运动矢量。如果前向后向运动类型都出现了，运动矢量按照以下顺序传输：水平前向→垂直前向→水平后向→垂直后向。

如果表13-29中的[coded pattern]列值为1，6位长的编码块模式将采用变长编码方式传输，并指示解码器当前宏块的6个块中哪一块已经被编码（设为1），哪一块还没有被编码（设置为0）。表13-32列出了63种可能组合对应的码字。如果还没有块被编码，就没有码字。编码块模式由宏块类型决定。对于P帧、B帧中的帧内编码宏块，编码块模式不需要传输，且假定为63（表示所有块都已经被编码）。对于4:2:2和4:4:4宏块，分别有附加的2位或者6位用来扩展编码块模式。

#### 2. 编码

块的DCT系数被转换成量化的系数，并且采用与P图像中相同的方式进行编码。

## 13.4 视频比特流

图13-5对视频比特流进行了举例说明，从上到下总共7个层次，各层分别是：视频序列、序列头、图像组（GOP）、图像、条带、宏块（MB）、块。

591 以下扩展可以用来支持不同级别的能力：序列扩展、序列显示扩展、序列可分级扩展、图像编码扩展、量化矩阵扩展、图像显示扩展、图像时域可分级扩展、图像空域可分级扩展。

如果视频序列的第一个序列头不是跟在扩展起始码（0x00001B5）之后，视频比特流必须符合MPEG-1视频比特流规范。

对于MPEG-2视频比特流，扩展起始码（0x000001B5）和序列扩展之后必须跟每一个序列头。

注意，起始码（0x000001xx）必须是字节对齐的，其实现方式是在起始码之前插入0~7个0位。

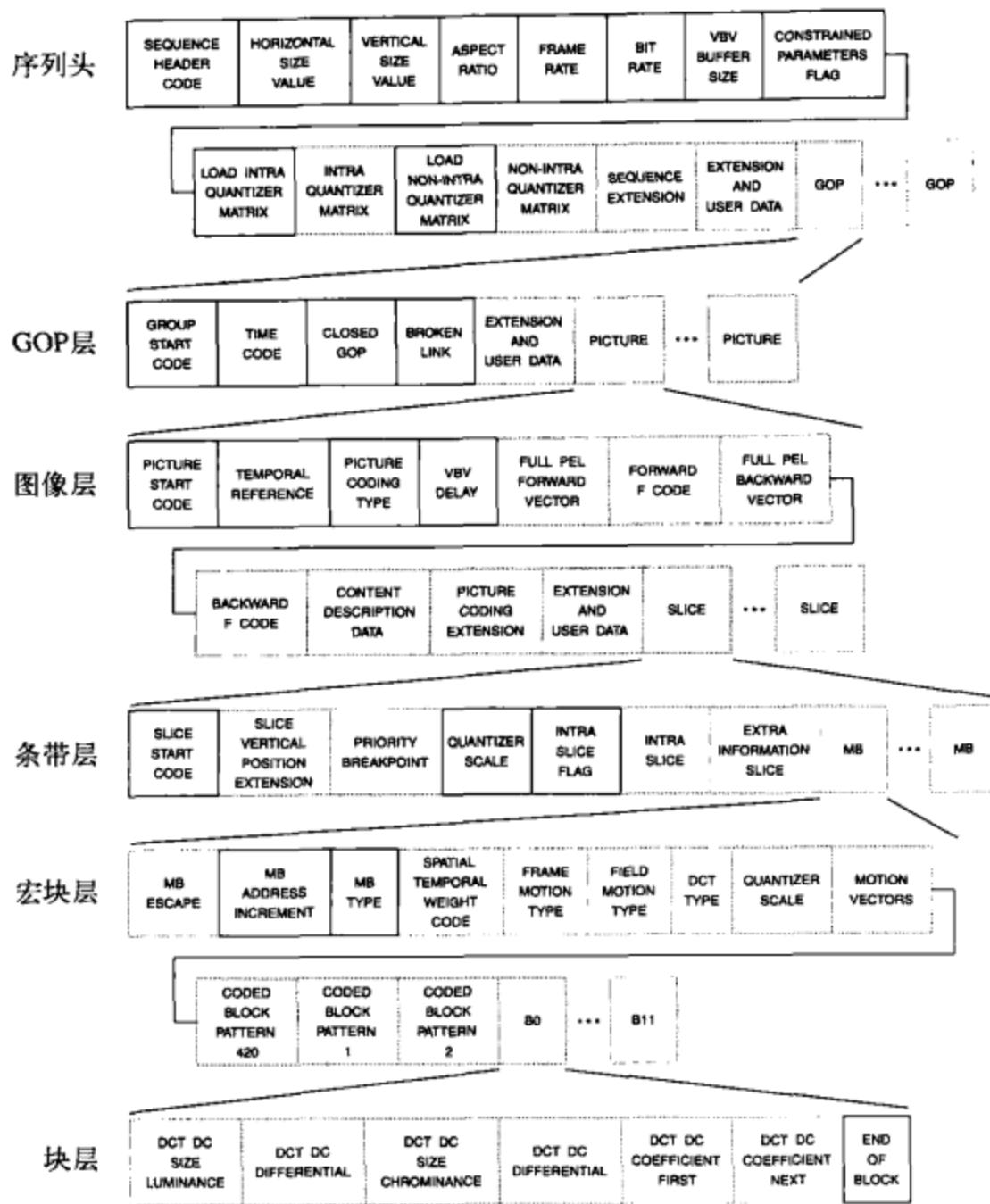


图13-5 MPEG-2视频比特流层次结构。没有给出标记和保留位

### 13.4.1 视频序列

`sequence_end_code`

长度为32位的字符串，值为0x000001B7，标志一个视频序列的结束。

### 13.4.2 序列头

一般来说，平均每秒钟应出现两次序列头。序列头结构如图13-5所示。如果其后没有跟一个序列扩展，那么比特流符合MPEG-1规范。

1. `sequence_header_code`

长度为32位的字符串，值为0x000001B3，表示序列头的开始。

2. `horizontal_size_value`

这是一个12位最低有效位，表示Y分量可见部分的宽度（像素）。14位值的两个最高有效位在`horizontal_size_extension`中指定了。零值是不允许的。

3. `vertical_size_value`

这是一个12位最低有效位，表示Y分量可见部分的高度（扫描线）。14位值的两个最高有效位

在vertical\_size\_extension中指定了。零值是不允许的。

#### 4. aspect\_ratio\_information

这是一个4位的码字，指示如表13-9所示的像素宽高比（SAR）或者显示宽高比（DAR）。

如果sequence\_display\_extension没有出现，SAR由下面的公式确定：

$$\text{SAR} = \text{DAR} \times (\text{horizontal\_size}/\text{vertical\_size})$$

如果sequence\_display\_extension出现了，SAR由下面的公式确定：

$$\text{SAR} = \text{DAR} \times (\text{display\_horizontal\_size}/\text{display\_vertical\_size})$$

#### 5. frame\_rate\_code

这是一个4位的码字，指示如表13-10所示的帧率。

实际帧率由下面的公式确定：

$$\text{frame\_rate} = \text{frame\_rate\_value} \times (\text{frame\_rate\_extension\_n} + 1) / (\text{frame\_rate\_extension\_d} + 1)$$

592  
593当指定表13-10中的某一个条目时，frame\_rate\_extension\_n和frame\_rate\_extension\_d均为00。如果progressive\_sequence为1，两帧到达解码器输出点之间的时间为帧率（frame\_rate）的倒数。如果progressive\_sequence为0，两帧到达解码器输出点之间的时间为帧率（frame\_rate）的倒数的一半。

表13-9 MPEG-2 aspect\_ratio\_information码字

SAR	DAR	编 码
禁止	禁止	0000
1.0000	—	0001
—	3/4	0010
—	9/16	0011
—	1/2.21	0100
—		0101
—		0110
—		0111
—		1000
—		1001
—		1010
—		1011
—		1100
—		1101
—		1110
—		1111

表13-10 MPEG-2 frame\_rate\_code码字

帧/秒	编 码
禁止	0000
24/1.001	0001
24	0010
25	0011
30/1.001	0100
30	0101
50	0110
60/1.001	0111
60	1000
保留	1001
保留	1010
保留	1011
保留	1100
保留	1101
保留	1110
保留	1111

#### 6. bit\_rate\_value

30位二进制数中的18位最低有效位，12位最高有效位在bit\_rate\_extension中。bit\_rate\_value表示比特流的比特率，以400bit/s为单位度量，不足时取上界。该值不允许为0。对于ATSC标准，值必须小于等于48500<sub>D</sub>（对于高数据率模式为小于等于97000<sub>D</sub>）。对于OpenCable™标准中的64QAM系统，值必须小于等于67500<sub>D</sub>（对于256QAM系统为小于等于97000<sub>D</sub>）。

#### 7. marker\_bit

标记位，值恒为1。

#### 8. vbu\_buffer\_size\_value

18位二进制数中的10位最低有效位。8位最高有效位在vbu\_buffer\_extension中。vbu\_buffer\_size

定义了解码器对视频序列解码时，视频缓存校验器（Video Buffering Verifier）需要的最小存储容量。按照如下定义：

$$B = 16 \times 1024 \times \text{vbv\_buffer\_size}$$

对于ATSC和OpenCable™标准，该值必须小于等于488<sub>D</sub>。

#### 9. constrained\_parameters\_flag

此位设为0，因为它对于MPEG-2没有任何意义。

#### 10. load\_intra\_quantizer\_matrix

如果其后紧跟intra\_quantizer\_matrix，此位设为1。如果设为0，如下默认值将用于内部块（Y和CbCr），直到下一个序列头或者quant\_matrix\_extension出现：

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

#### 11. intra\_quantizer\_matrix

可选的64个8位值列表，用来替换当前的帧内量化器值。不允许为0。intra\_quant[0,0]通常为8。在下一个序列头或者quant\_matrix\_extension出现之前，这些数值一直有效。对于4：2：2和4：4：4数据格式，新的值将用于Y和CbCr内部矩阵，除非装载了一个不同的CbCr内部矩阵。

#### 12. load\_non\_intra\_quantizer\_matrix

如果其后紧跟non\_intra\_quantizer\_matrix，此位设为1。如果设为0，在下一个序列头或者quant\_matrix\_extension出现之前，如下默认值将用于内部块（Y和CbCr）：

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

#### 13. non\_intra\_quantizer\_matrix

可选的64个8位值列表，用来替换当前的非帧内量化器值。不允许为0。在下一个序列头或者quant\_matrix\_extension出现之前，这些数值一直有效。对于4：2：2和4：4：4数据格式，新的值将用于Y和CbCr非内部矩阵，除非装载了一个不同的CbCr非内部矩阵。

594  
595

### 13.4.3 用户数据

#### 1. user\_data\_start\_code

可选的32位字符串，值为0x000001B2，表示user\_data（用户数据）的开始。在探测到下一个起始码之前，user\_data一直持续。

## 2. user\_data

长度为有 $n \times 8$ 位，当user\_data\_start\_code出现时才出现。user\_data不能包含超过连续23个0的字符串。

### 13.4.4 序列扩展

序列扩展只可能在序列头之后出现，其结构如图13-6所示。

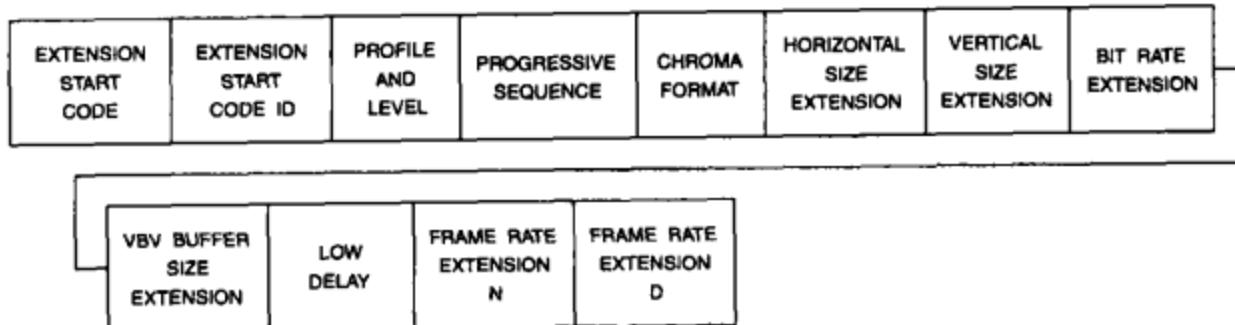


图13-6 MPEG-2序列扩展结构。没有给出标记位

#### 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示除MPEG-1之外的扩展数据的开始。

#### 2. extension\_start\_code\_ID

长度为4位的字段，值为0001，表示序列扩展的开始。对于MPEG-2视频比特流，每一个序列头都必须跟一个序列扩展。

#### 3. profile\_and\_level\_indication

长度为8位的字段，用于定义档次和级别，如表13-11所示。

位7表示转义位，位6~4表示档次ID，

位3~0表示级别ID。

表13-11 MPEG-2 profile\_and\_level\_indication码字

档次	档次ID编码	级别	级别ID编码
保留	000	保留	0000
高	001	保留	0001
空域可分级	010	保留	0010
SNR可分级	011	保留	0011
主要	100	高	0100
简单	101	保留	0101
保留	110	高1440	0110
保留	111	保留	0111
		主要	1000
		保留	1001
		低	1010
		保留	1011
		保留	1100
		保留	1101
		保留	1110
		保留	1111

**4. progressive\_sequence**

1位标志位，值为1时指示仅逐行图像出现，值为0时表示帧和场图像都可能出现，且帧图像可能是逐行的，也可能是隔行的。对于SVCD标准，该值必须为0。

**5. chroma\_format**

2位的码字，表示CbCr格式，如表13-12所示。对于ASTC和OpenCable™标准，该值必须为01。

**6. horizontal\_size\_extension**

horizontal\_size的两位最高有效位。对于ASTC和OpenCable™标准，该值必须为00。

**7. vertical\_size\_extension**

vertical\_size的两位最高有效位。对于ASTC和OpenCable™标准，该值必须为00。

**8. bit\_rate\_extension**

bit\_rate的12位最高有效位。对于ASTC和OpenCable™标准，该值必须为0000 0000 0000。

**9. marker\_bit**

标记位，值恒为1。

**10. vbv\_buffer\_size\_extension**

vbv\_buffer\_size的8位最高有效位。对于ASTC和OpenCable™标准，该值必须为0000 0000。

**11. low\_delay**

1位标志位，该位为1时表示没有B图像出现，因此没有帧重排序时延。对于SVCD，该值必须为0。

596

**12. frame\_rate\_extension\_n**

参看frame\_rate\_code关于这2位的二进制值。对于ASTC和OpenCable™标准，该值必须为00。

**13. frame\_rate\_extension\_d**

参看frame\_rate\_code关于这5位的二进制值。对于ASTC和OpenCable™标准，该值必须为00000。

### 13.4.5 序列显示扩展

这个可选的扩展仅可能在序列扩展之后出现，其结构如图13-7所示。

EXTENSION START CODE	EXTENSION START CODE ID	VIDEO FORMAT	COLOR DESCRIPTION	COLOR PRIMARIES	TRANSFER CHARACTERISTICS	MATRIX COEFFICIENTS	DISPLAY HORIZONTAL SIZE	DISPLAY VERTICAL SIZE
----------------------------	-------------------------------	-----------------	----------------------	--------------------	-----------------------------	------------------------	-------------------------------	-----------------------------

图13-7 MPEG-2序列显示扩展结构。没有显示出标记位

**1. extension\_start\_code**

长度为32位的字符串，值为0x000001B5，表示新的扩展数据集的开始。

**2. extension\_start\_code\_ID**

长度为4位的字段，值为0010，表示序列显示扩展的开始。由该扩展提供的信息并不影响解码过程，且可以被忽略。该扩展使得能够尽可能正确地显示解码的图像。

**3. video\_format**

长度为3位的码字，指示编码之前的图像源格式，如表13-13所示。对于ASTC和OpenCable™标准，该值必须为000。

表13-12 MPEG-2 chroma\_format码字

色度格式	编 码
保留	00
4:2:0	01
4:2:2	10
4:4:4	11

表13-13 MPEG-2 video\_format码字

视频格式	编 码
分量	000
PAL	001
NTSC	010
SECAM	011
MAC	100
未指定	101
保留	110
保留	111

#### 4. color\_description

1位的标志位，值为1时表示在比特流中将出现color\_primaries、transfer\_characteristics和matrix\_coefficients。

#### 5. color\_primaries

可选的8位码字，描述了源基色的色度坐标，如表13-14所示。如果sequence\_display\_extension没有出现，或者color\_description=0，表示必须采用默认值。

该信息可以在MPEG-2解码之后用来调整颜色处理过程，从而补偿显示的基色。

#### 6. transfer\_characteristics

可选的8位码字，描述了源图像的光电传输特性，如表13-15所示。如果sequence\_display\_extension没有出现，或者color\_description=0，表示必须采用默认值。

该信息可以在MPEG-2解码之后用来调整颜色处理过程，从而补偿显示的光电转换特性(gamma)。

表13-14 MPEG-2 color\_primaries码字

Color Primaries	编 码	默认应用
禁止	0000 0000	
BT.709、SMTPE 274M、 BT.1361、IEC 61966-2-4	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
BT.470系统M	0000 0100	30Hz SDTV
BT.470系统B、G	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
保留	0000 1000	
:	:	:
保留	1111 1111	

表13-15 MPEG-2\_transfer\_characteristic码字

光电传输特性	编 码	默认应用
禁止	0000 0000	
BT.709、SMPTE 274M、 BT.1361	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
BT.470系统M	0000 0100	30Hz SDTV
BT.470系统B、G	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
线性	0000 1000	
对数(100:1范围)	0000 1001	
对数(316:1范围)	0000 1010	
IEC 61966-2-4	0000 1011	
BT.1361	0000 1100	
保留	0000 1101	
:	:	:
保留	1111 1111	

#### 7. matrix\_coefficients

可选的8位码字，描述了用于从R'G'B'推导YCbCr的系数，如表13-16所示。如果

sequence\_display\_extension没有出现，或者color\_description = 0，表示必须采用默认值。

根据情况需要，该信息可以在MPEG-2解码之后用来选择合适的YCbCr到RGB转换的矩阵。

表13-16 MPEG-2 matrix\_coefficients码字

矩阵系数	编 码	默 认 应 用
禁止	0000 0000	
BT.709、SMPTE 274M、 BT.1361、 IEC 61966-2-4 (xvYCC <sub>709</sub> )	0000 0001	HDTV
未指定	0000 0010	
保留	0000 0011	
FCC	0000 0100	30Hz SDTV
BT.470系统B、G、I、 IEC 61966-2-4 (xvYCC <sub>601</sub> )	0000 0101	25Hz SDTV
SMPTE 170M	0000 0110	30Hz SDTV
SMPTE 240M	0000 0111	
YCgCo	0000 1000	
保留	0000 1001	
:	:	:
保留	1111 1111	

#### 8. display\_horizontal\_size

参看display\_vertical\_size，查看关于这个14位的二进制数的相关情况。

#### 9. marker\_bit

标记位，值恒为1。

#### 10. display\_vertical\_size

14位的二进制数，与display\_horizontal\_size一起定义了显示的有效区域。如果显示区域小于解码的图像大小，只能显示图像的一部分。如果显示区域大于解码图像大小，图像将显示在显示区域的一部分上。

### 13.4.6 序列可分级扩展

这个可选的扩展仅可能在序列扩展之后出现，其结构如图13-8所示。

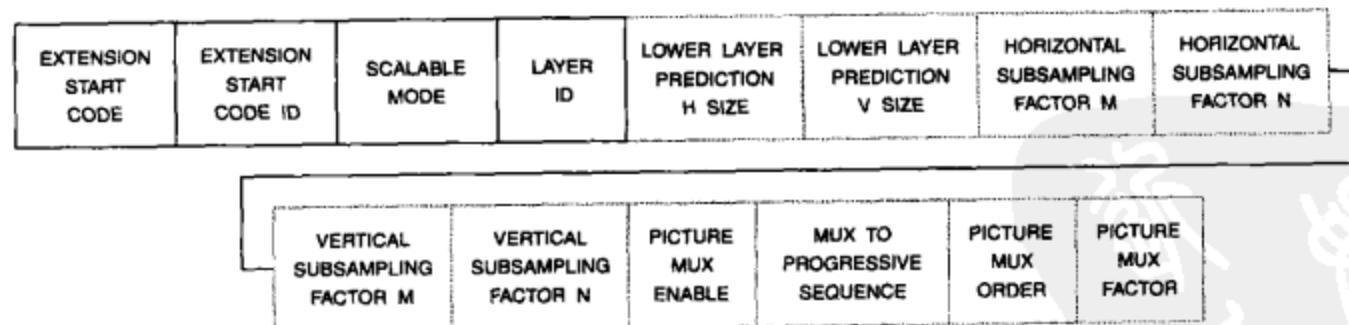


图13-8 MPEG-2序列可分级扩展结构。标记位没有给出

#### 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示新的扩展数据集的开始。

#### 2. extension\_start\_code\_ID

长度为4位的字段，值为0101，表示序列可分级扩展的开始。该扩展定义了位视频比特流实现的可分级模式。如果sequence\_scalable\_extension没有出现在比特流中，就没有可分级性被采用。

