

MPEG视频编码I



讲师：肖俊
(肖俊)
软件与技术学院

内容

- MPEG-1

- MPEG - 1中的运动补偿——与H.261的其他主要差异——MPEG - 1视频比特流

- MPEG-2

- 简介——隔行扫描视频
- MPEG - 2的可伸缩性
- 与MPEG - 1的其他主要差异

MPEG - 1简介

•1991年 国际标准化组织/国际电工委员会

- 用于最高约1.5Mbit/s数字存储媒体的运动图像及相关音频编码——CD和VCD；视频1.2M，音频256K——5个部分：系统、视频、音频、一致性、软件——采用CCIR601数字电视格式——SIF（源输入格式）——仅支持非隔行视频——NTSC制式30帧/秒时为352*240——PAL制式25帧/秒时为352*288——4:2:0色度子采样

1.1 MPEG - 1中的运动补偿

- 基于运动补偿 (MC) 的H. 261视频编码工作原理如下：

- 在运动估计 (ME) 中，目标P帧的每个宏块 (MB) 会从之前编码的I帧或P帧预测中分配一个最佳匹配的宏块。
- 预测误差：宏块与其匹配宏块之间的差异，该差异会被发送到离散余弦变换 (DCT) 及其后续编码步骤。
- 预测来自前一帧——前向预测。

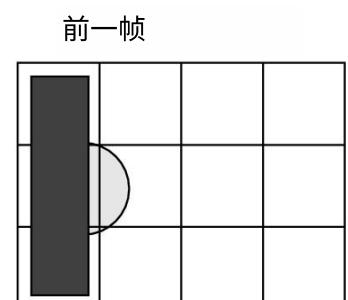
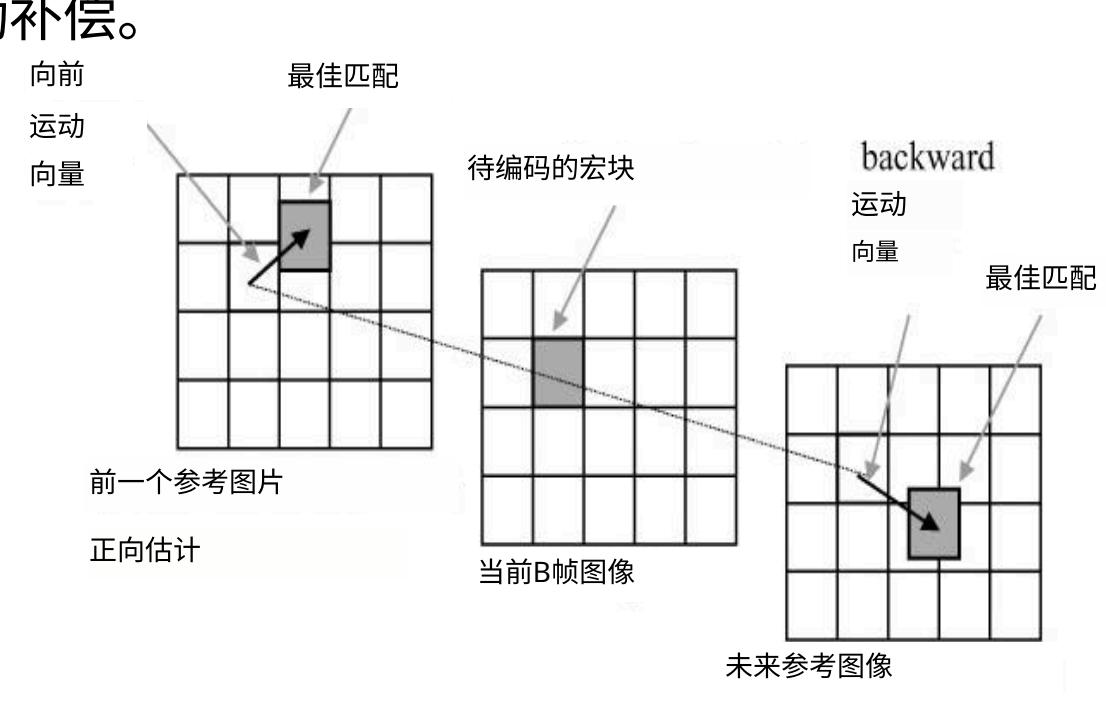


图11.1：双向搜索的必要性。

目标帧中包含部分球的宏块 (MB) 无法在前一帧中找到良好的匹配宏块，因为球的一半被另一个物体遮挡。然而，可以很容易地从下一帧中获得匹配。

- MPEG引入了第三种帧类型 - B帧，以及与之配套的双向运动补偿。



概述

- MPEG

- 运动图像专家组，1998年成立，旨在提供数字视频和音频

•成员情况

- 从1988年的25位专家，发展到来自约200家公司的350多位专家

1. MPEG-1

1.1 MPEG - 1中的运动补偿

1.1 MPEG - 1中的运动补偿

• 基于运动补偿的B帧编码理念如图11.2所示：

- B帧中的每个宏块最多有两个运动矢量（MV）（一个来自前向预测，一个来自后向预测）。
- 如果两个方向的匹配都成功，则会发送两个运动矢量，并在与目标宏块进行比较以生成预测误差之前，对两个相应的匹配宏块进行平均（图中用“%”表示）。- 如果仅在其中一个参考帧中找到可接受的匹配，则仅从前向或后向预测中使用一个运动矢量及其对应的宏块。