除了在数据分区可分级模式中外，其他可分级层次结构的基本层都没有sequence\_scalable\_extension。

### 3. scalable\_mode

两位的码字，表示视频序列的可分级类型，如表13-17所示。

### 4. layer\_ID

4位的二进制数，标识在可分级层次结构中的层次。基本层ID为0000。在数据分区中，0000的layer\_ID被分配给了分区层0，而0001的layer\_ID被分配给了分区层1。

### 5. lower\_layer\_prediction\_horizontal\_size

可选的14位二进制数，仅当scalable\_mode=01时才出现。它表示用于预测的低层帧的水平大小。它包含了在低层比特流中的horizontal\_size值。

### 6. marker\_bit

标记位，值恒为1。它只有在scalable\_mode=01时才出现。

### 7. lower\_layer\_prediction\_vertical\_size

可选的14位二进制数，仅当scalable\_mode=01时才出现。它表示用于预测的低层帧的垂直大小。它包含了在低层比特流中的vertical\_size值。

### 8. horizontal\_subsampling\_factor\_m

可选的5位二进制数，仅当scalable\_mode=01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

### 9. horizontal\_subsampling\_factor\_n

可选的5位二进制数，仅当scalable\_mode=01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

### 10. vertical\_subsampling\_factor\_m

可选的5位二进制数，仅当scalable\_mode=01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

### 11. vertical\_subsampling\_factor\_n

可选的5位二进制数，仅当scalable\_mode=01时才出现，该参数对于空域上采样过程有影响。取值不能为00000。

### 12. picture\_mux\_enable

可选的1位字段，仅当scalable\_mode=11时才出现。如果设置为1，picture\_mux\_order和picture\_mux\_factor参数将用于在显示之前的重复用（remultiplexing）。

### 13. mux\_to\_progressive\_sequence

可选的1位标志位，仅当scalable\_mode=11且picture\_mux\_enable=1时才出现。设置为1表示解码图像将在时域上复用，以生成用于显示的逐行序列。当时域复用将生成隔行序列时，该标志位需要设置为0。

### 14. picture\_mux\_order

可选的3位二进制数，仅当scalable\_mode=11时才出现。它定义了在第一个基本层图像之前的增强层图像数目。该参数将用于帮助解码器在显示之前以适当的方式重复用图像。

### 15. picture\_mux\_factor

可选的3位二进制数，仅当scalable\_mode=11时才出现。它表示在两个连续的基本层图像之

表13-17 MPEG-2 scalable\_mode码字

可分级模式	编 码
数据分区	00
空域可分级性	01
SNR可分级性	10
时域可分级性	11

间的增强层图像数目。该参数将用于帮助解码器在显示之前以适当的方式重复用图像。

### 13.4.7 图像组 (GOP) 层

GOP头应该每两秒钟出现一次。每一个图像组的数据都由一个GOP头和跟随其后的图像数据组成，其结构如图13-5所示。DVD标准在这一层为隐藏字幕数据采用了用户数据扩展。

#### 1. group\_start\_code

长度为32位的字符串，值为0x000001B8，表示一个图像组的开始。

#### 2. time\_code

表示时间码信息的25位字段，如表13-18所示。  
仅当帧率为30/1.001(29.97)Hz时，  
drop\_frame\_flag才可以被设置为1。

#### 3. closed\_gop

1位标志位，当图像组没有参考前一个图像组运动矢量进行编码时，该标志位设置为1。该位使得能够支持对压缩比特流的编辑。

#### 4. broken\_link

1位标志位，在解码过程中设置为0。在编辑过程中，当跟在图像组的第一个I帧之后的B帧不能正确地解码时，该标志位设置为1。

表13-18 MPEG-2 time\_code字段

Timecode	取值范围	位数
drop_frame_flag		1
time_code_hours	0~23	5
time_code_minutes	0~59	6
marker_bit	1	1
time_code_seconds	0~59	6
time_code_pictures	0~59	6

603

### 13.4.8 图像层

每一个图像的数据由一个图像头与跟随在之后的条带数据组成，其结构如图13-5所示。如果序列扩展出现了，每一个图像头将跟在图像编码扩展之后。

在这一层，通过用户数据扩展，有些实现使得帧与帧之间可以进行精准的高宽比信息切换。ATSC标准同样在这一层为CEA-708隐藏字幕数据采用了用户数据扩展。

#### 1. picture\_start\_code

长度为32位的字符串，值为0x00000100。

#### 2. temporal\_reference

10位的二进制数，对于GOP中的第一帧，temporal\_reference值为零。对于随后的显示顺序中每一帧，temporal\_reference以1024为模依次递增（增量为1）。当某一帧以两场编码时，两个场的时域参考是一致的。

#### 3. picture\_coding\_type

3位的码字，表示图像类型（I图像、P图像或者B图像），如表13-19所示。

#### 4. vbv\_delay

16位的二进制数，对于恒定比特率，在开始解码一个图像时，该参数将设置解码缓存的初始值，从而不会导致解码时发生上溢或者下溢的情况。对于ATSC和OpenCable™标准，除非vbv\_delay值为0xFFFF，否则其值必须小于等于45000<sub>D</sub>。

表13-19 MPEG-2 picture\_coding\_type码字

图像类型	编 码
禁止	000
I图像	001
P图像	010
B图像	011
禁止	100
保留	101
保留	110
保留	111

604

**5. full\_pel\_forward\_vector**

可选的1位字段，没有用于MPEG-2，因此其值为0。仅当picture\_coding\_type = 010或者011时才出现。

**6. forward\_f\_code**

可选的3位字段，没有用于MPEG-2，因此其值为111。仅当picture\_coding\_type = 010或者011时才出现。

**7. full\_pel\_backward\_vector**

可选的1位字段，没有用于MPEG-2，因此其值为0。仅当picture\_coding\_type = 011时才出现。

**8. forward\_f\_code**

可选的3位字段，没有用于MPEG-2，因此其值为111。仅当picture\_coding\_type = 011时才出现。

**9. extra\_bit\_picture**

1位标志位，当设置为1时，表示其后跟随content\_description\_data。当设置为0时，表示其后没有content\_description\_data。

**10. content\_description\_data**

如果extra\_bit\_picture = 1，该可选的可变长字段将会出现，其中每隔8位有一位值为1。

**11. extra\_bit\_picture**

1位标志位，值为0，且仅当content\_description\_data出现时才出现。

### 13.4.9 内容描述数据

可选的数据，且仅当图像头中的extra\_bit\_picture指示其出现时才出现。

**1. data\_type\_upper**

8位的字段，包含了定义内容描述数据类型的16位二进制参数data\_type的8位最高有效位，如表13-20所示。

**2. marker\_bit**

标记位，值恒为1。

**3. data\_type\_lower**

8位字段，定义内容描述数据类型的16位二进制参数data\_type的8位最低有效位，如表13-20所示。

**4. marker\_bit**

标记位，值恒为1。

**5. data\_length**

8位的二进制数，定义了接下来的数据长度，以9位为单位表示。

**605** 注意：当data\_type为“填充字节（padding\_bytes）”时，下列字段将出现。这两个字段将重复多次，次数由data\_length字段指定。

表13-20 MPEG-2 data\_type码字

数据类型	编 码
保留	0000 0000 0000 0000
填充字节	0000 0000 0000 0001
捕获时间码	0000 0000 0000 0010
摇摄扫描参数	0000 0000 0000 0011
有效区域窗口	0000 0000 0000 0100
编码图像长度	0000 0000 0000 0101
保留	0000 0000 0000 0110
:	:
保留	1111 1111 1111 1111

**6. marker\_bit**

标记位，值恒为1。

**7. padding\_byte**

8位字段，值为0000 0000。所有其他值都被禁止使用。

注意：当data\_type为“捕获时间码（capture timecode）”时，下列字段将出现。它包含了表示帧或场的源捕获或创建时间的时间戳。这两个字段将重复多次，次数由data\_length字段指定。它并不优先于任何时间码出现在系统级。

**8. marker\_bit**

标记位，值恒为1。

**9. timecode\_type**

2位码字，表示与图像相关的时间码数目，如表13-21所示。

**10. counting\_type**

可选的3位码字，它指定了用于修正nframes计数参数的方法，从而减少漂移的累积（drift accumulation）。

**11. reserved\_bit**

保留位，值恒为0。

**12. reserved\_bit**

保留位，值恒为0。

**13. reserved\_bit**

保留位，值恒为0。

**14. marker\_bit**

可选标记位，值恒为1。仅当counting\_type不等于000时该字段出现。

**15. nframes\_conversion\_code**

可选位，在确定由nframes参数表示的时间量时，该位定义转换因子（1000 + nframes\_conversion\_code）。仅当counting\_type不等于000时该字段出现。

**16. clock\_divisor**

可选的7位二进制数，定义了27MHz系统时钟的因子数，从而用于产生相同的时间戳。仅当counting\_type不等于000时该字段出现。

**17. marker\_bit**

可选标记位，值恒为1。仅当counting\_type不等于000时该字段出现。

**18. nframes\_multiplier\_upper**

可选的8位值，为16位nframes\_multiplier值的8位最高有效位。仅当counting\_type不等于000时该字段出现。

**19. marker\_bit**

可选标记位，值恒为1。仅当counting\_type不等于000时该字段出现。

**20. nframes\_multiplier\_lower**

可选的8位值，为16位nframes\_multiplier值的8位最低有效位。仅当counting\_type不等于000时该字段出现。

下面是“场或帧捕获时间戳”信息。

表13-21 MPEG-2 timecode\_type码字

时间码类型	编 码
整个帧一个时间码	00
第一场或仅有的一场一个时间码	01
第二场一个时间码	10
两个时间码，两个场各一个	11

**21. marker\_bit**

可选标记位，值恒为1。仅当counting\_type不等于000时该字段出现。

**22. nframes**

可选的8位二进制数，在导出等价时间戳时，指定了增加的帧时间增量数。仅当counting\_type不等于000时该字段出现。

**23. marker\_bit**

标记位，值恒为1。

**24. time\_discontinuity**

1位标志位，值为1表示时间码序列中发生了中断。

**25. prior\_count\_dropped**

1位标志位，表示是否丢弃了一个或多个nframes值计数。

**26. time\_offset\_part\_a**

6位值，包含了time\_offset的6位最高有效位。time\_offset是一个30位的符号值，定义了偏移自其他时间戳参数（当前场或帧被丢弃时，这些参数定义了相同时间戳）定义的时间的时钟周期数。

**27. marker\_bit**

标记位，值恒为1。

**28. time\_offset\_part\_b**

8位值，包含了time\_offset的第二个8位最高有效位。

**29. marker\_bit**

标记位，值恒为1。

**30. time\_offset\_part\_c**

8位值，包含了time\_offset的第三个8位最高有效位。

**31. marker\_bit**

标记位，值恒为1。

**32. time\_offset\_part\_d**

8位值，包含了time\_offset的8位最低有效位。

607

**33. marker\_bit**

标记位，值恒为1。

**34. units\_of\_seconds**

4位的二进制数，表示秒钟时间戳值。其值可以为0000~1001。

**35. tens\_of\_seconds**

4位的二进制数，表示10秒钟时间戳值。其值可以为0000~0101。

**36. marker\_bit**

标记位，值恒为1。

**37. units\_of\_minutes**

4位的二进制数，表示分钟时间戳值。其值可以为0000~1001。

**38. tens\_of\_minutes**

4位的二进制数，表示10分钟时间戳值。其值可以为0000~0101。

**39. marker\_bit**

标记位，值恒为1。

**40. units\_of\_hours**

4位的二进制数，表示小时时间戳值。其值可以为0000~1001。

**41. tens\_of\_hours**

4位的二进制数，表示10小时时间戳值。其值可以为0000~0010。

当timecode\_type = 11时，“场或帧捕获时间戳”字段又再次出现，从而为第二个字段传递信息。

注意：当date\_type为“摄像参数”时，下面的字段将出现。这些字段将为有不同的高宽比的显示传输附加的摄像信息。

**42. marker\_bit**

标记位，值恒为1。

**43. aspect\_ratio\_information**

4位码字，与序列头所使用的高宽比信息一致。

**44. reserved\_bit**

保留位，值恒为0。

**45. reserved\_bit**

保留位，值恒为0。

**46. reserved\_bit**

保留位，值恒为0。

**47. display\_size\_present**

1位标志位，表示接下来是否有display\_horizontal\_size和display\_vertical\_size字段。

**48. marker\_bit**

标记位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**49. reserved\_bit**

保留位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

[608]

**50. reserved\_bit**

保留位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**51. display\_horizontal\_size\_upper**

该字段为display\_horizontal\_size的6位最高有效位。仅当display\_size\_present = 1时该可选字段才出现。

**52. marker\_bit**

标记位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**53. display\_horizontal\_size\_lower**

该字段为display\_horizontal\_size的8位最低有效位。仅当display\_size\_present = 1时该可选字段才出现。

**54. marker\_bit**

标记位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**55. reserved\_bit**

保留位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**56. reserved\_bit**

保留位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**57. display\_vertical\_size\_upper**

该字段为display\_vertical\_size的6位最高有效位。仅当display\_size\_present = 1时该可选字段才出现。

**58. marker\_bit**

标记位，值恒为1。仅当display\_size\_present = 1时该可选字段才出现。

**59. display\_vertical\_size\_lower**

该字段为display\_vertical\_size的8位最低有效位。仅当display\_size\_present = 1时该可选字段才出现。

**注意：**对于出现的每一个帧中心偏移 (frame center offsets)，下面的字段都将出现。

**60. marker\_bit**

标记位，值恒为1。

**61. frame\_center\_horizontal\_offset\_upper**

frame\_center\_horizontal\_offset的8位最高有效位。frame\_center\_horizontal\_offset的定义在图像显示扩展中给出了。

**62. marker\_bit**

标记位，值恒为1。

**63. frame\_center\_horizontal\_offset\_lower**

frame\_center\_horizontal\_offset的8位最低有效位。

**64. marker\_bit**

标记位，值恒为1。

**65. frame\_center\_vertical\_offset\_upper**

frame\_center\_vertical\_offset的8位最高有效位。frame\_center\_vertical\_offset的定义在图像显示扩展中给出了。

**66. marker\_bit**

标记位，值恒为1。

**67. frame\_center\_vertical\_offset\_lower**

frame\_center\_vertical\_offset的8位最低有效位。

609

**注意：**当data\_type指定为“有效区域窗口 (active region window)”时，下列字段将出现。active\_region\_window定义了解码图像将要显示的矩形区域。该窗口不能大于由horizontal\_size和vertical\_size定义的区域。

**68. marker\_bit**

标记位，值恒为1。

**69. top\_left\_x\_upper**

8位的字段，是16位的top\_left\_x的8位最高有效位。top\_left\_x定义了Y采样点数目，它与top\_left\_y一起定义了active\_region\_window的左上角。

**70. marker\_bit**

标记位，值恒为1。

**71. top\_left\_x\_lower**

8位的字段，是16位的top\_left\_x的8位最低有效位。

**72. marker\_bit**

标记位，值恒为1。

**73. top\_left\_y\_upper**

8位的字段，是16位的top\_left\_y的8位最高有效位。top\_left\_y定义了Y扫描行数目，它与top\_left\_x一起定义了active\_region\_window的左上角。

**74. marker\_bit**

标记位，值恒为1。

**75. top\_left\_y\_lower**

8位的字段，是16位的top\_left\_y的8位最低有效位。

**76. marker\_bit**

标记位，值恒为1。

**77. active\_region\_horizontal\_size\_upper**

8位字段，是16位的active\_region\_horizontal\_size的8位最高有效位。active\_region\_horizontal\_size与active\_region\_vertical\_size一起定义了active\_region\_window的右下角。当active\_region\_horizontal\_size字段值为0x000时，表示其大小未知。

**78. marker\_bit**

标记位，值恒为1。

**79. active\_region\_horizontal\_size\_lower**

8位的字段，是16位的active\_region\_horizontal\_size的8位最低有效位。

**80. marker\_bit**

标记位，值恒为1。

**81. active\_region\_vertical\_size\_upper**

8位的字段，是16位的active\_region\_vertical\_size的8位最高有效位。active\_region\_vertical\_size与active\_region\_horizontal\_size一起定义了active\_region\_window的右下角。当active\_region\_vertical\_size字段值为0x0000时，表示其大小未知。

**82. marker\_bit**

标记位，值恒为1。

**83. active\_region\_vertical\_size\_lower**

8位的字段，是16位的active\_region\_vertical\_size的8位最低有效位。

610

注意：当data\_type指定为“编码图像长度（coded picture length）”时，下面的字段将出现。

**84. marker\_bit**

标记位，值恒为1。

**85. picture\_byte\_count\_part\_a**

8位字段，是32位的picture\_byte\_count的8位最高有效位。picture\_byte\_count指示了从第一个slice\_start\_code的第一个字节开始，到紧跟在图像最后一个宏块之后的起始码前缀之前的那个字节之间的字节数目。当值为0x0000时，表示长度是不可知的。

**86. marker\_bit**

标记位，值恒为1。

**87. picture\_byte\_count\_part\_b**

8位字段，是32位的picture\_byte\_count的第二个8位最高有效位。

## 88. marker\_bit

标记位，值恒为1。

## 89. picture\_byte\_count\_part\_c

8位字段，是32位的picture\_byte\_count的第三个8位最高有效位。

## 90. marker\_bit

标记位，值恒为1。

## 91. picture\_byte\_count\_part\_d

8位字段，是32位的picture\_byte\_count的8位最低有效位。

注意：当没有其他数据出现（由data\_type指示决定）时，下面两个字段将出现。这两个字段会重复多次，重复次数由data\_length字段指定。

## 92. marker\_bit

标记位，值恒为1。

## 93. reserved\_content\_description\_data

保留位，长度为8位。

### 13.4.10 图像编码扩展

图像编码扩展只可能出现在图像头之后，其结构如图13-9所示。

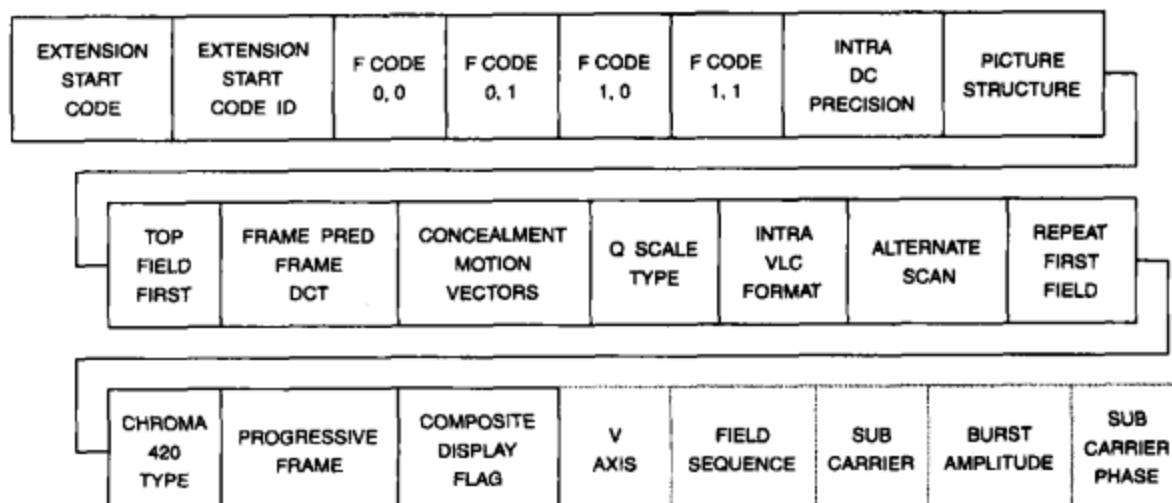


图13-9 MPEG-2图像编码扩展结构。没有给出标记位

## 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

## 2. extension\_start\_code\_ID

长度为4位的字段，值为1000，表示图像编码扩展的开始。

## 3. f\_code[0,0]

4位的二进制数，取值范围为0001~1001，用于前向水平运动矢量的解码。其值不能为0000，而值为1111表示该字段将被忽略。

## 4. f\_code[0,1]

4位的二进制数，取值范围为0001~1001，用于前向垂直运动矢量的解码。其值不能为0000，而值为1111表示该参数将被忽略。

## 5. f\_code[1,0]

4位的二进制数，取值范围为0001~1001，用于后向水平运动矢量的解码。其值不能为0000，

而值为1111表示该字段将被忽略。

#### 6. f\_code[1,1]

4位的二进制数，取值范围为0001~1001，用于后向垂直运动矢量的解码。其值不能为0000，而值为1111表示该字段将被忽略。

#### 7. intra\_dc\_precision

2位的码字，定义intra DC的精度，如表13-22所示。

#### 8. picture\_structure

2位的码字，定义图像结构，如表13-23所示。

表13-22 MPEG-2 intra\_dc\_precision码字

内部DC精度（位）	编 码
8	00
9	01
10	10
11	11

表13-23 MPEG-2 picture\_structure码字

图 像 结 构	编 码
保留	00
顶场	01
底场	10
帧图像	11

612

#### 9. top\_field\_first

如果progressive\_sequence=0，这一位表示解码器将最先输出哪一个场。在场中，这一位的值为0。在帧中，该位值为1表示解码帧的第一个场是顶场（top field）；该位值为0表示解码帧的第一个场是底场（bottom field）。

如果progressive\_sequence=1、repeat\_first\_field=0，且该位为0，那么解码器产生一个逐行扫描帧。

如果progressive\_sequence=1、repeat\_first\_field=1，且该位为0，那么解码器产生两个完全相同的逐行扫描帧。

如果progressive\_sequence=1、repeat\_first\_field=1，且该位为1，那么解码器产生三个完全相同的逐行扫描帧。

#### 10. frame\_pred\_frame\_dct

如果这一位为1，只有frame-DCT和帧预测被采用。对于场图像，该位恒为0。如果progressive\_frame为1，该参数为1。

#### 11. concealment\_motion\_vectors

这一位为1，表示运动矢量是为内部宏块而编码的。

#### 12. q\_scale\_type

这一位表示quantizer\_scale\_code和quantizer\_scale之间的两个映射中的哪一个被解码器所采用。

#### 13. intra\_vlc\_format

这一位表示对于内部块，DCT系数采用哪一个表。当intra\_vlc\_format=0时，采用表13-36；当intra\_vlc\_format=1时，采用表13-37。对于非内部块，无论intra\_vlc\_format值为多少，都是采用表13-36。

#### 14. alternate\_scan

这一位表示对于变换系数数据，解码器将采用哪一个扫描模式。当alternate\_scan=0时，采用图7-57所示模式；当alternate\_scan=1时，采用图7-58所示模式。

#### 15. repeat\_first\_field

对于这一位的使用，请参看top\_filed\_first。对于场图像，这一位值为0。

**16. chroma\_420\_type**

如果chroma\_format是4：2：0，那么这一位与progressive\_frame是一样的。否则，这一位值为0。

**17. progressive\_frame**

如果值为0，这一位表示帧的两个场是隔行场，两场之间有一个时间间隔。如果值为1，帧的两场来自同一时刻。

**18. composite\_display\_flag**

1位的标志位，表示v\_axis、field\_sequence、sub\_carrier、burst\_amplitude和sub\_carrier\_phase等是否会在比特流中。

**19. v\_axis**

仅当Composite\_display\_flag = 1时，这一位才出现。当初始源是一个PAL视频信号时，将采用这一位。当V符号为正时，v\_axis = 1；否则v\_axis = 0。

这一信息可以从正在驱动MPEG-2编码器的PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将PAL编码器的V开关设置成与初始源相同。  
613

**20. field\_sequence**

仅当composite\_display\_flag = 1时，这3位码字才会出现。该码字定义了在最初的4场或8场序列中的场编号，如表13-24所示。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的场序列设置成与初始源相同。

**21. sub\_carrier**

仅当composite\_display\_flag = 1时，这一位才出现。该位为0表示原始的副载波扫描线（subcarrier-to-line）的频率关系是正确的。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。

**22. burst\_amplitude**

仅当composite\_display\_flag = 1时，这个7位的二进制数才出现。该参数定义了每个BT.601量化时（忽略MSB），初始PAL或者NTSC的脉冲幅度。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的彩色脉冲幅度设置成与初始源相同。

**23. sub\_carrier\_phase**

仅当composite\_display\_flag = 1时，这个8位的二进制数才出现。该参数定义了初始PAL或者NTSC副载波相位，如BT.470所定义的一样。其值定义为： $(360^\circ / 256) \text{sub\_carrier\_phase}$ 。

这一信息可以从正在驱动MPEG-2编码器的NTSC/PAL解码器获得。这一信息也可用来使MPEG-2解码器能够将NTSC/PAL编码器的彩色副载波相位设置成与初始源相同。

### 13.4.11 量化矩阵扩展

每一个量化矩阵都有默认值。当对序列头解码时，所有的矩阵复位为它们的默认值。在序列头期间或者使用这个扩展的过程中，用户定义的矩阵可能会被下载。这个可选的扩展只可能出现

表13-24 MPEG-2 field\_sequence码字

帧序	列	场序	列	编 码
1		1		000
1		2		001
2		3		010
2		4		011
3		5		100
3		6		101
4		7		110
4		8		111

在图像编码扩展之后，其结构如图13-10所示。

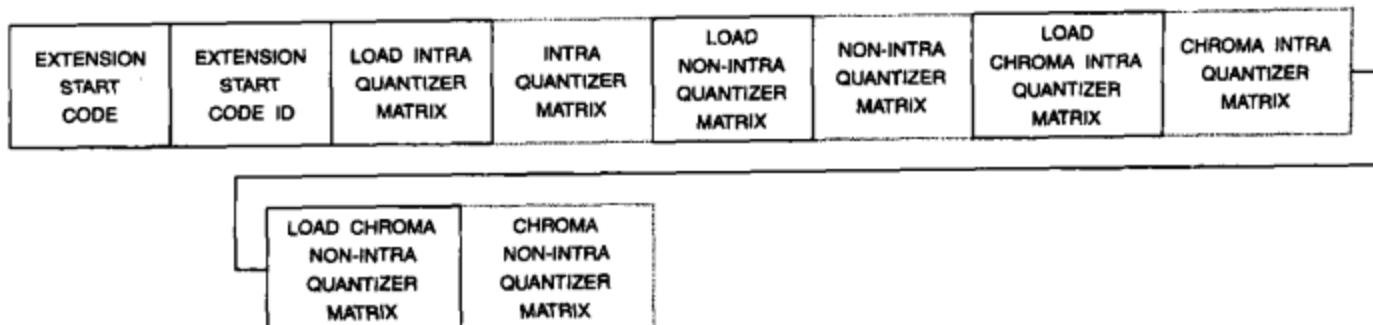


图13-10 MPEG-2量化矩阵扩展结构。没有给出标记位

#### 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示一个新扩展数据集的开始。

#### 2. extension\_start\_code\_ID

长度为4位的字段，值为0011，表示quant\_matrix\_extension的开始。这个扩展同样允许量化器矩阵以4：2：2和4：4：4色度格式传送。

#### 3. load\_intra\_quantizer\_matrix

如果在之后跟随一个intra\_quantizer\_matrix，这一位设置为1。如果设置为0，那么下面的默认值将用于内部块，直到序列头或者quant\_matrix\_extension的下一次出现。

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

#### 4. intra\_quantizer\_matrix

可选的64个8位值列表，用来替换上面给出的默认值。不允许为0。intra\_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant\_matrix\_extension的下一次出现。出现的顺序如图7-57所示。

对于4：2：2和4：4：4数据格式，新的值同时用于Y和CbCr内部矩阵，除非装载了一个不同的CbCr内部矩阵。

#### 5. load\_non\_intra\_quantizer\_matrix

如果其后紧跟non\_intra\_quantizer\_matrix，此位设为1。如果设为0，采用如下默认值，直到序列头或者quant\_matrix\_extension的下一次出现。

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

614

615

**6. non\_intra\_quantizer\_matrix**

可选的64个8位值列表，用来替换上面给出的默认值。不允许为0。这些数值一直有效，直到序列头或者quant\_matrix\_extension的下一次出现。出现的顺序如图7-57所示。

对于4：2：2和4：4：4数据格式，新的值同时用于Y和CbCr非内部矩阵，除非装载了一个新的CbCr非内部矩阵。

**7. load\_chroma\_intra\_quantizer\_matrix**

如果其后紧跟chroma\_intra\_quantizer\_matrix，此位设为1。如果设为0，在使用的值中没有发生改变。如果chroma\_format为4：2：0，这一位为0。

**8. chroma\_intra\_quantizer\_matrix**

可选的64个8位值列表，用来替换用于CbCr数据的前一个或者默认值。不允许为0。chroma\_intra\_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant\_matrix\_extension的下一次出现。出现的顺序如图7-57所示。

**9. load\_chroma\_non\_intra\_quantizer\_matrix**

如果其后紧跟chroma\_non\_intra\_quantizer\_matrix，此位设为1。如果设为0，在使用的值中没有发生改变。如果chroma\_format为4：2：0，这一位为0。

**10. chroma\_non\_intra\_quantizer\_matrix**

可选的64个8位值列表，用来替换用于CbCr数据的前一个或者默认值。不允许为0。chroma\_intra\_quant[0,0]恒为8。这些数值一直有效，直到序列头或者quant\_matrix\_extension的下一次出现。出现的顺序如图7-57所示。

### 13.4.12 图像显示扩展

这一扩展准许显示矩阵的位置在图像之间发生移动。一个典型的应用可能就是实现摇摄和扫描。这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-11所示。

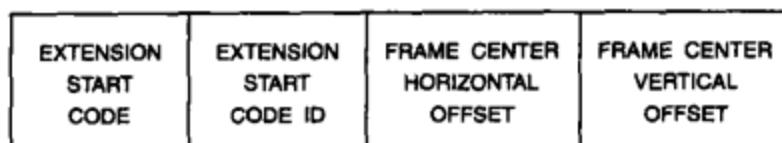


图13-11 MPEG-2图像显示扩展结构。没有给出标记位

**1. extension\_start\_code**

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

**2. extension\_start\_code\_ID**

长度为4位的字段，值为0111，表示图像显示扩展的开始。

在隔行扫描序列情况下，一个图像可能与一个、两个或者三个解码的场相关。因此，在比特流中有可能至少出现3个由下面4个字段组成的集。

**3. frame\_center\_horizontal\_offset**

16位二进制补码，定义了水平偏移（单位为十六分之一采样点）。正数值决定解码图像的中心位置位于显示区域中心的右侧。

**4. marker\_bit**

标记位，值恒为1。

**5. frame\_center\_vertical\_offset**

16位二进制补码，定义了垂直偏移（单位为十六分之一扫描线）。正数值决定解码图像的中心位置位于显示区域中心的下方。

### 6. marker\_bit

标记位，值恒为1。

#### 13.4.13 图像时域可分级扩展

这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-12所示。

EXTENSION START CODE	EXTENSION START CODE ID	REFERENCE SELECT CODE	FORWARD TEMPORAL REFERENCE	BACKWARD TEMPORAL REFERENCE
----------------------------	-------------------------------	-----------------------------	----------------------------------	-----------------------------------

图13-12 MPEG-2图像时域可分级扩展结构。标记位没有给出

### 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

### 2. extension\_start\_code\_ID

长度为4位的字段，值为1010，表示图像时域可分级扩展的开始。

### 3. reference\_select\_code

2位的码字，表示用于预测的参考帧或场

### 4. forward\_temporal\_reference

10位的二进制数，指示用于提供前向预测的低层的时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

### 5. marker\_bit

标记位，值恒为1。

### 6. backward\_temporal\_reference

10位的二进制数，指示用于提供后向预测的低层的时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

617

#### 13.4.14 图像空域可分级扩展

这个可选的扩展只可能出现在图像编码扩展之后，其结构如图13-13所示。

EXTENSION START CODE	EXTENSION START CODE ID	LOWER LAYER TEMPORAL REFERENCE	LOWER LAYER HORIZONTAL OFFSET	LOWER LAYER VERTICAL OFFSET	Spatial TEMPORAL WEIGHT CODE	LOWER LAYER PROGRESSIVE FRAME
LOWER LAYER DEINTERLACED FIELD SELECT						

图13-13 MPEG-2图像空域可分级扩展结构。标记位没有给出

### 1. extension\_start\_code

长度为32位的字符串，值为0x000001B5，表示一个新的扩展数据集的开始。

### 2. extension\_start\_code\_ID

长度为4位的字段，值为1001，表示图像空域可分级扩展的开始。

### 3. lower\_layer\_temporal\_reference

10位的二进制数，指示用于提供预测的低层时域参考。如果需要多于10位来定义时域参考，仅10位最低有效位被采用。

**4. marker\_bit**

标记位，值恒为1。

**5. lower\_layer\_horizontal\_offset**

15位的二进制数，指示上采样低层图像的左上角相对于增强层图像的水平偏移量。对于4:2:0和4:2:2格式，这个参数必须是一个偶数。

**6. marker\_bit**

标记位，值恒为1。

**7. lower\_layer\_vertical\_offset**

15位的二进制数，指示上采样低层图像的左上角相对于增强层图像的垂直偏移量。对于4:2:0和4:2:2格式，这个参数必须是一个偶数。

**8. spatial\_temporal\_weight\_code\_table\_index**

2位的码字，指示将采用哪一个空时域权重码。

**9. lower\_layer\_progressive\_frame**

如果低层图像是逐行的，这一位将为1。

**10. lower\_layer\_deinterlaced\_field\_select**

这一位将与其他参数联合使用，以辅助解码器解码。参见表13-25。

**表13-25 MPEG-2图像空域可分级扩展上采样过程**

低层去隔行扫描场选择	低层逐行扫描帧	逐行扫描帧	应用去隔行扫描过程	用于预测
0	0	1	是	顶场
1	0	1	是	底场
1	1	1	否	帧
1	1	0	否	帧
1	0	0	是	两个场

### 13.4.15 版权扩展

这个可选的扩展只可能出现在图像编码扩展之后。

**1. extension\_start\_code**

长度为32位的字符串，值为0x0000001B5，表示一个新的扩展数据集的开始。

**2. extension\_start\_code\_ID**

长度为4位的字段，值为0100，表示版权扩展的开始。

**3. copyright\_flag**

1位的标志位，如果为1，定义随后的视频内容是受版权保护的，直到下一个版权扩展的出现。如果为0，没有指示随后的视频内容是否受到版权保护。

**4. copyright\_identifier**

8位的二进制数，指示版权拥有者。值为0000 0000表示这一信息没有提供。当copyright\_flag=0时，copyright\_identifier必须为0000 0000。

**5. original\_or\_copy**

值为1表示为原始内容，为0表示为副本。

**6. reserved**

7位保留位，值恒为000 0000。

**7. marker\_bit**

标记位，值恒为1。

**8. copyright\_number\_1**

表示版权编码的第44~63位的20位数。

**9. marker\_bit**

标记位，值恒为1。

**10. copyright\_number\_2**

表示版权编码的第22~43位的22位数。

**11. marker\_bit**

标记位，值恒为1。

**12 copyright\_number\_3**

表示版权编码的第0~21位的22位数。64位的copy\_number唯一标识了受版权保护的内容。当copyright\_identifier = 0000 0000或者copyright\_flag = 0时，copyright\_number必须为零。

619

### 13.4.16 摄像头参数扩展

这个可选的扩展只可能出现在图像编码扩展之后。

在值为0x000001B5的32位extension\_start\_code和值为1011的4位extension\_start\_code\_ID之后，有几个字段定义所采用摄像头的位置和特性。

### 13.4.17 ITU-T ext. D扩展

这个可选的扩展只可能出现在图像编码扩展之后。

在值为0x000001B5的32位extension\_start\_code和值为1011的4位extension\_start\_code\_ID之后，有1位数据。这一扩展的使用方法在ITU-T H.320 Annex A中给出了定义。

### 13.4.18 条带层

每一个条带层的数据都是由一个条带头及跟在之后的宏块数据组成的。条带层的结构在图13-5中给出了。

**1. slice\_start\_code**

32位的条带起始码，前24位值为0x000001。后8位为slice\_vertical\_position，其取值范围为0x01~0xAF。

slice\_vertical\_position定义了第一个宏块在条带中的垂直位置（以宏块为单位）。宏块的第一行的slice\_vertical\_position为1。

**2. slice\_vertical\_position\_extension**

可选的3位二进制数，如果帧的垂直大小大于2800行，该字段表示为11位的slice\_vertical\_position值的3个最高有效位；否则，这个字段将不会出现。

**3. priority\_breakpoint**

可选的7位二进制数，仅当sequence\_scalable\_extension出现在比特流中且scalable\_mode = “数据分区”时才出现。它定义了在比特流中的哪个位置进行分区。

**4. quantizer\_scale\_code**

5位的二进制数，取值为1~31（0值被禁用）。它定义了接收到的DCT系数的重构非零值的比例因子。解码器一直利用这个值，直到在条带层或者宏块层接收到另一个quantizer\_scale\_code为

止。

#### 5. slice\_extension\_flag

1位可选的标志位，如果值为1，其后将出现intra\_slice、slice\_picture\_ID\_enable和slice\_picture\_ID字段。

#### 6. intra\_slice

可选位，仅当slice\_extension\_flag=1时才出现。如果条带中的任一宏块为非内部宏块，这一位必须设置为0。  
[620]

#### 7. slice\_picture\_ID\_enable

可选位，仅当slice\_extension\_flag=1时才出现。值为1表示采用了slice\_picture\_ID。

#### 8. slice\_picture\_ID

这些可选的6位字段用于辅助恢复严重的突发错误。对于一幅图像中的所有条带，slice\_picture\_ID值必须相同。仅当slice\_extension\_flag=1时这个字段才出现。如果slice\_picture\_ID\_enable=0，这6位必须为00 0000

#### 9. extra\_bit\_slice

1位的标志位，设置为1表示extra\_information\_slice跟随在其后。设置为0表示在这个字段之后没有数据。

#### 10. extra\_information\_slice

如果extra\_bit\_slice=1，那么这个长为9位的字段跟随在其后，其中8位为数据(extra\_information\_slice)，第9位为另一个extra\_bit\_slice，指示是否有另一个9位跟随在其后，并依此循环。

### 13.4.19 宏块层

每一个宏块层的数据都是由一个宏块头与跟随在其后的运动矢量和块数据组成的，如图13-5所示。

#### 1. macroblock\_escape

这个可选的11位字段是一个固定的比特字符串0000 0001 000，在当前宏块地址与前一个宏块地址之间差值大于33时采用。该字段强制macroblock\_address\_increment的值增加33。任意个连续的macroblock\_escape字段都可能被采用。

#### 2. macroblock\_address\_increment

这是一个变长的码字，定义当前宏块地址与前一个宏块地址之间的差值。它的最大值为33。当值大于33时，采用macroblock\_escape字段来编码。表13-26中给出了这些变长码字。

表13-26 MPEG-2 macroblock\_address\_increment可变长编码表

增 量 值	编 码	增 量 值	编 码
1	1	9	0000 110
2	011	10	0000 1011
3	010	11	0000 1010
4	0011	12	0000 1001
5	0010	13	0000 1000
6	0001 1	14	0000 0111
7	0001 0	15	0000 0110
8	0000 111	16	0000 0101 11

(续)

增量值	编 码	增量值	编 码
17	0000 0101 10	25	0000 0100 000
18	0000 0101 01	26	0000 0011 111
19	0000 0101 00	27	0000 0011 110
20	0000 0100 11	28	0000 0011 101
21	0000 0100 10	29	0000 0011 100
22	0000 0100 011	30	0000 0011 011
23	0000 0100 010	31	0000 0011 010
24	0000 0100 001	32	0000 0011 001
		33	0000 0011 000
	macroblock_escape		0000 0001 000

### 3. macroblock\_type

可变长的码字，表示编码的方法和宏块的内容，如表13-27、表13-28和表13-29所示。

表13-27 MPEG-2中I图像macroblock\_type的可变长编码表

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
intra	0	0	0	0	1	0	0	1
intra, quant	1	0	0	0	1	0	0	01
空域可分级的I图像								
coded, compatible	0	0	0	1	0	0	4	1
coded, compatible, quant	1	0	0	1	0	0	4	01
intra	0	0	0	0	1	0	0	0011
intra, quant	1	0	0	0	1	0	0	0010
not coded, compatible	0	0	0	0	0	0	4	0001
SNR可分级的I图像								
coded	0	0	0	1	0	0	0	1
coded, quant	1	0	0	1	0	0	0	01
not coded	0	0	0	0	0	0	0	001

表13-28a MPEG-2中P图像macroblock\_type的可变长编码表

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
mc, coded	0	1	0	1	0	0	0	1
no mc, coded	0	0	0	1	0	0	0	01
mc, not coded	0	1	0	0	0	0	0	001
intra	0	0	0	0	1	0	0	0001 1

(续)

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
mc,coded, quant	1	1	0	1	0	0	0	0001 0
no mc,coded, quant	1	0	0	1	0	0	0	0000 1
intra,quant	1	0	0	0	1	0	0	0000 01
SNR可分级的P图像								
coded	0	0	0	1	0	0	0	1
coded, quant	1	0	0	1	0	0	0	01
not coded	0	0	0	0	0	0	0	001

表13-28b MPEG-2中P图像macroblock\_type的可变长编码表

类 型	Macroblock Quant	Motion Forward	Motion Backward	Coded Pattern	Intra Macro-block	Spatial Temporal Weight Code Flag	Permitted Spatial Temporal Weight Class	编码
空域可分级的P图像								
mc, coded	0	1	0	1	0	0	0	10
mc.coded, compatible	0	1	0	1	0	1	1, 2, 3	011
no mc, coded	0	0	0	1	0	0	0	0000 100
no mc, coded, compatible	0	0	0	1	0	1	1, 2, 3	0001 11
mc, not coded	0	1	0	0	0	0	0	0010
intra	0	0	0	0	1	0	0	0000 111
mc, not coded, compatible	0	1	0	0	0	1	1, 2, 3	0011
mc, coded, quant	1	1	0	1	0	0	0	010
no mc,coded, quant	1	0	0	1	0	0	0	0001 00
intra,quant	1	0	0	0	1	0	0	0000 110
mc, coded, compatible, quant	1	1	0	1	0	1	1, 2, 3	11
no mc, coded, compatible, quant	1	0	0	1	0	1	1, 2, 3	0001 01
no mc, not coded, compatible	0	0	0	0	0	1	1, 2, 3	0001 10
coded, compatible	0	0	0	1	0	0	4	0000 101
coded, compatible, quant	1	0	0	1	0	0	4	0000 010
not coded, compatible	0	0	0	0	0	0	4	0000 0011

表13-29 MPEG-2中B图像macroblock\_type的可变长编码表

**4. spatial\_temporal\_weight\_code**

可选的2位码字，表示在空域可分级情况下，时域和空域预测是如何联合为宏块进行预测的。仅当表13-27、表13-28和表13-29中的[spatial temporal weight class] = 1且spatial\_temporal\_weight\_code\_table\_index ≠ 00时该字段才出现。

**5. frame\_motion\_type**

可选的2位码字，表示如表13-30所示的宏块运动预测。仅当picture\_structure = frame, frame\_pred\_frame\_dct = 0且表13-27、表13-28和表13-29中的[motion forward]或者[motion backward] = 1时该字段才出现。

表13-30 MPEG-2 frame\_motion\_type码字

空域时域权重类	预测类型	运动矢量统计	运动矢量格式	编 码
	保留			00
0,1	场	2	场	01
2,3	场	1	场	01
0,1,2,3	帧	1	帧	10
0,2,3	对偶	1	场	11

**6. field\_motion\_type**

可选的2位码字，表示如表13-31所示的宏块运动预测。仅当表13-27、表13-28和表13-29中的[motion forward]或者[motion backward] = 1且frame\_motion\_type字段没有出现时该字段才出现。  
621

表13-31 MPEG-2 field\_motion\_type码字

空域时域权重类	预测类型	运动矢量统计	运动矢量格式	编 码
	保留			00
0,1	场	1	场	01
0,1	16 × 8 mc	2	场	10
0	对偶	1	场	11

**7. dct\_type**

可选标志位，表示宏块是帧DCT编码还是场DCT编码。1表示是场，0表示是帧。仅当picture\_structure = 11、frame\_pred\_frame\_dct = 0并且表13-27、表13-28和表13-29中的[intra macroblock]或者[coded pattern] = 1时该标志位才出现。

**8. quantizer\_scale\_code**

可选的5位二进制数，取值为1~31（0值被禁用）。它定义了接收到的DCT系数的重构非零值的比例因子。解码器利用这个值直到接收到另一个quantizer\_scale\_code为止。仅当表13-27、表13-28和表13-29中的[macroblock quant] = 1时这个字段才出现。

以下为可选的运动矢量。

**9. marker\_bit**

标记位，值恒为1。仅当concealment\_motion\_vectors = 1且表13-27、表13-28和表13-29中的[intra macroblock] = 1时该标记位才出现。

**10. coded\_block\_pattern\_420**

这个可选的可变长码字用于推导4 : 2 : 0编码块模式（CBP），如表13-32所示。仅当表13-27、表13-28和表13-29中的[coded pattern] = 1时该字段才出现，它指示宏块中哪一个块至少有一个变换系数已被传送。编码块模式编号表示如下：

$P_1P_2P_3P_4P_5P_6$

上述编号中，对于块[n]，其任一系数出现则 $P_n=1$ ，否则 $P_n=0$ 。图13-2中给出了块编号（十进制格式）。

表13-32 MPEG-2 coded\_block\_pattern\_420可变长编码表

编码块模式 (CBP)	编 码		编码块模式 (CBP)	编 码	
60	111		62	0100	0
4	1101		24	0011	11
8	1100		36	0011	10
16	1011		3	0011	01
32	1010		63	0011	00
12	1001	1	5	0010	111
48	1001	0	9	0010	110
20	1000	1	17	0010	101
40	1000	0	33	0010	100
28	0111	1	6	0010	011
44	0111	0	10	0010	010
52	0110	1	18	0010	001
56	0110	0	34	0010	000
1	0101	1	7	0001	1111
61	0101	0	11	0001	1110
2	0100	1	19	0001	1101
35	0001	1100	38	0000	1100
13	0001	1011	29	0000	1011
49	0001	1010	45	0000	1010
21	0001	1001	53	0000	1001
41	0001	1000	57	0000	1000
14	0001	0111	30	0000	0111
50	0001	0110	46	0000	0110
22	0001	0101	54	0000	0101
42	0001	0100	58	0000	0100
15	0001	0011	31	0000	0011 1
51	0001	0010	47	0000	0011 0
23	0001	0001	55	0000	0010 1
43	0001	0000	59	0000	0010 0
25	0000	1111	27	0000	0001 1
37	0000	1110	39	0000	0001 0
26	0000	1101	0*	0000	0000 1

\* 不被4:2:0色度结构所采用。

#### 11. coded\_block\_pattern\_1

仅当chroma\_format=4:2:2并且表13-27、表13-28和表13-29中的[coded pattern]=1时才出现。这个可选的2位字段被用来扩展编码块模式。

#### 12. coded\_block\_pattern\_2

仅当chroma\_format=4:2:4并且表13-27、表13-28和表13-29中的[coded pattern]=1时才出现。

这个可选的6位字段被用来扩展编码块模式。

### 13.4.20 块层

每一个块层的数据是由系数数据所组成的。图13-5给出了块层的结构。

#### 1. dct\_dc\_size\_luminance

仅当当前块为Y内部编码块时这个可选的可变长编码才出现，它定义了在接下来的dct\_dc\_differential中的位数。表13-33中给出了其取值。

表13-33 MPEG-2 dct\_dc\_size\_luminance 可变长编码表

DCT DC Size Luminance	编 码	DCT DC Size Luminance	编 码
0	100	6	1111 0
1	00	7	1111 10
2	01	8	1111 110
3	101	9	1111 1110
4	110	10	1111 1111 0
5	1110	11	1111 1111 1

#### 2. dct\_dc\_differential

如果dct\_dc\_luminance ≠ 0，这个可选的可变长码将出现。表13-34中给出了其取值。

表13-34 MPEG-2 dct\_dc\_size\_chrominance 可变长编码表

DCT DC Size Chrominance	编 码	DCT DC Size Chrominance	编 码
0	00	6	1111 10
1	01	7	1111 110
2	10	8	1111 1110
3	110	9	1111 11110
4	1110	10	1111 1111 10
5	11110	11	1111 1111 11

#### 3. dct\_dc\_size\_chrominance

仅当当前块为CbCr内部编码块时这个可选的可变长码才出现，它定义了在接下来的dct\_dc\_differential中的位数。表13-35中给出了其取值。

表13-35 MPEG-2 dct\_dc\_differential 可变长编码表

DCT DC Differential	DCT DC Size	编码 (Y)	编码 (CbCr)	附加 编 码
-2048~-1024	11	111111111	111111111	0000000000~0111111111
-1023~-512	10	111111110	111111110	000000000~0111111111
-511~-256	9	11111110	111111110	000000000~011111111
-255~-128	8	1111110	11111110	00000000~01111111
-127~-64	7	111110	1111110	0000000~0111111
-63~-32	6	11110	111110	000000~011111
-31~-16	5	1110	11110	00000~01111
-15~-8	4	110	1110	0000~0111
-7~-4	3	101	110	000~011
-3~-2	2	01	10	00~01

(续)

DCT DC Differential	DCT DC Size	编码 (Y)	编码 (CbCr)	附加编码
-1	1	00	01	0
0	0	100	00	
1	1	00	01	1
2~3	2	01	10	10~11
4~7	3	101	110	100~111
8~15	4	110	1110	1000~1111
16~31	5	1110	11110	10000~11111
32~63	6	11110	111110	100000~111111
64~127	7	111110	1111110	100000~01111111
128~255	8	1111110	11111110	10000000~11111111
256~511	9	11111110	111111110	100000000~111111111
512~1023	10	111111110	1111111110	1000000000~1111111111
1024~2047	11	111111111	1111111111	10000000000~111111111111

**4. dct\_dc\_differential**

如果dct\_dc\_chrominance≠0，这个可选的可变长码将出现。表13-35中给出了其取值。

622

**5. dct\_coefficient\_first**

这个可选的可变长码字用于非内部编码块中的第一个DCT系数，其定义在表13-36、表13-37、表13-38和表13-39中给出了。

表13-36 MPEG-2 dct\_coefficient\_first 和dct\_coefficient\_next的可变长编码表零

RUN	LEVEL	编 码			
EOB		10			
0	1	1s	如果是第一个系数		
0	1	11s	不是第一个系数		
0	2	0100	s		
0	3	0010	1s		
0	4	0000	110s		
0	5	0010	0110	s	
0	6	0010	0001	s	
0	7	0000	0010	10s	
0	8	0000	0001	1101	s
0	9	0000	0001	1000	s
0	10	0000	0001	0011	s
0	11	0000	0001	0000	s
0	12	0000	0000	1101	0s
0	13	0000	0000	1100	1s
0	14	0000	0000	1100	0s
0	15	0000	0000	1011	1s
0	16	0000	0000	0111	11s
0	17	0000	0000	0111	10s
0	18	0000	0000	0111	01s
0	19	0000	0000	0111	00s
0	20	0000	0000	0110	11s
0	21	0000	0000	0110	10s

(续)

RUN	LEVEL	编 码			
0	22	0000	0000	0110	01s
0	23	0000	0000	0110	00s
0	24	0000	0000	0101	11s
0	25	0000	0000	0101	10s
0	26	0000	0000	0101	01s
0	27	0000	0000	0101	00s
0	28	0000	0000	0100	11s
0	29	0000	0000	0100	10s
0	30	0000	0000	0100	01s
0	31	0000	0000	0100	00s
0	32	0000	0000	0011	000s
0	33	0000	0000	0010	111s
0	34	0000	0000	0010	110s
0	35	0000	0000	0010	101s
0	36	0000	0000	0010	100s
0	37	0000	0000	0010	011s
0	38	0000	0000	0010	010s
0	39	0000	0000	0010	001s
0	40	0000	0000	0010	000s
1	1	011s			
1	2	0001	10s		
1	3	0010	0101	s	
1	4	0000	0011	00s	
1	5	0000	0001	1011	s
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
1	8	0000	0000	0011	111s
1	9	0000	0000	0011	110s
1	10	0000	0000	0011	101s
1	11	0000	0000	0011	100s
1	12	0000	0000	0011	011s
1	13	0000	0000	0011	010s
1	14	0000	0000	0011	001s
1	15	0000	0000	0001	0011s
1	16	0000	0000	0001	0010s
1	17	0000	0000	0001	0001s
1	18	0000	0000	0001	0000s
2	1	0101	s		
2	2	0000	100s		
2	3	0000	0010	11s	
2	4	0000	0001	0100	s
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0100	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0011	0s		

(续)

RUN	LEVEL	编 码			
4	2	0000	0011	11s	
4	3	0000	0001	0010	s
5	1	0011	11s		
5	2	0000	0010	01s	
5	3	0000	0000	1001	0s
6	1	0011	01s		
6	2	0000	0001	1110	s
6	3	0000	0000	0001	0100s
7	1	0001	00s		
7	2	0000	0001	0101	s
8	1	0000	111s		
8	2	0000	0001	0001	s
9	1	0000	101s		
9	2	0000	0000	1000	1s
10	1	0010	0111	s	
10	2	0000	0000	1000	0s
11	1	0010	0011	s	
11	2	0000	0000	0001	1010s
12	1	0010	0010	s	
12	2	0000	0000	0001	1001s
13	1	0010	0000	s	
13	2	0000	0000	0001	1000s
14	1	0000	0011	10s	
14	2	0000	0000	0001	0111s
15	1	0000	0011	01s	
15	2	0000	0000	0001	0110s
16	1	0000	0010	00s	
16	2	0000	0000	0001	0101s
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
27	1	0000	0000	0001	1111s
28	1	0000	0000	0001	1110s
29	1	0000	0000	0001	1101s
30	1	0000	0000	0001	1100s
31	1	0000	0000	0001	1011s
ESC		0000	01		

注: s表示非零系数 (level) 的符号, 0表示正, 1表示负。

表13-37 MPEG-2 dct\_coefficient\_first和dct\_coefficient\_next的可变长编码表一

RUN	LEVEL	编 码		
EOB		0110		
0	1	10s		
0	2	110s		
0	3	0111	s	
0	4	1110	0s	
0	5	1110	1s	
0	6	0001	01s	
0	7	0001	00s	
0	8	1111	011s	
0	9	1111	100s	
0	10	0010	0011	s
0	11	0010	0010	s
0	12	1111	1010	s
0	13	1111	1011	s
0	14	1111	1110	s
0	15	1111	1111	s
0	16	0000	0000	0111 11s
0	17	0000	0000	0111 10s
0	18	0000	0000	0111 01s
0	19	0000	0000	0111 00s
0	20	0000	0000	0110 11s
0	21	0000	0000	0110 10s
0	22	0000	0000	0110 01s
0	23	0000	0000	0110 00s
0	24	0000	0000	0101 11s
0	25	0000	0000	0101 10s
0	26	0000	0000	0101 01s
0	27	0000	0000	0101 00s
0	28	0000	0000	0100 11s
0	29	0000	0000	0100 10s
0	30	0000	0000	0100 01s
0	31	0000	0000	0100 00s
0	32	0000	0000	0011 000s
0	33	0000	0000	0010 111s
0	34	0000	0000	0010 110s
0	35	0000	0000	0010 101s
0	36	0000	0000	0010 100s
0	37	0000	0000	0010 011s
0	38	0000	0000	0010 010s
0	39	0000	0000	0010 001s
0	40	0000	0000	0010 000s
1	1	010s		
1	2	0011	0s	

(续)

RUN	LEVEL	编 码			
1	3	1111	001s		
1	4	0010	0111	s	
1	5	0010	0000	s	
1	6	0000	0000	1011	0s
1	7	0000	0000	1010	1s
1	8	0000	0000	0011	111s
1	9	0000	0000	0011	110s
1	10	0000	0000	0011	101s
1	11	0000	0000	0011	100s
1	12	0000	0000	0011	011s
1	13	0000	0000	0011	010s
1	14	0000	0000	0011	001s
1	15	0000	0000	0001	0011s
1	16	0000	0000	0001	0010s
1	17	0000	0000	0001	0001s
1	18	0000	0000	0001	0000s
2	1	0010	1s		
2	2	0000	111s		
2	3	1111	1100	s	
2	4	0000	0011	00s	
2	5	0000	0000	1010	0s
3	1	0011	1s		
3	2	0010	0110	s	
3	3	0000	0001	1100	s
3	4	0000	0000	1001	1s
4	1	0001	10s		
4	2	1111	1101	s	
4	3	0000	0001	0010	s
5	1	0001	11s		
5	2	0000	0010	0s	
5	3	0000	0000	1001	0s
6	1	0000	110s		
6	2	0000	0001	1110	s
6	3	0000	0000	0001	0100s
7	1	0000	100s		
7	2	0000	0001	0101	s
8	1	0000	101s		
8	2	0000	0001	0001	s
9	1	1111	000s		
9	2	0000	0000	1000	1s
10	1	1111	010s		
10	2	0000	0000	1000	0s
11	1	0010	0001	s	

(续)

RUN	LEVEL	编 码			
11	2	0000	0000	0001	1010s
12	1	0010	0101	s	
12	2	0000	0000	0001	1001s
13	1	0010	0100	s	
13	2	0000	0000	0001	1000s
14	1	0000	0010	1s	
14	2	0000	0000	0001	0111s
15	1	0000	0011	1s	
15	2	0000	0000	0001	0110s
16	1	0000	0011	01s	
16	2	0000	0000	0001	0101s
17	1	0000	0001	1111	s
18	1	0000	0001	1010	s
19	1	0000	0001	1001	s
20	1	0000	0001	0111	s
21	1	0000	0001	0110	s
22	1	0000	0000	1111	1s
23	1	0000	0000	1111	0s
24	1	0000	0000	1110	1s
25	1	0000	0000	1110	0s
26	1	0000	0000	1101	1s
27	1	0000	0000	0001	1111s
28	1	0000	0000	0001	1110s
29	1	0000	0000	0001	1101s
30	1	0000	0000	0001	1100s
31	1	0000	0000	0001	1011s
ESC		0000	01		

注: s表示非零系数 (level) 的符号, 0表示正, 1表示负。

表13-38 dct\_coefficient\_first和dct\_coefficient\_next的游程编码及其后面的转义码

游程编码	编 码	
0	0000	00
1	0000	01
2	0000	10
:	:	:
62	1111	10
63	1111	11

表13-39 dct\_coefficient\_first和dct\_coefficient\_next的LEVEL编码及其后面的转义码

LEVEL	编 码		
-2047	1000	0000	0001
-2046	1000	0000	0010
⋮	⋮	⋮	⋮
-1	1111	1111	1111
禁止	0000	0000	0000
1	0000	0000	0001
⋮	⋮	⋮	⋮
2047	0111	1111	1111

#### 6. dct\_coefficient\_next

最多63个可选的可变长码字, 仅对I、P和B帧出现。它们是第一个DCT系数之后的其他DCT系数, 其定义在表13-36、表13-37、表13-38和表13-39中给出了。

### 7. end\_of\_block

2位或者4位值，用于指示没有其他非零系数出现。这个参数的值为10或者0110。

## 13.5 运动补偿

图13-14说明了基本的运动补偿过程。运动补偿形成了从前面的解码图像得到的预测，这些预测结果依次与来自IDCT的系数数据（误差项）组合。

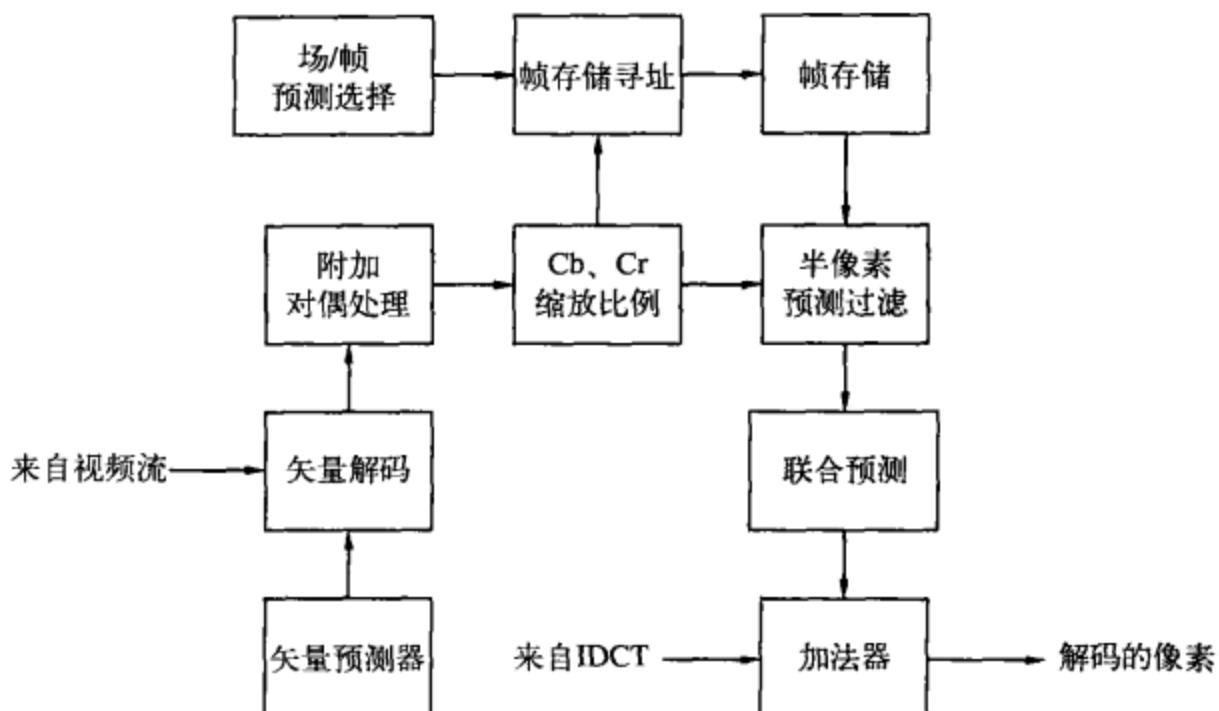


图13-14 简化的运动补偿过程

### 13.5.1 场预测

P图像的预测是由两个最近解码的参考场所构成的。图13-15给出了最简单的情况，当预测一帧的第一个图像或者在一帧内采用场预测时使用。

预测一帧的第二个场同样需要两个最近解码的参考场。图13-16和图13-17给出了这种情况的例子，两图中给出的第二个图像分别是底场和顶场。

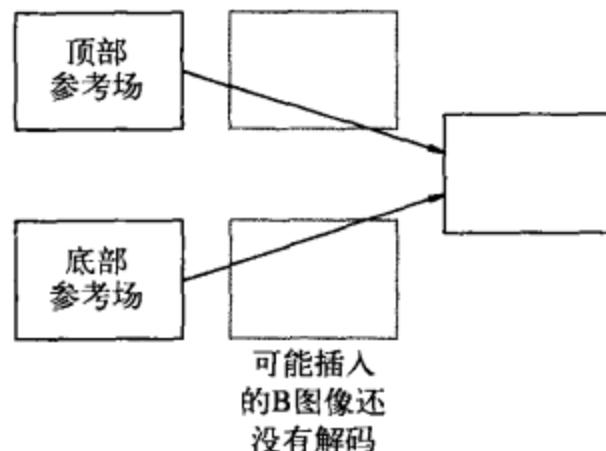


图13-15 第一场的P图像预测或者帧图像中的场预测

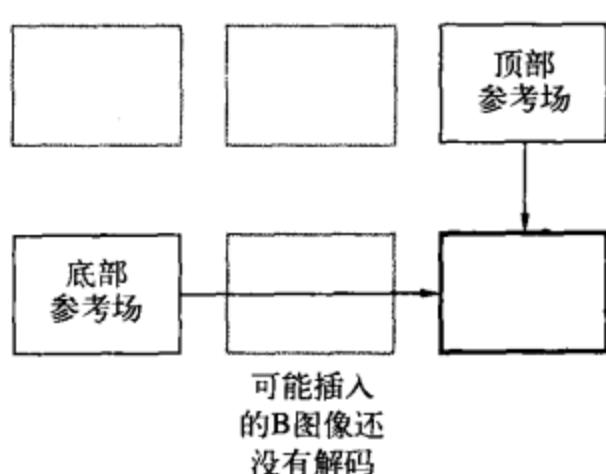


图13-16 第二场图像的P图像预测 (底场)

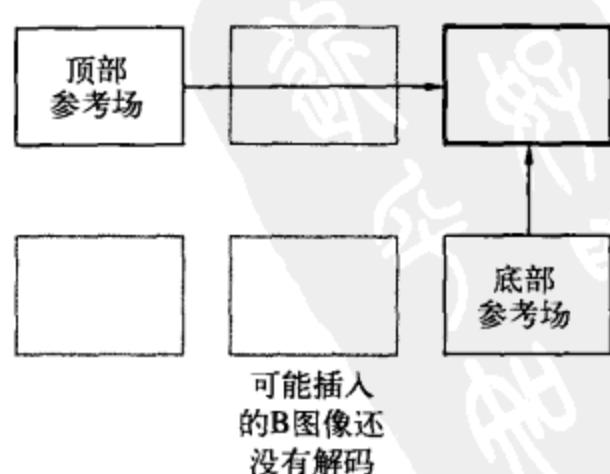


图13-17 第二场图像的P图像预测 (顶场)

B图像的场预测是由两个最近的参考帧的两个场所构成的，如图13-18所示。

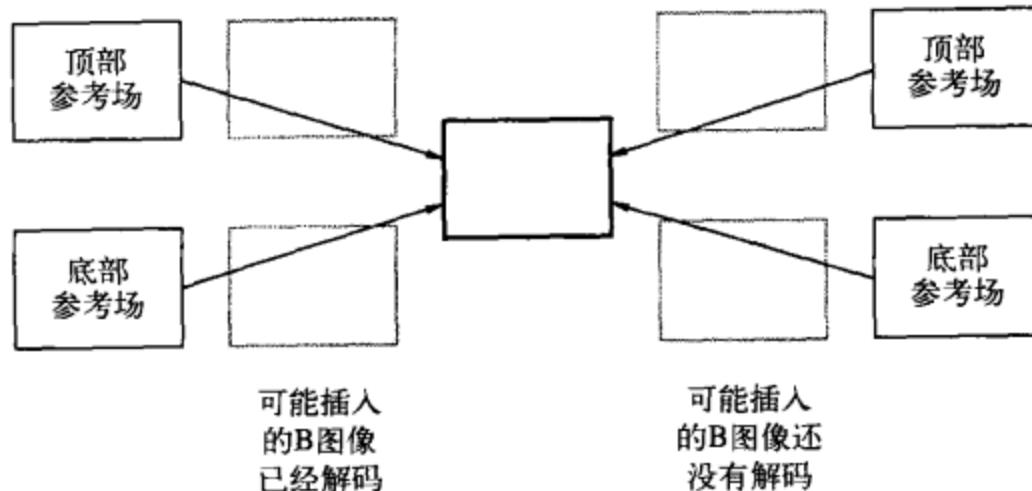


图13-18 B场或者帧图像的场预测

### 13.5.2 帧预测

P图像的预测是由两个最近的解码图像所构成的，如图13-19所示。参考的图像可能已经作为两个场或者一个单独的帧而解码了。

B图像的预测是由两个最近的参考帧所构成的，如图13-20所示。每一个参考帧可能已经作为两个场或者一个单独的帧而解码了。



图13-19 P图像的帧预测

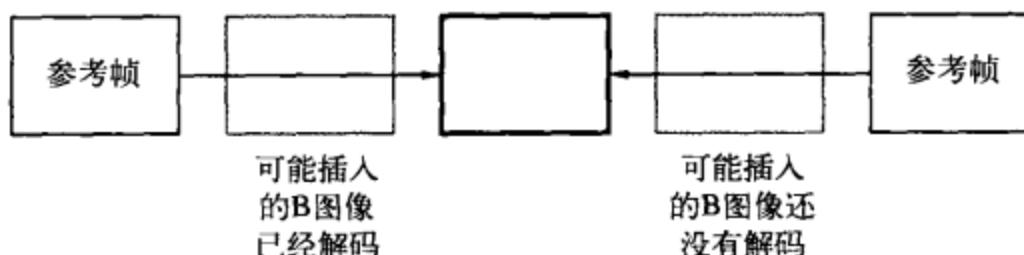


图13-20 B图像的帧预测

## 13.6 PES包

一个打包基本流（PES）由一个单独的基本流组成，该基本流已经被封装成数据包，每个数据包由一个附加的包头开始。一个PES包含来自一个源的唯一类型的数据（音频、视频等）。

图13-21给出了PES包的一般格式。注意，起始码（0x000001xx）必须是字节对齐，其实现方式是通过在起始码之前插入0~7个0。

### 1. packet\_start\_code\_prefix

长度为24位的字段，值为0x000001，与stream\_ID一起指示一个包的开始。

### 2. stream\_ID

长度为8位的码，定义了基本流的类型和数目，如表13-40所示。对于ATSC和OpenCable™标准，音频流的值必须是1011 1101，用来指示Dolby® Digital。

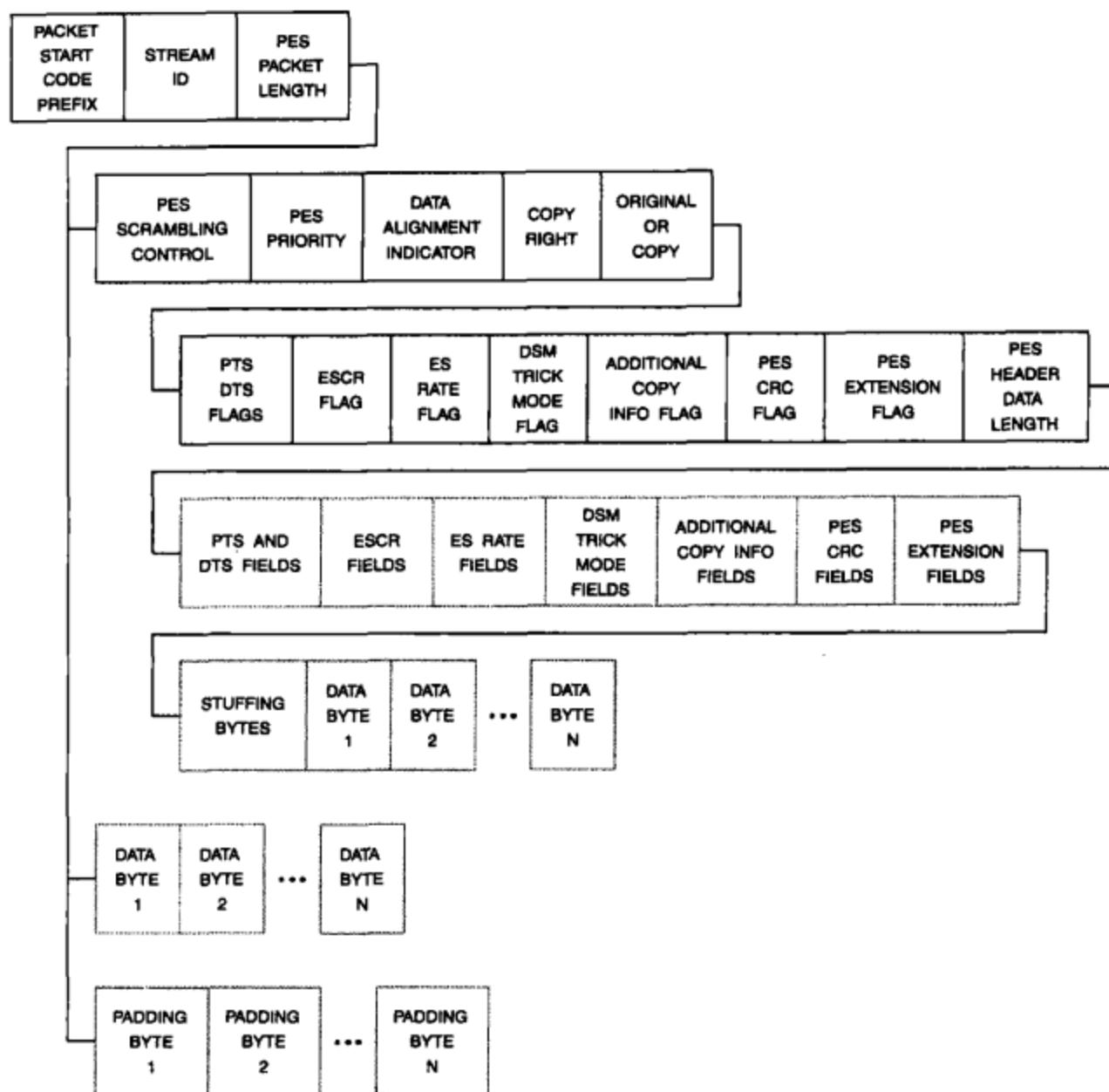


图13-21 MPEG-2 PES包结构。标记位和保留位没有给出

表13-40 MPEG-2 stream\_ID码字

流	编 码
所有音频流	1011 1000
所有视频流	1011 1001
节目流映射	1011 1100
私有流1（包括LPCM、Dolby Digital、Dolby Digital Plus、DTS、DTS-HD和MLP audio）	1011 1101
填充流	1011 1110
私有流2	1011 1111
MPEG-1.3、-2.3、-4.3或者MPEG-2.7音频流	110x xxxx
MPEG-1.2、-2.2、-4.2或者MPEG-4.10（H.264）视频流	1110 xxxx
ECM流	1111 0000
EMM流	1111 0001
DSM-CC流	1111 0010
ISO/IEC 13552流	1111 0011
ITU-T H.222.1类型A	1111 0100
ITU-T H.222.1类型B	1111 0101

(续)

流	编 码
ITU-T H.222.1类型C	1111 0110
ITU-T H.222.1类型D	1111 0111
ITU-T H.222.1类型E	1111 1000
辅助流	1111 1001
MPEG-4 SL打包流	1111 1010
MPEG-4 FlexMux流	1111 1011
元数据流	1111 1100
扩展流ID	1111 1101
保留	1111 1110
节目流目录PSD	1111 1111

**3. PES\_packet\_length**

16位的二进制数，定义了在这个字段之后PES包中的字节数。值为零表示既没有定义也不是有界的，且零值仅用于传输流。对于ATSC标准的视频流，其值必须是0x0000。

注意：如果stream\_ID=program stream map、padding stream、private stream 2、ECM stream、EMM stream、DSM-CC stream、H.222.1 type E或者program stream directory，下面的字段（直到下一个节点）将不会出现。

**4. marker\_bits**

可选的2位标记位，值为10。

**5. PES\_scrambling\_control**

可选的两位编码，定义了加扰模式（scrambling mode）。00表示不加扰，01表示保留，10表示每一个键都加扰，11表示奇数键加扰。对于SVCD、ATSC和OpenCable™标准，其值必须为00。

**6. PES\_priority**

可选位，定义了PES包载荷的优先级。1的优先级比0高。对于DVB标准，这个字段是可选的，如果出现了，解码器可以忽略掉。对于SVCD标准，其值必须为0。

**7. data\_alignment\_indicator**

可选位，值为1表示PES包头之后紧跟视频起始码，或者由数据流对齐描述符（如果有）定义的音频同步字。对于SVCD标准，其值必须为0。

表13-41 某些常用的MPEG-2 stream\_ID\_extension码字

流	编 码
IPMP控制信息流	000 0000
IPMP流	000 0001
保留数据流	000 0010
⋮	⋮
保留数据流	011 1111
私有流	100 0000
⋮	⋮
私有流	101 0100
SMPTE421M (VC-1) 视频流	101 0101
Dolby Digital、Dolby Digital Plus、DTS或DTS-HD核心音频流	111 0001
MLP或DTS-HD扩展音频流	111 0010

**8. copyright**

标志位，1表示素材是有版权的。对于SVCD，这个值必须为0。

**9. original\_or\_copy**

标志位，1表示该素材是原件；0表示该素材是一份副本。对于SVCD标准，这个值必须为1。

**10. PTS\_DTS\_flags**

两位标志位，值10表示一个PTS（显示时间戳）字段将出现在PES包头中。值11表示PTS和DTS（解码时间戳）字段将同时出现。值00表示PTS和DTS字段都不出现。对于SVCD标准，这个值必须为00、10或者11（11仅对视频）。

648  
649

**11. ESCR\_flag**

标志位，1表示ESCR（基本流时钟参考）基准和扩展字段将出现在PES包头中。对于SVCD、ATSC和OpenCable™标准，ESCR字段是可选的，如果出现了，解码器可以将其忽略。

**12. ES\_rate\_flag**

标志位，1表示ES\_rate（基本流速率）将出现在PES包头中。对于SVCD、ATSC和OpenCable™标准，这个值必须为0。对于DVB标准，ES字段是可选的，如果出现了，解码器可以将其忽略。

**13. DSM\_trick\_mode\_flag**

标志位，1表示trick\_mode\_flag字段将出现。对于SVCD标准，这个值必须为0。

**14. additional\_copy\_info\_flag**

标志位，1表示additional\_copy\_info字段将出现。对于SVCD标准，这个值必须为0。

**15. PES\_CRC\_flag**

标志位，1表示previous\_PES\_packet\_CRC字段将出现。对于SVCD、ATSC和OpenCable™标准，这个值必须为0。

**16. PES\_extension\_flag**

标志位，1表示一个扩展字段将出现在PES包头中。当传递SMPTE 421M（VC-1）视频流时，这一位必须为1，从而在PES包头中插入扩展。

**17. PES\_header\_data\_length**

8位的二进制数，定义了PES包头中的可选字段和填充位的字节数。

**18. marker\_bits**

可选的4位标记位，值为0010。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**19. PTS [32~30]**

可选字段，仅当PTS\_DTS\_flag = 10时才出现。

**20. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**21. PTS [29~15]**

可选字段，仅当PTS\_DTS\_flag = 10时才出现。

**22. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**23. PTS [14~0]**

可选的33位显示时间戳（PTS），指示解码器计划的显示时间。它所采用的单位是27MHz时钟除以300。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**24. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 10时这个字段才出现。

650

**25. marker\_bits**

可选的4位标记位，值为0011。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**26. PTS [32~30]**

可选字段，仅当PTS\_DTS\_flag = 11时才出现。

**27. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**28. PTS [29~15]**

可选字段，仅当PTS\_DTS\_flag = 11时才出现。

**29. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**30. PTS [14~0]**

可选字段，仅当PTS\_DTS\_flag = 11时才出现。

**31. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**32. marker\_bits**

可选的4位标记位，值为0001。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**33. DTS [32~30]**

可选字段，仅当PTS\_DTS\_flag = 11时才出现。

**34. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**35. DTS [29~15]**

可选字段，仅当PTS\_DTS\_flag = 11时才出现。

**36. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**37. DTS [14~0]**

可选的33位解码时间戳 (DTS)，指示计划的解码时间。它所采用的单位是27MHz时钟除以300。仅当PTS\_DTS\_flag = 11时这个字段才出现。

**38. marker\_bit**

可选的标记位，值为1。仅当PTS\_DTS\_flag = 10时这个字段才出现。

**39. reserved\_bits**

可选的两位保留位，值为11。仅当ESCR\_flag = 1时这个字段才出现。

**40. ESCR\_base [32~30]**

可选字段，仅当ESCR\_flag = 1时才出现。

**41. marker\_bit**

可选的标记位，值为1。仅当ESCR\_flag = 1时这个字段才出现。

**42. ESCR\_base [29~15]**

可选字段，仅当ESCR\_flag = 1时才出现。

**43. marker\_bit**

可选的标记位，值为1。仅当ESCR\_flag = 1时这个字段才出现。

**44. ESCR\_base [14~0]**

可选字段，仅当ESCR\_flag = 1时才出现。

**45. marker\_bit**

可选的标记位，值为1。仅当ESCR\_flag = 1时这个字段才出现。

**46. ESCR\_extension**

可选的9位基本流时钟参考（ESCR）扩展和33位ESCR基本信息组合成一个42位的值。它表示包含ESCR\_base最后一位的字节计划到达的时间。ESCR\_base的值定义了90kHz时钟周期的数目。

ESCR\_extension的值定义了在90kHz时钟周期启动之后27MHz时钟周期的数目。

这个字段仅当ESCR\_flag = 1时才出现。

**47. marker\_bit**

可选的标记位，值为1。仅当ESCR\_flag = 1时这个字段才出现。

**48. marker\_bit**

可选的标记位，值为1。仅当ESCR\_rate\_flag = 1时这个字段才出现。

**49. ES\_rate**

可选的22位基本流速率（ES\_rate），指示解码器接收PES包字节的速率。其定义的单位为50字节/秒。仅当ESCR\_rate\_flag = 1时这个字段才出现。

**50. marker\_bit**

可选的标记位，值为1。仅当ESCR\_rate\_flag = 1时这个字段才出现。

**51. trick\_mode\_control**

这个可选的3位码字指示哪一个窍门模式（trick mode）被应用到了视频流，如表13-42所示。仅当DSM\_trick\_mode\_flag = 1时这个字段才出现。

表13-42 MPEG-2 trick\_mode\_control码字

窍门模式	编 码	窍门模式	编 码
快进	000	慢退	100
慢进	001	保留	101
冻结帧	010	保留	110
快退	011	保留	111

**52. field\_ID**

可选的2位码字，指示哪一个场将被显示，如表13-43所示。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 000或者011时这个字段才出现。

**53. intra\_slice\_refresh**

可选的1位标志位，1表示条带之间可能有宏块丢失。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 000或者011时这个字段才出现。

652

**54. frequency\_truncation**

可选的2位码字，表示一个受限的系数集可能被用于数据编码中，如表13-44所示。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 000或者011时这个字段才出现。

表13-43 MPEG-2 field\_ID码字

场ID	编 码
仅顶场	00
仅底场	01
完全帧	10
保留	11

表13-44 MPEG-2 frequency\_truncation码字

描 述	编 码
仅DC系数是非零值	00
前3个系数是非零值	01
前6个系数是非零值	10
所有系数都可能是非零值	11

**55. rep\_ctrl**

可选的5位二进制数，表示每一个隔行场或者逐行帧应该显示的次数。值00000被禁用。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 001或者100时这个字段才出现。

**56. field\_ID**

表13-42中给出的可选的2位码字，表示将要显示哪一个场。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 010时这个字段才出现。

**57. reserved\_bits**

可选的3位保留位，值为111。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 010时这个字段才出现。

**58. reserved\_bits**

可选的5位保留位，值为1 1111。仅当DSM\_trick\_mode\_flag = 1且trick\_mode\_control = 101、110或者111时这个字段才出现。

**59. marker\_bit**

可选的标记位，值恒为1。仅当additional\_copy\_info\_flag = 1时这个字段才出现。

**60. additional\_copy\_info**

可选的7位字段，包含关于版权信息的私有数据。仅当additional\_copy\_info\_flag = 1时这个字段才出现。

**61. previous\_PES\_packet\_CRC**

可选的16位校验位，仅当PES\_CRC\_flag = 1时出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**62. PES\_private\_data\_flag**

可选的1位标志位，值为1时表示私有数据将出现。仅当PES\_extension\_flag = 1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。

**63. packet\_header\_field\_flag**

1位标志位，值为1时表示一个MPEG-1包头或者节目流包头出现在PES包头中。仅当PES\_extension\_flag = 1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。  
[653]

**64. program\_packet\_sequence\_counter\_flag**

可选的1位标志位，值为1时表示program\_packet\_sequence\_counter、MPEG1\_MPEG2\_identifier和original\_stuff\_length等字段将出现。仅当PES\_extension\_flag = 1时这个字段才出现。对于SVCD、ATSC和OpenCable™标准，其值必须为0。

**65. P-STD\_buffer\_flag**

可选的1位标志位，值为1时表示P-STD\_buffer\_scale和P-STD\_buffer\_size将出现。仅当PES\_extension\_flag = 1时这个字段才出现。对于ATSC和OpenCable™标准，其值必须为0。对于SVCD标准，其值必须为1。

**66. reserved\_bits**

可选的3位保留位，值恒为111。仅当PES\_extension\_flag = 1时这个字段才出现。

**67. PES\_extension\_flag\_2**

可选的1位标志位，值为1时表示PES\_extension\_field\_length及其相关字段将出现。仅当PES\_extension\_flag = 1时这个字段才出现。对于SVCD标准，其值必须为0。当传输SMPTE 421M(VC-1)视频流时，这一位必须为1，从而使得能够将第二个扩展组插入到PES包头中。

**68. PES\_private\_data**

可选的128位私有数据，与其前后的字段组合在一起，不能与packet\_start\_code\_prefix相似。仅当PES\_extension\_flag = 1并且PES\_private\_data\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**69. pack\_field\_length**

可选的8位二进制数，表示紧跟在包组头之后的字段的长度，单位是字节。仅当PES\_extension\_flag = 1并且pack\_header\_field\_flag = 1时这个字段以及紧跟在包组头之后的字段才出现。

**70. marker\_bit**

可选的标记位，值恒为1。仅当PES\_extension\_flag = 1并且program\_packet\_sequence\_counter\_flag = 1时这个字段才出现。

**71. program\_packet\_sequence\_counter**

可选的7位二进制数，对于节目流或者MPEG-1系统流中每一个连续的PES包，其值增加1。当达到最大值后，其值反转到0。对于两个连续的PES包，其值不能相同。仅当PES\_extension\_flag = 1并且program\_packet\_sequence\_counter\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**72. marker\_bit**

可选的标记位，值恒为1。仅当PES\_extension\_flag = 1并且program\_packet\_sequence\_counter\_flag = 1时这个字段才出现。

**73. MPEG1\_MPEG2\_identifier**

可选标志位，值为1时表示PES包有来自MPEG-1系统流的信息。值为0表示PES包有来自节目流的信息。仅当PES\_extension\_flag = 1并且program\_packet\_sequence\_counter\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**74. original\_stuff\_length**

可选的6位二进制数，指定了用于初始PES或者MPEG-1包头中填充的字节数。仅当PES\_extension\_flag = 1并且program\_packet\_sequence\_counter\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

654

**75. marker\_bits**

可选的2位标记位，值恒为01。仅当PES\_extension\_flag = 1并且P-STD\_buffer\_flag = 1时这个字段才出现。

**76. P-STD\_buffer\_scale**

可选位，表示接下来的P-STD\_buffer\_size参数的比例因子。对于音频流，其值为0。对于视频流，其值为1。对于所有其他类型的流，其值既可为0，也可为1。仅当PES\_extension\_flag = 1并且P-STD\_buffer\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**77. P-STD\_buffer\_size**

可选的13位二进制数，指定了解码器输入缓冲区的大小。如果P-STD\_buffer\_scale值为0，该字段为128字节。如果P-STD\_buffer\_scale值为1，该字段为1024字节。仅当PES\_extension\_flag = 1并且P-STD\_buffer\_flag = 1时这个字段才出现。对于DVB标准，这个字段是可选的，如果出现了，解码器可以将其忽略掉。

**78. marker\_bit**

可选的标记位，值恒为1。仅当PES\_extension\_flag = 1并且PES\_extension\_flag\_2 = 1时这个字段

才出现。

#### 79. PES\_extension\_field\_length

可选的7位二进制数，表示下面三个字段总的字节数。仅当PES\_extension\_flag = 1并且PES\_extension\_flag\_2 = 1时这个字段才出现。

#### 80. Stream\_ID\_extension\_flag

可选的1位标志位，值为0时表示接下来是一个stream\_ID\_extension字段。仅当PES\_extension\_flag = 1并且PES\_extension\_flag\_2 = 1时这个字段才出现。当传输SMPTE 421M (VC-1) 视频流时，这一位必须为0，从而使得能够插入一个有效的stream\_id\_extension。

#### 81. stream\_ID\_extension

7位码字，作为stream\_ID的扩展使用，用来指定表13-41中所定义的基本流类型。仅当stream\_ID = 1111 1101时这个字段才被采用。仅当PES\_extension\_flag和PES\_extension\_flag\_2 = 1并且stream\_ID\_extension\_flag = 0时这个可选的字段才出现。

#### 82. reserved\_byte

[n]字节保留数据，值为1111 1111。仅当PES\_extension\_flag和PES\_extension\_flag\_2 = 1并且stream\_ID\_extension\_flag = 0时这个可选的字段才出现。

#### 83. stuffing\_byte

可选的[n]字节填充数据，值为1111 1111。最多可以采用32字节填充数据。这些数据将被解码器忽略。

655

#### 84. PES\_packet\_data\_byte

来自音频流、视频流、私有流1、辅助流、H.222.1类型A-D流或者ISO/IEC 13552流的[n]字节数据。字节数由PES\_packet\_length字段推导而得到。

注意：当stream\_ID = 节目流映射、私有流2、ECM流、EMM流、DSM-CC流、H.2221类型E或者节目流目录时，下面的字段将出现。

#### 85. PES\_packet\_data\_byte

来自节目流映射、私有流2、ECM流、EMM流、DSM-CC流、H.2221类型E流或者节目流目录描述符的[n]字节数据。字节数由PES\_packet\_length字段推导而得到。

注意：当stream\_ID = padding stream时，下面的字段将出现。

#### 86. padding\_byte

[n]字节，值为1111 1111。字节数由PES\_packet\_length字段指定。这一字段将被解码器忽略。

## 13.7 节目流

节目流是设计用于相对无差错的环境中的，它被DVD和SVCD标准所采用。节目流由一个或多个PES包复用到一起所组成，并且与数据一起编码，从而使得它们能够同时解码。节目流包可以是可变长的，且可以相对非常长。

图13-22给出了节目流的一般格式。注意，起始码（0x0000001xx）必须是字节对齐的，可以通过在起始码之前插入0~7个0位来实现。

656

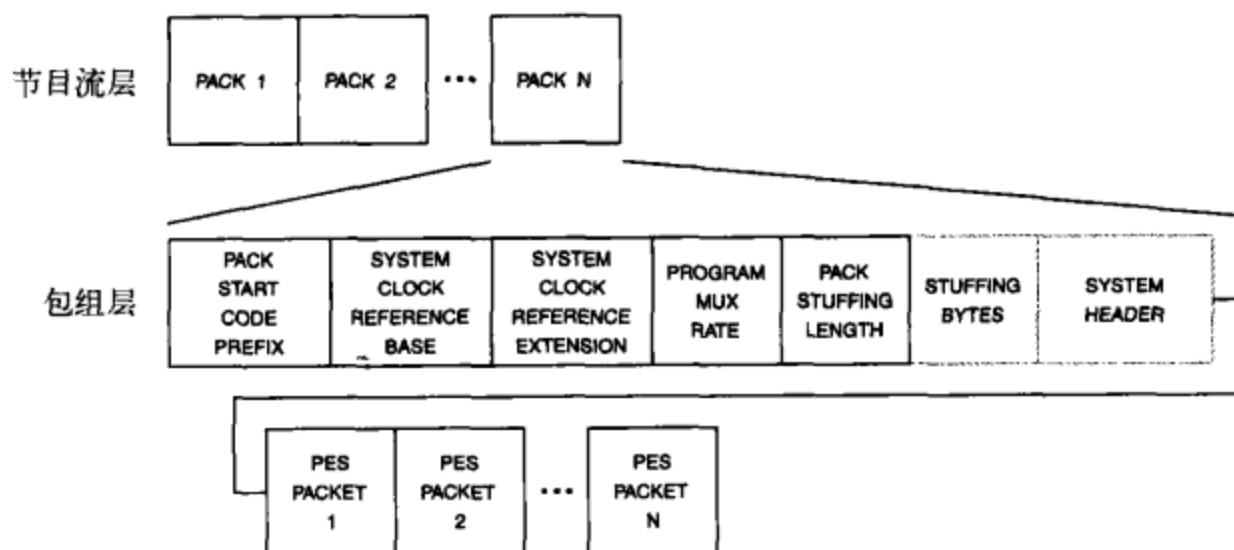


图13-22 MPEG-2节目流结构。没有给出标记和保留位

### 13.7.1 包组层

每个包组的数据由一个包组头跟一个可选的系统头和一个或多个PES包组成。

#### 1. pack\_start\_code

32位字符串，值为0x000001BA，指示一个包组的开始。

#### 2. marker\_bits

2位标记位，值为01。

#### 3. system\_clock\_reference\_base [32~30]

#### 4. marker\_bit

标记位，值为1。

#### 5. system\_clock\_reference\_base [29~15]

#### 6. marker\_bit

标记位，值为1。

#### 7. system\_clock\_reference\_base [14~0]

#### 8. marker\_bit

标记位，值为1。

#### 9. system\_clock\_reference\_extension

9位字段，与system\_clock\_reference\_base字段一起组成系统时钟参考（SCR）。system\_clock\_reference\_base字段采用的单位是90kHz乘以1/300。system\_clock\_reference\_extension字段采用单位是27MHz。SCR指示包含system\_clock\_reference\_base字段最后一位的字节的预期达到时间。

#### 10. marker\_bit

标记位，值为1。

#### 11. program\_mux\_rate

22位二进制数，定义了一个以每秒50字节为度量的值，该值禁止取零。它指示解码器接收节目流的速度。对于SVCD标准，该值必须小于等于6972<sub>D</sub>。

#### 12. marker\_bit

标记位，值为1。

#### 13. marker\_bit

标记位，值为1。

**14. reserved\_bits**

5位保留位，值为1 1111。

**15. pack\_stuffing\_length**

3位二进制数，定义了这一字段之后stuffing\_byte字段的个数。

**16. stuffing\_byte**

可能出现0~7个填充字节。解码器忽略这些字节。每个字节的值为1111 1111。

### 13.7.2 系统头

该字段包含比特流参数的梗概。在第一个包组头后必然会跟随一个系统头，但在未来的包组头中它可能会可选地重复出现。

**1. system\_header\_start\_code**

657 32位字符串，值为0x000001BB，指示一个系统头的开始。

**2. header\_length**

16位二进制数，定义了这个字段之后系统头的字节数。

**3. marker\_bit**

标记位，值为1。

**4. rate\_bound**

22位二进制数，定义了一个大于或等于program\_mux\_rate最大值的值。对于SVCD标准，该值必须为6972<sub>D</sub>。

**5. marker\_bit**

标记位，值为1。

**6. audio\_bound**

6位二进制数，定义了一个大于或等于有效音频流最大个数的值（0~32）。对于SVCD标准，该值必须为0。

**7. fixed\_flag**

1位标志位，值为1时表示固定的位速操作。值为0时表示可变的位速操作。对于SVCD标准，该值必须为0。

**8. CSPS\_flag**

1位标志位。值为1表示节目流是“受约束的系统参数流”。

**9. system\_audio\_lock\_flag**

1位标志位。如果在音频采样速率和解码器系统时钟频率之间存在已定义的、稳定的关系，该字段值为1。对于SVCD标准，该值必须为1。

**10. system\_video\_lock\_flag**

1位标志位。如果在视频图像速率和解码器系统时钟频率之间存在已定义的、稳定的关系，该位值为1。对于SVCD标准，该值必须为1。

**11. marker\_bit**

标记位，值为1。

**12. video\_bound**

5位二进制数，定义了一个大于或等于有效视频流最大个数的值（0~16）。对于SVCD标准，该值必须为0或1。

**13. reserved\_bits**

7位保留位，值为111 1111。

**14. stream\_ID**

可选的8位编码，如表13-40所示，这一字段指示应用于PSTD\_buffer\_bound\_scale 和 PSTD\_buffer\_size\_bound字段的流。

**15. marker\_bits**

可选的2位标记位，值为11。仅当stream\_ID出现时这个字段才出现。

**16. P-STD\_buffer\_bound\_scale**

仅当stream\_ID出现时这个可选位才出现，指示用于P-STD\_buffer\_size\_bound的缩放比例因子。值为0表示stream\_ID定义一个音频流。值为1表示stream\_ID定义一个视频流。对于其他类型的流ID，该值要么为0，要么为1。658

**17. P-STD\_buffer\_size\_bound**

可选的13位二进制数，定义了一个大于或等于解码器输入缓冲最大容量的值。仅当stream\_ID 出现时这个字段才出现。如果P-STD\_buffer\_bound\_scale的值为0，该字段包含128字节。如果P-STD\_buffer\_bound\_scale的值为1，该字段包含1024字节。

### 13.7.3 节目流映射

节目流映射（Program Stream Map，PSM）提供节目流中的比特流以及比特流相互之间关系的描述。如果stream\_ID = program stream map，那么PSM以PES包数据的形式出现。

**1. packet\_start\_code\_prefix**

24位字符串，值为0x000001，指示一个流映射的开始。

**2. map\_stream\_ID**

8位字符串，值为1011 1100。

**3. program\_stream\_map\_length**

16位二进制数，指示这个字段之后的字节数。最大值为1018<sub>D</sub>。

**4. current\_next\_indicator**

标志位。值为1表示节目流映射目前可用，值为0表示节目流映射还不可用，但是下一个将有效。

**5. reserved\_bits**

2位保留位，值为11。

**6. program\_stream\_map\_version**

5位二进制数，定义了节目流映射的版本号。当节目流映射变化时这一字段必须加一，且其值一旦达到31就重新归零。

**7. reserved\_bits**

7位保留位，值为111 1111。

**8. marker\_bit**

标记位，值为1。

**9. program\_stream\_info\_length**

16位二进制数，定义了紧跟在该字段之后的描述符总长（单位是字节）。

**10. descriptor\_loop**

deescriptor\_loop中可能出现不同类型的描述符。

**11. elementary\_stream\_map\_length**

16位二进制数，指示节目流映射中所有基本流信息的字节数。

注意：对于每个具有唯一stream\_type值的流，下列4个字段出现。

#### 12. stream\_type

8位码字，定义了表13-45给出的流类型。

#### 13. elementary\_stream\_ID

8位字段，定义了stream\_ID的值，如表13-40所示，该字段位于包含比特流的PES包的PES包

659 头中。

表13-45 常用的stream\_type码字。[www.atsc.org](http://www.atsc.org)网站的Code Point Registry 提供stream\_type编码的完整列表

流类型	编 码	流类型	编 码
保留	0000 0000	MPEG-4.2视觉	0001 0000
MPEG-1.2视频	0000 0001	MPEG-4.3音频	0001 0001
MPEG-2.2视频	0000 0010	由PES包传送的MPEG-4 SL 打包流或FlexMux流	0001 0010
MPEG-1.3音频	0000 0011	由MPEG-4分段传送的MPEG-4 SL 打包流或FlexMux流	0001 0011
MPEG-2.3音频	0000 0100	MPEG-2.6同步下载协议	0001 0100
MPEG-2私有分段	0000 0101	由PES包传送的元数据	0001 0101
包含私有数据的MPEG-2 PES包	0000 0110	由元数据分段传送的元数据	0001 0110
ISO/IEC 13552 MHEG	0000 0111	由MPEG-2.6数据轮播传送的元数据	0001 0111
MPEG-2 DSM CC	0000 1000	由MPEG-2.6对象轮播传送的元数据	0001 1000
MPEG-1、MPEG-2辅助流	0000 1001	由MPEG-2.6同步下载协议传送的元数据	0001 1001
MPEG-2.6类型A	0000 1010	MPEG-2.11 IPMP流	0001 1010
MPEG-2.6类型B	0000 1011	MPEG-4.10视频	0001 1011
MPEG-2.6类型C	0000 1100	保留	0001 1100~0111 1110
MPEG-2.6类型D	0000 1101	IPMP流	0111 1111
MPEG-2辅助流	0000 1110	用户私有	1000 0000~1111 1111
MPEG-2.7音频	0000 1111		

#### 某些常用的用户私有细节

DigiCipher II视频	1000 0000	ATSC数据服务表、网络资源表	1001 0101
Dolby Digital音频	1000 0001	SCTE IP数据	1010 0000
SCTE标准字幕	1000 0010	ATSC同步数据流或SCTE同步数据	1100 0010
SCTE同步数据	1000 0011	SCTE异步数据	1100 0011
ATSC节目标识符	1000 0101		
Dolby Digital Plus音频	1000 0111		

#### 14. elementary\_stream\_info\_length

16位二进制数，定义了紧跟在该字段之后的描述符总长（单位是字节）。

#### 15. descriptor\_loop

descriptor\_loop中可能出现不同类型的描述符。

#### 16. CRC\_32

32位CRC，用于整个节目流映射。

### 13.7.4 节目流目录

节目流目录（Program Stream Directory, PSD）提供关于节目流中的比特流以及比特流相互之

660

间关系的描述。如果stream\_ID = program stream map, PSDI以PES包数据的形式出现。

## 13.8 传输流

传输流设计用于易出错的环境中，如远距离传输或嘈杂环境中，它被ARIB、ATSC、DVB、数字有线电视以及OpenCable™标准所采用。

传输流将一个或多个节目以及一个或多个独立时间基组合成一个单一流。每个传输流中的节目都能拥有属于自己的时间基。同一传输流中不同的节目，其时间基可能各不相同。

传输流由一个或多个188B的包组成。每个包的数据来自于PES包、PSI（节目相关信息）分段、填充字节或私有数据。除了MPEG-2数据之外，MPEG-4.2、MPEG-4.10（H.264）、SMPTE 421M（VC-1）以及其他类型数据都可采用MPEG-2传输流传送。

在每个包的开始是一个包标识符（PID），PID能够使解码器确定如何处理该数据包。如果MPEG数据采用“单载波多通道”传送，解码器通过PID确定哪些包是当前观察或记录通道的一部分，从而处理这些包，而丢弃其他包。系统信息（SI）（如节目指南、通道频率等）同样被分配唯一的PID值。

图13-23给出了传输流的一般格式。注意，起始码（0x000001xx）必须是字节对齐的，可以通过在起始码之前插入0~7个0位来实现。

### 13.8.1 包层

每个包由一个包头跟一个可选的适配域和一个或多个数据包组成。

#### 1. sync\_byte

8位字符串，值为0100 0111。

#### 2. transport\_error\_indicator

标志位。值为1表示包中出现至少一个不可纠正的比特错误。

#### 3. payload\_unit\_start\_indicator

该标志位的含义取决于有效载荷。

对于PES包数据，该字段值为1表示包的数据块以PES包的字节开始。值为0表示包的数据块中没有以PES包开始的。

对于PSI数据，该字段值为1表示包的数据块包含PSI分段的第一个字节。

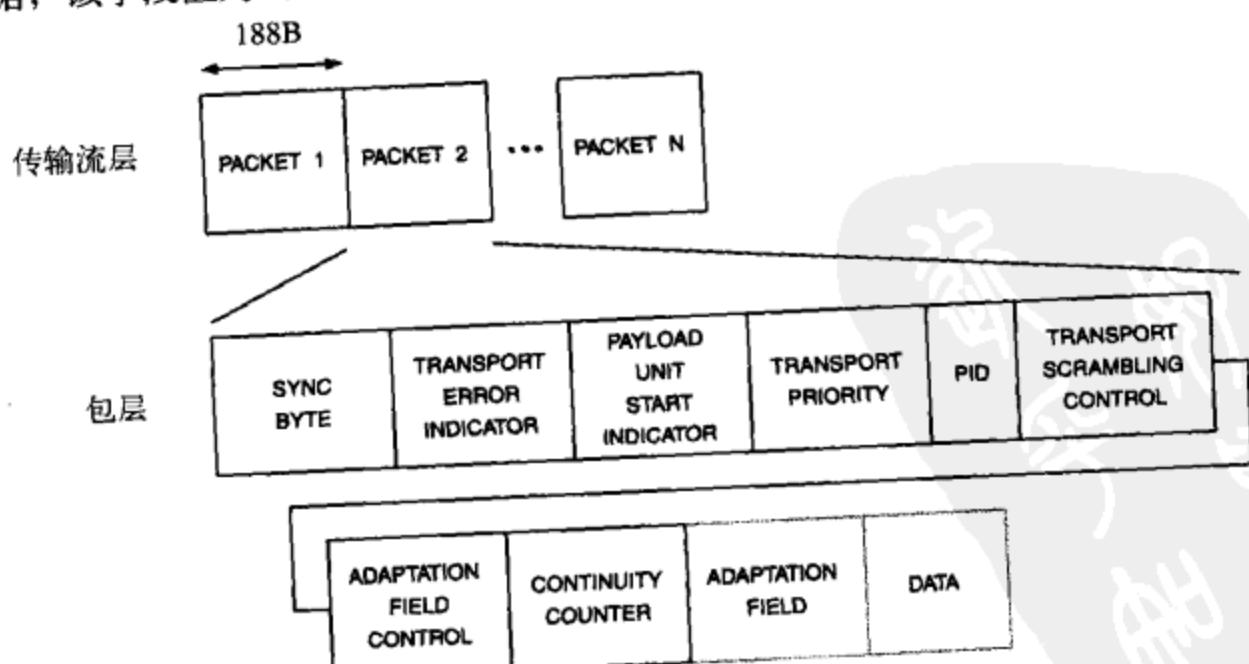


图13-23 MPEG-2传输流结构。没有给出标记和保留位。某些应用增加一个4字节TP\_extra\_header前缀（产生192B），由到达时间戳和复制允许信息组成

**4. transport\_priority**

标志位。值为1表示与其他具有相同PID的包相比，这个包拥有更高的优先级。对于DVB标准，该字段是可选的，而且即使出现也可被解码器忽略。

**5. PID**

13位码字，指示数据块中的数据类型，如表13-46所示。[www.atsc.org](http://www.atsc.org)网站的Code Point Registry中提供PID编码的完全列表。

**6. transport\_scrambling\_control**

2位编码，指示有效载荷的加扰模式。00=不加扰，01=不加扰（私有用途），10=偶数健加扰，11=奇数键加扰。当该字段取00之外的值时，要求在基本码流中出现CA解码器。

表13-46 常用PID码字

描    述	编    码
节目关联表	0 0000 0000 0000
条件接收表	0 0000 0000 0001
传输流描述表	0 0000 0000 0010
IPMP控制信息表	0 0000 0000 0011
MPEG-2保留	0 0000 0000 0100~0 0000 0000 1111
由DVB使用	0 0000 0001 0000~0 0000 0001 1111
由ARIB使用	0 0000 0010 0000~0 0000 0010 1111
由ATSC、CEA和SCTE使用	1 1111 1111 0111~1 1111 1111 1110
MPEG-2空包	1 1111 1111 1111

662

**7. adaptation\_field\_control**

2位编码，指示包头之后出现适配域还是数据块，如表13-47所示。

**8. continuity\_counter**

4位二进制数，随PID相同的包数目增加而增加。达到最大值后，它归零重新计数。当包中不存在数据块时该字段值不会增大。

**9. adaptation\_field**

参见13.8.2节。

**10. data\_byte**

[n]数据字节，是来自于PES包、PSI分段、填充字节或私有数据的毗邻数据字节。[n]等于184减去适配域（如果出现）中的数据字节个数。当adaptation\_field\_control=01或11时该字段才出现。

### 13.8.2 适配域

这个域包含42位节目时钟基准（PCR）、随机访问指示符、接合倒计时以及其他MPEG和私有信息。在解码器中，PCR用来重新生成27MHz时间基时钟，该时钟与编码器中使用的相同。这个时钟也是显示时间戳（PTS）的基础。在流中PCR一般每0.1秒出现一次。当adaptation\_field\_control=01或11时该域才出现。

**1. adaptation\_field\_length**

8位二进制数，定义了该字段后紧随出现的字节个数。值0000 0000用来在传输流包中插入单个的填充字节。当adaptation\_field\_control=11时，该字段的值范围为0~182<sub>D</sub>。当adaptation\_

表13-47 MPEG-2 adaptation\_field\_control码字

描    述	编    码
保留	00
仅数据	01
仅适配域	10
适配域跟数据	11

field\_control = 10时，该字段的值为183D。

注意：如果adaptation\_field\_length = 0000 0000，以下字段不会出现。

#### 2. discontinuity\_indicator

1位标志位。值为1表示当前传输流包中存在不连续状态。

#### 3. random\_access\_indicator

1位标志位，用来指示当前传输流包以及后续可能出现的具有相同PID的传输流包是否包含帮助随机访问的某些信息。

#### 4. elementary\_stream\_priority\_indicator

1位标志位。在具有相同PID的包中，该字段用来指示传输流包中携带基本码流数据的优先级。值为1表示该有效载荷的优先级高于其他传输流包的有效载荷。值为0表示该有效载荷与其他使这个标志位值不为1的包具有相同的优先级。

#### 5. PCR\_flag

1位标志位。值为1表示适配域包含PCR字段。值为0表示适配域不包含PCR字段。

#### 6. OPCR\_flag

可选的1位标志位。值为1表示适配域包含OPCR字段。值为0表示适配域不包含任何OPCR字段。

#### 7. splicing\_point\_flag

1位标志位。值为1表示splice\_countdown字段在关联的适配域中出现，指定了接合点的出现。

#### 8. transport\_private\_data\_flag

1位标志位。值为1表示适配域包含一个或多个私有字节。

#### 9. adaptation\_field\_extension\_flag

1位标志位。值为1表示存在适配域的扩展。

#### 10. program\_clock\_reference\_base

可选42位program\_clock\_reference字段的33位LSB。当PCR\_flag = 1时该字段才出现。

#### 11. reserved\_bits

6位可选保留位，值为111111。当PCR\_flag = 1时该字段才出现。

#### 12. program\_clock\_reference\_extension

9位可选字段，与program\_clock\_reference\_base字段一起组成42位节目时钟基准（PCR）。PCR指示字节预期到达的时间，该字节包含作为解码器输入的program\_clock\_reference\_base字段的最后一一位。当PCR\_flag = 1时该字段才出现。

#### 13. original\_program\_clock\_reference\_base

可选42位original\_program\_clock\_reference字段的33位LSB。当OPCR\_flag = 1时该字段才出现。

#### 14. reserved\_bits

6位可选保留位，值为111111。当OPCR\_flag = 1时该字段才出现。

#### 15. original\_program\_clock\_reference\_extension

9位可选字段，与original\_program\_clock\_reference\_base字段一起组成42位原始节目时钟基准（OPCR）。OPCR仅在包含PCR的传输流包中出现。OPCR支持单一节目传输流从另一传输流中重建。当OPCR\_flag = 1时该字段才出现。

#### 16. splice\_countdown

2的反码，长度为8位，代表一个可正可负的值。正值定义了在接合点之前与关联传输流包具

664

有相同PID值的传输流包个数。负值指示在接合点之后的关联传输流包是第n个包。当 splicing\_point\_flag = 1时这个字段才出现。

#### 17. transport\_private\_data\_length

8位可选的二进制数[n]，定义了这个字段后紧接着出现的字节数。当transport\_private\_data\_flag = 1时该字段才出现。

#### 18. private\_data\_byte

可选的[n]数据字节，MPEG-2没有定义该字段。当transport\_private\_data\_flag = 1时该字段才出现。

#### 19. adaptation\_field\_extension\_length

可选的8位二进制数，指示这个字段后紧接着出现的扩展适配域数据的字节数，如果出现，还包括保留字节。当adaptation\_field\_extension\_flag = 1时该字段才出现。

#### 20. ltw\_flag

1位可选标志位（合法的时间窗口标志），值为1指示ltw\_offset字段的出现。当adaptation\_field\_extension\_flag = 1时该字段才出现。

#### 21. piecewise\_rate\_flag

1位可选标志位，值为1指示piece\_rate字段的出现。当adaptation\_field\_extension\_flag = 1时该字段才出现。

#### 22. seamless\_splice\_flag

1位可选的标志位，值为1指示splice\_type和DTS\_next\_AU字段的出现。当adaptation\_field\_extension\_flag = 1时该字段才出现。

#### 23. reserved\_bits

5位可选保留位，值为11111。当adaptation\_field\_extension\_flag = 1时该字段才出现。

#### 24. ltw\_valid\_flag

1位可选标志位，指示ltw\_offset的值是否有效。当adaptation\_field\_extension\_flag = 1以及ltw\_flag = 1时该字段才出现。

#### 25. ltw\_offset

15位可选的二进制数，定义了单位为 $(300/f_s)$ 秒的合法时间窗口偏移，这里 $f_s$ 是该PID所属节目的系统时钟频率。当adaptation\_field\_extension\_flag = 1以及ltw\_flag = 1时该字段才出现。

#### 26. reserved\_bits

2位可选保留位，值为11。当adaptation\_field\_extension\_flag = 1以及piecewise\_rate\_flag = 1时该字段才出现。

#### 27. piecewise\_rate

可选的22位二进制数，定义了一个假定的比特率，该比特率用来界定具有相同PID值的传输流包中合法时间窗口的结束时间，这些传输流包在当前包之后出现且不包含legal\_time\_window\_offset字段。当adaptation\_field\_extension\_flag = 1以及piecewise\_rate\_flag = 1时该字段才出现。

#### 28. splice\_type

4位可选的二进制数，在它出现的所有后续具有相同PID值的传输流包中，该字段具有相同的值，直到包中的splice\_countdown字段值归零为止（包括当前包）。如果带入该PID的基本流是音频流，该字段的值将为0000。如果带入该PID的基本流是视频流，该字段指示将引起用于接合的基本流注意的状态。当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时该字段才出现。

665

**29. DTS\_next\_AU[32...30]**

可选的33位字段，指示接合点后第一个访问单元的解码时间。当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时该字段才出现。

**30. marker\_bit**

可选的1位标记位。值恒为1。当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时该字段才出现。

**31. DTS\_next\_AU[29...15]**

可选字段，当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时才出现。

**32. marker\_bit**

可选的1位标记位。值恒为1。当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时该字段才出现。

**33. DTS\_next\_AU[14...0]**

可选字段，当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时才出现。

**34. marker\_bit**

可选的1位标记位。值恒为1。当adaptation\_field\_extension\_flag = 1以及seamless\_splice\_flag = 1时该字段才出现。

**35. reserved\_bits**

可选的[n]数据字节，值为1111 1111。当ltw\_flag、piecewise\_rate\_flag以及seamless\_splice\_flag = 0，而adaptation\_field\_extension\_flag = 1时，该字段才出现。

**36. stuffing\_byte**

可选的[n]数据字节，值为1111 1111。当OPCR\_flag、adaptation\_field\_extension\_flag、PCR\_flag、transport\_private\_data\_flag以及splicing\_point\_flag = 0时该字段才出现。

### 13.8.3 节目相关信息

节目相关信息（PSI）是一类附加信息，使解码器能够更加高效地在单一传输流中找出所需的内容，并集成为用户友好的电子节目指南（EPG）。

节目由一个或多个的打包音频、视频、数据等基本流组成，每个流都分配一个13位的包标识号（PID）。此外，携带相同PES的传输流包分配相同但唯一的PID。

只有当获得正确的PID后MPEG-2解码器才能处理正确的包。这是PSI的功能。传输流包中携带的PSI具有唯一的PID，因而MPEG-2解码器能够方便地找到它。

PSI包括6个表，如表13-48所示。节目关联表（PAT）、条件接收表（CAT）、传输流描述表（TSDT）、IPMP控制信息表（ICIP）以及空包是仅有的固定的PID。MPEG-2解码器通过访问适当的表来确定剩余的PID。

如果首次接收传输流，MEPG-2解码器将查找PAT、CAT、TSDT和ICIT。如图13-24所示，从PIA中读出网络信息表（NIT）的PID和每个PMT。再从PMT中读出每个基本流的PID。如果节目是加密的，还要求访问CAT。

表13-48 节目相关信息表

流类型	缩写	PID	描述
节目关联表	PAT	0x0000	关联节目编号与节目映射表PID
条件接收表	CAT	0x0001	关联一个或多个（私有）EMM流，每一个都有一个唯一的PID值
传输流描述表	TSDT	0x0002	关联一个或多个描述符到一个完整的传输流

(续)

流类型	缩写	PID	描述
IPMP控制信息表	ICIT	0x0003	包含MPEG-2.11中定义的IPMP工具列表、权利箱、工具箱
节目映射表	PMT	由PAT分配	为一个或多个节目的组件定义PID值
网络信息表	NIT	由PAT分配*	物理网络参数,如FDM频率、转发器数量等

\* 对于ARIB和许多DVB系统1PID = 0x0010。

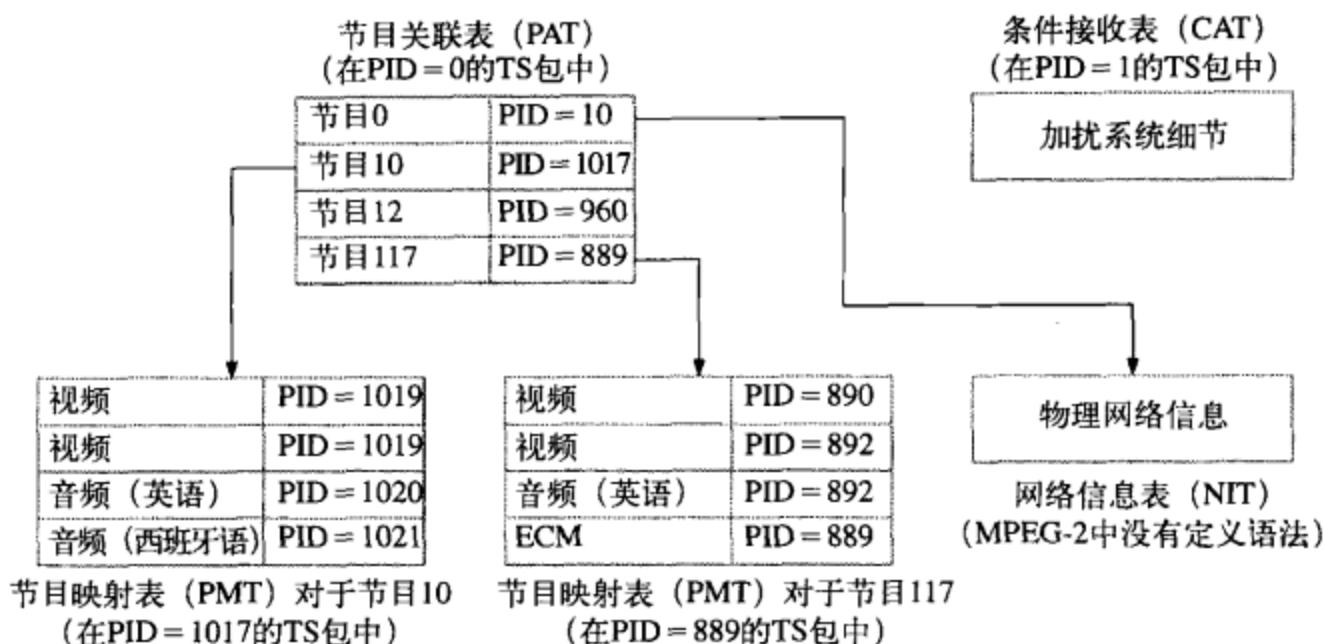


图13-24 MPEG-2 PAT和PMT示例

### 13.8.4 节目关联表

每个传输流包含一个或多个PID值为0x0000的传输流包。所有这些包一起组成了完整的节目关联表 (PAT)。

如图13-24所示, PAT提供传输流中所有节目的完全列表。每个节目都包括传输流包的PID值, 这些包中包含它对应的节目映射表 (PMT)。

由于没有PAT就不可能分离信号, 新节目的锁定速度取决于PAT传送的频率。MPEG-2在PAT和任意在PAT中提到的PMT之间定义了一个最大为0.5秒的间隔值。

采用下列语法, PAT在插入传输流包前可以分割成一个或多个分段 (section)。

#### 1. table\_ID

8位码字, 标识了内容的类型, 值为0x00, 如表13-49所示。www.atsc.org网站的Code Point Registry中提供table\_ID编码的完全列表。

表13-49 常用的Table\_ID码字

描述	编 码
PAT分段	0x00
CAT分段	0x01
PMT分段	0x02
TSDT分段	0x03
MPEG-4场景描述分段	0x04
MPEG-4对象描述符分段	0x05
元数据分段	0x06
ICIT分段	0x07

(续)

描述	编码
MPEG-2保留	0x08~0x38
MPEG-2.6 DSM-CC可寻址分段	0x39
包含多协议数据的MPEG-2.6 DSM-CC分段	0x3A
包含U-N消息的MPEG-2.6 DSM-CC分段	0x3B
包含下载数据消息的MPEG-2.6 DSM-CC分段	0x3C
包含流描述符的MPEG-2.6 DSM-CC分段	0x3D
包含私有数据的MPEG-2.6 DSM-CC分段	0x3E
MPEG-2.6 DSM-CC可寻址分段	0x3F
由DVB使用	0x40~0x7F
由ARIB、ATSC CA和DVB CA所使用	0x80~0x8F
由ATSC和SCTE使用	0xC0~0xFE
禁用	0xFF

**2. section\_syntax\_indicator**

1位标志位，值恒为1。

**3. reserved\_bits**

2位保留位，值为01。

**4. section\_length**12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过1021<sub>D</sub>或0x3FD。**5. transport\_stream\_ID**

16位二进制数，作为标签标识网络中来自于其他多路通道的传输流。

**6. reserved\_bits**

2位保留位，值为11。

**7. version\_number**

5位二进制数，它是整个PAT的版本号。任何时候只要PAT的定义发生变化，版本号就加1(模数32)。

当current\_next\_indicator值为1时，version\_number将是当前可用PAT的版本号。当current\_next\_indicator值为0时，version\_number将是下一个可用PAT的版本号。

**8. current\_next\_indicator**

1位标志位，值为1表示PAT传送当前可用。值为0表示PAT传送还不可用，但下一个PAT传送将变为有效。

**9. section\_number**

8位二进制数，定义了分段的编号。PAT中第一分段的section\_number值必须为0x00。每当出现附加分段，PAT中该字段的值就加1。

**10. last\_section\_number**

8位二进制数，定义了完整PAT最后一个分段的编号（也就是section\_number值最高的分段）。

注意：对于每个节目，下列4个字节会重复出现。

**11. program\_number**

16位二进制数，定义了program\_map\_PID分配的节目。

**12. reserved\_bits**

3位保留位，值为111。

**13. network\_PID**

13位二进制数，定义了包含网络信息表的传输流包的PID。它的值可能为0x0010~0x1FFE。仅当program\_number = 0x0000时该字节才出现。

**14. program\_map\_PID**

13位二进制数，定义了包含PMT分段的传输流包的PID，对于program\_number定义的节目，这个PMT分段是可用的。它的值可能为0x0010~0x1FFE。仅当program\_number ≠ 0x0000时该字节才出现。

**15. CRC\_32**

32位CRC值。

669

### 13.8.5 节目映射表

如图13-24所示，节目映射表（PMT）提供了节目编号和节目元素（视频、音频等）之间的映射关系。这项工作通过标识音频、视频和其他属于给定节目的流的PID值完成。

需要注意的是，对于ATSC、DVB和OpenCable™标准，某些PMT要求特定的PID。因此，MPEG-2和DVB/ATSC/OpenCable™比特流不能完全交互。

采用下列语法，PMT仅有一个由program\_number标识的分段。

**1. table\_ID**

8位码字，标识了内容的类型，值为0x02，如表13-49所示。[www.atsc.org](http://www.atsc.org)网站的Code Point Registry中提供table\_ID编码的完全列表。

**2. section\_syntax\_indicator**

1位标志位，值恒为1。

**3. reserved\_bits**

3位保留位，值为011。

**4. section\_length**

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过1021<sub>D</sub>或0x3FD。

**5. program\_number**

16位二进制数，定义了program\_map\_PID适用的节目。一个节目映射分段传送一个节目定义。节目定义绝不能长于1016<sub>D</sub> (0x3F8)。

**6. reserved\_bits**

2位保留位，值为11。

**7. version\_number**

5位二进制数，它是节目映射分段的版本号。当分段出现时携带的信息发生变化，版本号加1(模数32)。

当current\_next\_indicator值为1时，version\_number是当前可用节目映射分段。当current\_next\_indicator值为0时，version\_number将是下一个可用节目映射分段。

**8. current\_next\_indicator**

1位标志位，值为1表示节目映射分段传送当前可用。值为0表示节目映射分段传送还不可用，但下一个节目映射分段传送变为有效。

**9. section\_number**

8位二进制数，值为0x00。

**10. last\_section\_number**

8位二进制数，值为0x00。

**11. reserved\_bits**

3位保留位，值为111。

670

**12. PCR\_PID**

13位二进制数，指示包含PCR字段传输流包的PID，对于program\_number定义的节目，这一PCR字段是有效的。这一字段的值可能为0x0010~0x1FFE。如果对于私有流，没有PCR与节目定义关联，那么这一字段的值为0xFFFF。

**13. reserved\_bits**

4位保留位，值为1111。

**14. program\_info\_length**

12位二进制数，定义了紧接该字段之后出现的描述符总长（单位为字节）。

**15. descriptor\_loop**

[n]描述符，可能在descriptor\_loop中出现。

注意：对于任意出现的流类型，下列6个字段会重复出现。

**16. stream\_type**

8位码字，定义了包中节目元素的类型，该包的PID由elementary\_PID定义。stream\_type的值在表13-45中定义。

**17. reserved\_bits**

3位保留位，值为111。

**18. elementary\_PID**

13位二进制数，定义了携带相关节目元素的传输流包的PID。

**19. reserved\_bits**

4位保留位，值为1111。

**20. ES\_info\_length**

12位二进制数，定义了紧接该字段后出现的描述符总长（单位为字节）。

**21. descriptor\_loop**

[n]描述符，可能在descriptor\_loop中出现。

**22. CRC\_32**

32位CRC值。

### 13.8.6 传输流描述表

可选的传输流描述表（TSDT）用来包含应用于整个传输流的描述符。

采用下列语法，在插入传输流包之前，TSDT可以分割成一个或多个分段。

**1. table\_ID**

8位码字，标识了内容的类型，值为0x03，如表13-49所示。www.atsc.org网站的Code Point Registry中提供table\_ID编码的完全列表。

**2. section\_syntax\_indicator**

1位标志位，值恒为1。

[671]

**3. reserved\_bit**

1位保留位，值为0。

**4. reserved\_bits**

2位保留位，值为11。

**5. section\_length**

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字段的值不能超过 $1021_D$ 或0x3FD。

**6. reserved\_bits**

18位保留位，值为11 1111 1111 1111 1111。

**7. version\_number**

5位二进制数，它是节目映射分段的版本号。当分段出现时携带的信息就要发生变化，版本号加1（模数32）。

当current\_next\_indicator值为1时，version\_number是当前可用节目映射分段。当current\_next\_indicator值为0时，version\_number将是下一个可用节目映射分段。

**8. current\_next\_indicator**

1位标志位，值为1表示节目映射分段传送当前可用。值为0表示节目映射分段传送还不可用，但下一个节目映射分段传送变为有效。

**9. section\_number**

8位二进制数，定义了分段的编号。第一个分段的section\_number值必须为0x00。每当出现附加的分段时，该字段的值就加1。

**10. last\_section\_number**

8位二进制数，定义了最后一个分段的编号（也就是section\_number值最高的分段）。

**11. descriptor\_loop**

[n]描述符，可能在descriptor\_loop中出现。

**12. CRC\_32**

32位CRC值。

### 13.8.7 条件接收表

条件接收表（CAT）提供一个或多个条件接收（CA）系统、它们的权利管理消息（EMM）以及与之相关的特定参数之间的关联。CAT还包含权利控制消息（ECM）和权利管理消息（EMM）的PID。

采用下列语法，在插入传输流包之前，CAT可以分割成一个或多个分段。

**1. table\_ID**

8位码字，标识了内容的类型，值为0x01，如表13-49所示。[www.atsc.org](http://www.atsc.org)网站的Code Point Registry中提供table\_ID编码的完全列表。

**2. section\_syntax\_indicator**

1位标志位，值恒为1。

**3. reserved\_bits**

3位保留位，值为011。

**4. section\_length**

12位二进制数，定义了分段的字节数，该分段紧接着当前字段之后开始并包括CRC。这个字

段的值不能超过 $1021_{10}$ 或 $0x3FD$ 。

[672]

#### 5. reserved\_bits

18位保留位，值为11 1111 1111 1111 1111。

#### 6. version\_number

5位二进制数，它是整个CAT的版本号。当CAT变化时携带的信息发生变化，版本号加1（模数32）。

当current\_next\_indicator值为1时，version\_number将是当前可用CAT。当current\_next\_indicator值为0时，version\_number将是下一个可用CAT。

#### 7. current\_next\_indicator

当1位标志位为1时，发送的CAT当前可用。当1位标志位为0时，发送的CAT尚不可用，下一个CAT才有效。

#### 8. section\_number

8位二进制数，定义了分段的编号。CAT中第一个分段的section\_number值必须为 $0x00$ 。每当CAT中出现附加分段，该字段的值就加1。

#### 9. last\_section\_number

8位二进制数，定义了整个CAT最后一个分段的编号（也就是section\_number值最高的分段）。

#### 10. descriptor\_loop

[n]描述符，可能在descriptor\_loop中出现。

#### 11. CRC\_32

32位CRC值。

### 13.8.8 网络信息表

PAT中的首次登录是为网络数据保留的，且包含网络信息表（NIT）的PID。NIT包括其他可能可用传输流的信息，例如通过调谐至不同的RF频道或卫星。每个传输流可能包括定义射频、卫星轨道位置等信息的描述符。在MPEG-2中，对于这个目的只有NIT是强制要求的。在DVB中采用称为DVB-SI的附加元数据，而且NIT被视作DVB-SI的一部分。因此，术语PSI/SI是通用术语。

### 13.8.9 IPMP控制信息表

IPMP（知识产权管理和保护）控制信息表包含IPMP[也被称为数字权利管理(DRM) ]相关信息，其中包括工具列表、权利箱和工具箱。

工具列表表示并且启用处理内容所需的IPMP工具的选择。工具箱使内容流中能够携带二进制工具。权利箱可能包含权利描述，这些描述说明了与IPMP保护内容相关的适用规则。

IPMP流携带的所有类型IPMP数据（包括健、ECM、EMM）都会传送给这些工具。

[673]

## 13.9 知识产权管理和保护

知识产权管理和保护（IPMP）也被称作数字权利管理（DRM），它为知识产权管理的实现提供一类接口和工具，而不是一个完整的系统。

已提供的管理与保护的层次和类型取决于内容和业务模型的值。因此，IPMP系统的完整设计留待应用开发者来完成。

这一框架保证了开放和私有解决方案的使用，但同时保证互操作性，支持多种类型保护（如解密、水印、权利管理等）的使用，同时支持采用定义过的设备间消息（反映在家庭网络上的内

容分发问题) 实现设备间的内容传送。

对于受保护的内容, IPMP工具要求在显示开始前与解码器进行通信。工具配置和初始化信息通过IPMP描述符或IPMP基本流传送。所需工具可以通过内嵌、下载或其他方法获得。

在IPMP描述符中的控制点和顺序序列信息允许在系统的不同位置使用不同工具。IPMP描述符或IPMP基本流携带的IPMP数据包括权利箱、键箱和工具初始化数据。

### 13.10 在MPEG-2传输流上的MPEG-4.2视频

MPEG-2传输或节目流取代MPEG-2视频来传输MPEG-4.2视频。这使得现有的基础设施和设备能够方便地适应MPEG-4.2视频编解码。

对于MPEG-4.2视频, PES包的stream\_id = 1110 xxxx。PMT或PSM中的stream\_type = 0x10。MPEG-4.2流的传送也必须采用MPEG-4视频描述符发出信号告知。

MPEG-4.3音频、MPEG-4 SL打包流以及MPEG-4 FlexMux流也能够通过MPEG-2传输或节目流传送。

### 13.11 在MPEG-2传输流上的MPEG-4.10 (H.264)

MPEG-2传输或节目流取代MPEG-2视频在PES包中传送MPEG-4.10 (H.264) 视频。这使得现有的基础设施和设备能够方便地适应H.264视频编解码。

对于MPEG-4.10 (H.264) 视频, PES包的stream\_id = 1110 xxxx。PMT或PSM中的stream\_type = 0x1B。H.264流的传送也必须采用MPEG-2 AVC视频描述符或MPEG-2 AVC时序及HRD描述符发出信号告知。  
674

### 13.12 在MPEG-2传输流上的SMPTE 421M (VC-1) 视频

MPEG-2传输和节目流取代MPEG-2视频在PES包中传送SMPTE 421M (VC-1) 视频。这使得现有的基础设施和设备能够方便地适应SMPTE 421M (VC-1) 视频编解码。

对于SMPTE 421M视频, PES包stream\_ID = 0xFD, stream\_ID\_extension = 101 0101。在PMT或者PSM中, stream\_type = 0xEA。SMPTE 421M流的传送必须采用一个或多个MPEG-2注册描述符(Registration Descriptor) 来发出信号告知, 每一个描述符可能在additional\_identification\_info字段中包含一个可选的SMPTE 421M子描述符。

SMPTE 421M档次和级别子描述符(Profile and Level Subdescriptor) 可能被用来指定SMPTE 421M流的档次和级别。SMPTE 421M对齐子描述符(Alignment Subdescriptor) 可能被用来定义在编码的字节序列和一个PES包之间存在哪一种对齐类型。SMPTE 421M缓冲大小子描述符(Buffer Size Subdescriptor) 可能被用来定义在解码器中用于解码SMPTE 421M流所需的最小基本流缓冲区大小。

对于Simple或者Main档次流, VC-1\_SPMP\_PESpacket\_PayloadFormatHeader() 必须出现在每一个存取单元的起始处。

### 13.13 MPEG-2 PMT/PSM描述符

这些MPEG-2描述符被用来标识出现在MPEG-2传输或者节目流中的常用私有(非MPEG-2)信息。

描述符通常包含在MPEG-2 PMT或者PSM中的descriptor\_loop中。描述符的一般格式是:  
descriptor\_tag (8位)

`descriptor_length` (8位)

`data`

描述符标签值为0和1，而44~63保留。值19~26为MPEG-2.6数据保留。除非有其他指示，描述符可以同时出现在传输和节目流中。

### 13.13.1 MPEG-2 AAC音频描述符

对于在PES包中传送的独立MPEG-2.7音频流，MPEG-2 AAC描述符为标识这个音频基本流的编码参数提供了基本的信息。

#### 1. `descriptor_tag`

长度为8位的字段，值为0010 1011。

#### 2. `descriptor_length`

8位二进制数，定义了这个字段之后的字节数，值为0000 0011。

#### 3. `MPEG-2_AAC_profile`

8位字段，标识每个MPEG-2.7规范的AAC档次。

675

#### 4. `MPEG-2_AAC_channel_configuration`

8位字段，标识音频信道的数目和配置。

#### 5. `MPEG-2_AAC_additional_configuration`

8位字段，标识是否有带宽扩展数据被嵌入到音频流中。

### 13.13.2 音频流描述符

音频流描述符提供了标识一个音频基本流编码版本的基本信息。

#### 1. `descriptor_tag`

8位字段，值为0000 0011。

#### 2. `descriptor_length`

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

#### 3. `free_format_flag`

1位标志位，值为1指示音频流中的`bitrate_index`字段是0000。

#### 4. `ID`

这一位设置与音频流中的ID字段相同的值。

#### 5. `layer`

2位的二进制数，设置与任何音频流中的最高层相同或者更高的值。

#### 6. `variable_rate_audio_indicator`

值为0表示音频流的比特率在音频帧之间不发生变化。

#### 7. `reserved_bits`

3位保留位，值恒为111。

### 13.13.3 AVC时序和HRD描述符

对于在PES包中传送的独立MPEG-4.10 (H.264) 视频流，这个MPEG-2描述符描述了视频流时间信息和假定的参考解码器 (hypothetical reference decoder, HRD) 信息。当H.264视频流没有传递H.264视频可用信息 (video usability information, VUI) 参数时，这个描述符必须出现在PMT中。

**1. descriptor\_tag**

8位字段，值为0010 1010。

**2. descriptor\_length**

8位二进制数，定义了这个字段之后的字节数。

**3. HRD\_management\_valid\_flag**

1位标志位，值为1指示缓冲周期SEI（参看H.264）需要包含在H.264视频流中。值为0指示必须采用MPEG-2漏斗方法。  
676

**4. reserved\_bits**

6位保留位，值恒为11 1111。

**5. picture\_and\_timing\_info\_present**

值为1表示该描述符包含90kHz\_flag和用于映射到系统时钟的参数。

**6. 90kHz\_flag**

值为1表示H.264时间基是90kHz。0表示N和K字段将出现。仅当picture\_and\_timing\_info\_present = 1时这个字段才出现。

**7. reserved\_bits**

7位保留位，值恒为111 1111。仅当picture\_and\_timing\_info\_present = 1时这个字段才出现。

**8. N, K**

长度为32位的字段，描述了H.264的time\_scale和system\_clock\_reference之间的关系。仅当picture\_and\_timing\_info\_present = 1并且90kHz\_flag = 0时这些字段才出现。

**9. num\_units\_in\_tick**

长度为32位的字段，其定义请参看第14章。仅当picture\_and\_timing\_info\_present = 1时这个字段才出现。

**10. fixed\_frame\_rate\_flag**

1位标志位，值为1表示H.264编码视频帧率为固定值。值为0表示在此描述符中没有关于帧率的任何信息。

**11. temporal\_proc\_flag**

1位标志位，该位值为1且fixed\_frame\_rate\_flag = 1表示H.264视频流必须传送图像顺序计数(picture order count, POC)信息。该位值为0表示此描述符中没有任何关于H.264视频流的POC信息与时间之间关系的信息。

**12. picture\_to\_display\_conversion\_flag**

1位标志位，值为1表示H.264视频流包含关于显示编码图像的信息。当这一位值为0时，H.264视频流的pic\_struct\_present\_flag必须设置为0。

**13. reserved\_bits**

5位保留位，值恒为1 1111。

### 13.13.4 AVC视频描述符

对于在PES包中传送的独立MPEG-4.10（H.264）视频流，这个MPEG-2描述符描述了视频流的编码参数。当这个描述符没有出现在PMT中时，流应该不包含H.264静态图像或者H.264 24小时图片。

**1. descriptor\_tag**

8位字段，值为0010 1000。

**2. descriptor\_length**

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。

**3. profile\_IDC**

8位字段，标识了遵循H.264规范的H.264视频流的档次（profile）。

**4. constraint\_set0\_flag**

1位标志位，其定义在H.264规范中给出。

**5. constraint\_set1\_flag**

1位标志位，其定义在H.264规范中给出。

**6. constraint\_set2\_flag**

1位标志位，其定义在H.264规范中给出。

**7. AVC\_compatible\_flags**

5位标志位，与H.264规范中的序列参数集中的reserved\_zero\_5bits有相同的值。677

**8. level\_IDC**

8位字段，标识了遵循MPEG-4.10（H.264）规范的MPEG-4.10（H.264）视频流的级别（level）。

**9. AVC\_still\_present**

这一位为1表示H.264视频流包含H.264静态图像。值为0表示H.264视频流不应该包含H.264静态图像。

**10. AVC\_24\_hour\_picture\_flag**

1位标志位，值为1表示H.264视频流包含24小时图片，这些图片是显示时间超过24小时的存取单元。值为0表示H.264视频流不应该包含24小时图片。

**11. reserved\_bits**

6位保留位，值恒为11 1111。

### 13.13.5 CA描述符

这个MPEG-2描述符表示包含ECM、EMM或者SRM信息的传输流包的PID。如果出现在CAT中，那么存在一个系统范围内的条件接收管理系统。如果出现在PMT中，CA\_PID指向包含与节目相关的ECM包。

**1. descriptor\_tag**

8位字段，值为0000 1001。

**2. descriptor\_length**

8位二进制数，定义了这个字段之后的字节数。

**3. CA\_system\_ID**

16位的二进制数，定义了条件接收使用的类型。这个字段的编码是私下定义的。值为0x4ADD标明它是一个ATSC SRM参考描述符（Reference Descriptor）。

**4. reserved\_bits**

3位保留位，值恒为111。

**5. CA\_PID或SRM\_PID**

13位的二进制数，定义了传输流包的PID，这些传输流包或者包含ECM和EMM信息，或者包含关于由CA\_system\_ID定义的条件接收系统的SRM信息。

对于传输流，该字段值为0x0003表示在传输流中有组件使用IPMP。对于节目流，stream\_ID\_

extension值为0x00表示在节目流中有组件使用IPMP。

#### 6. private\_data\_byte

这些可选的[n]字节私有数据由CA所属者定义。

### 13.13.6 字幕服务描述符

对于OpenCable<sup>TM</sup>和一般的MPEG-2解码器，对于每一个有隐藏字幕（简称CC）的节目，这个CEA-708描述符必须出现。对于ATSC和OpenCable<sup>TM</sup>，这个描述符同样也必须出现在事件信息表(event information table, EIT) 的descriptor\_loop中。

MPEG-2最多支持16个独立的服务描述符，每一个服务描述符长度为6字节。

#### 1. descriptor\_tag

8位字段，值为1000 0110。

#### 2. descriptor\_length

8位二进制数，定义了这个字段之后的字节数。

#### 3. reserved\_bits

3位保留位，值恒为111。

#### 4. number\_of\_services

5位字段，取值为1~16，表示字幕服务出现的数目。

注意：[number\_of\_services]定义了下面9个字段重复的次数。

#### 5. language

3位编码，指定与每个ISO 639.2/B的字幕服务相关的语言。

#### 6. digital\_CC

1位标志位，值为1表示CEA-708字幕服务出现。值为0表示CEA-608字幕出现。

#### 7. reserved\_bit

保留位，值恒为1。

#### 8. reserved\_bits

5位可选保留位，值恒为1 1111，仅当caption\_type = 0时这些保留位才出现。

#### 9. line21\_field

1位可选标志位，值为1表示场2的CEA-608字幕出现。值为0表示场1的CEA-608字幕出现。仅当caption\_type = 0时这一位才出现。

#### 10. caption\_service\_number

可选的6位字段，取值范围为1~63，表示字幕流的服务数。仅当caption\_type = 1时这些位才出现。

#### 11. easy\_reader

1位标志位，值为1表示字幕服务包含了为新读者（beginning reader）提供的文本格式。值为0表示字幕服务不是为此制作的。

#### 12. wide\_aspect\_ratio

1位标志位，值为1表示字幕服务是为16 : 9显示器而格式化的。值为0表示字幕服务是为4 : 3显示器而格式化的，且可以可选地显示在16 : 9显示器的中心。

#### 13. reserved\_bits

14位保留位，值恒为11 1111 1111 1111。

### 13.13.7 版权描述符

这个MPEG-2描述符提供了一种为音视频作品标识身份的方法。对于DVB标准，这个描述符是可选的，如果存在，解码器可以将其忽略。

#### 1. descriptor\_tag

8位字段，值为0000 1101。

#### 2. descriptor\_length

8位二进制数，定义了这个字段之后的字节数。

679

#### 3. copyright\_ID

32位值，从版权登记机构获得。

#### 4. additional\_copyright\_info

可选的[n]字节数据，由版权所有者自行定义，永不变更。

### 13.13.8 数据流对齐描述符

这个MPEG-2描述符描述了相对于PES包载荷起始点的视频流语法的对齐准则。ATSC需要这个描述符出现在PMT分段的节目元素循环中，用来描述视频基本流。对于DVB标准，这个描述符是可选的，如果存在，解码器可将其忽略。

#### 1. descriptor\_tag

8位字段，值为0000 0110。

#### 2. descriptor\_length

8位二进制数，定义了这个字段之后的字节数，值为0000 0001。

#### 3. alignment\_type

8位的码字，定义了音频或者视频对齐类型，如表13-50所示。对于ATSC标准，这个值必须是0000 0010。

表13-50 alignment\_type码字

视频流对齐类型	音频流对齐类型	码
保留	保留	0000 0000
条带、图像、GOP或者序列	同步字	0000 0001
图像、GOP或序列		0000 0010
GOP或序列		0000 0011
序列		0000 0100
		0000 0101
保留	保留	:
		1111 1111

### 13.13.9 DTCP描述符

这个描述符用来控制HDCP保护和DTCP保护的数字输出，如IEEE 1394、USB、HDMI和IP网络。

#### 1. descriptor\_tag

8位字段，值为1000 1000。

#### 2. descriptor\_length

8位二进制数，定义了这个字段之后的字节数。

680

**3. CA\_system\_ID**

16位的二进制数，定义了条件接收使用的类型。值为0xFFFF（DTLA）。

**4. reserved\_bits**

5位保留位，值恒为1 1111。

**5. EPN**

1位加密加非声明（Encryption Plus NonAssertion，EPN）标志位，值为1说明非其他复制控制的内容将不会在Internet上重传。

**6. DTCP\_CCI**

2位码字，指定了数字复制代次管理：

00=自由复制

01=不再复制

10=复制1代

11=从不复制

**7. reserved\_bits**

5位保留位，值恒为1 1111。

**8. image\_constraint\_token**

1位标志位，值为1表示当输出到不受保护的高清晰模拟视频输出设备上时，高清晰内容必须限定最高总共520 000像素或者更少（例如，960×540p）。

**9. APS**

2位的码字，详细说明模拟保护服务（Analog Protection Service，APS）。

00=没有模拟保护服务

01=PSP开，颜色带关

10=PSP开，2行颜色带开

11=PSP开，4行颜色带开

### 13.13.10 DTS音频描述符

包含DTS<sup>®</sup>音频的PES包可能被包括在一个MPEG-2节目或者传输流中，其方式与MPEG或Dolby<sup>®</sup> Digital音频可以被包括的方式一样。

MPEG-2并不显式地支持DTS<sup>®</sup>比特流。同样，在PSI表中，MPEG-2音频流描述符并没有充分地描述DTS<sup>®</sup>比特流的内容。

因此，包含DTS<sup>®</sup>音频数据的PES包将采用私有流1来发送。此外，还需要一个注册描述符Registration Descriptor（descriptor\_tag = 0000 0101）和一个DTS音频描述符（descriptor\_tag = 1001 0001或0111 0011）。

### 13.13.11 层次描述符

这个MPEG-2描述符提供了标识节目元素的信息，这些节目元素包含了层次编码视频、音频和私有流的组件。

**1. descriptor\_tag**

8位字段，值为0000 0100。

**2. descriptor\_length**

8位二进制数，定义了这个字段之后的字节数，值为0000 0100。