1.1 MPEG - 1中的运动补偿

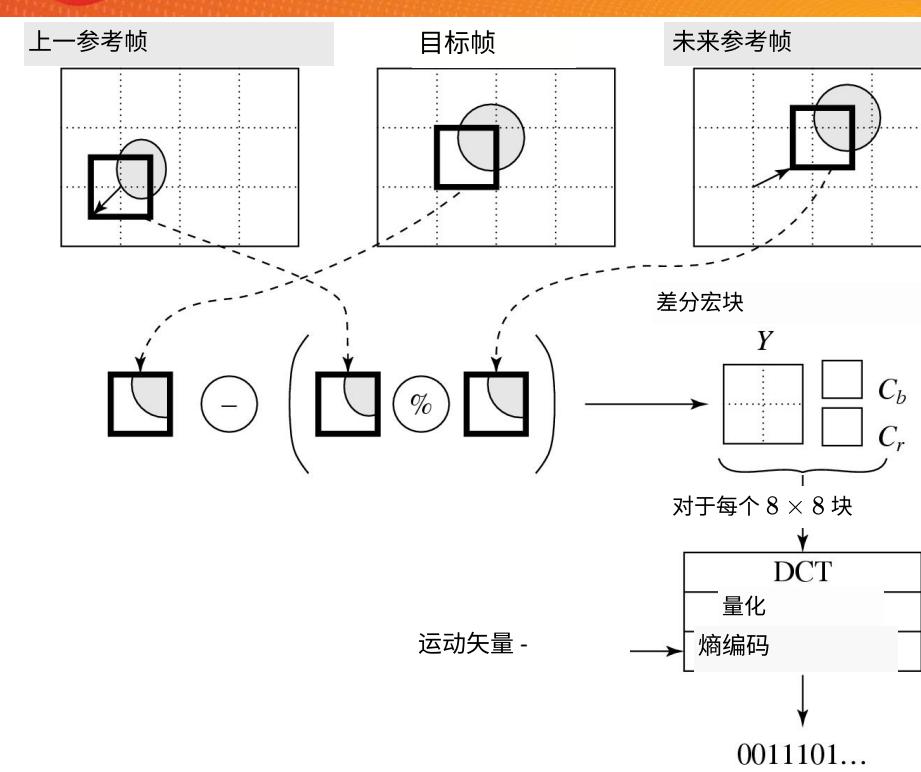


图11.2：基于双向运动的B帧编码
补偿。

1.1 MPEG - 1中的运动补偿

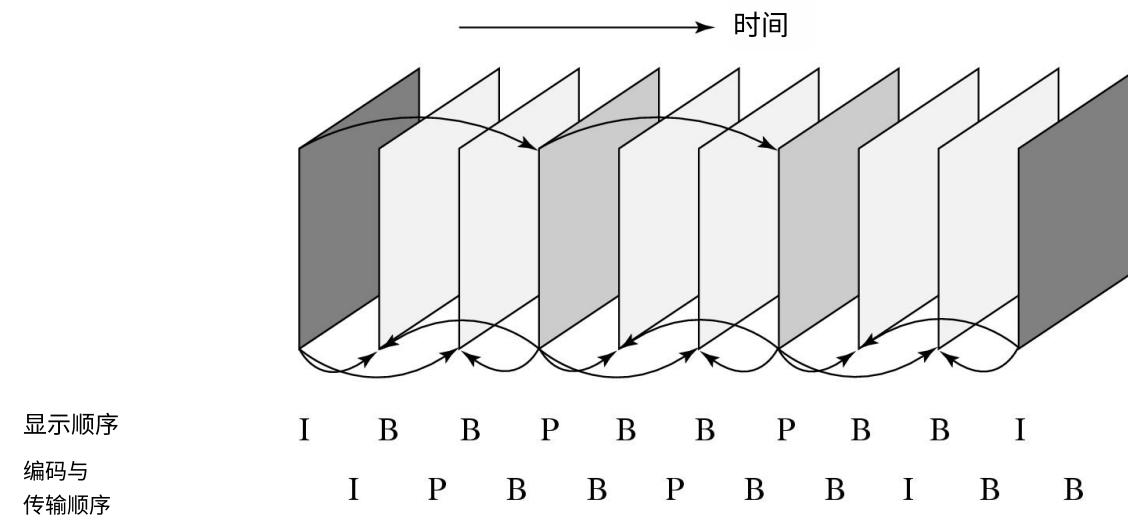
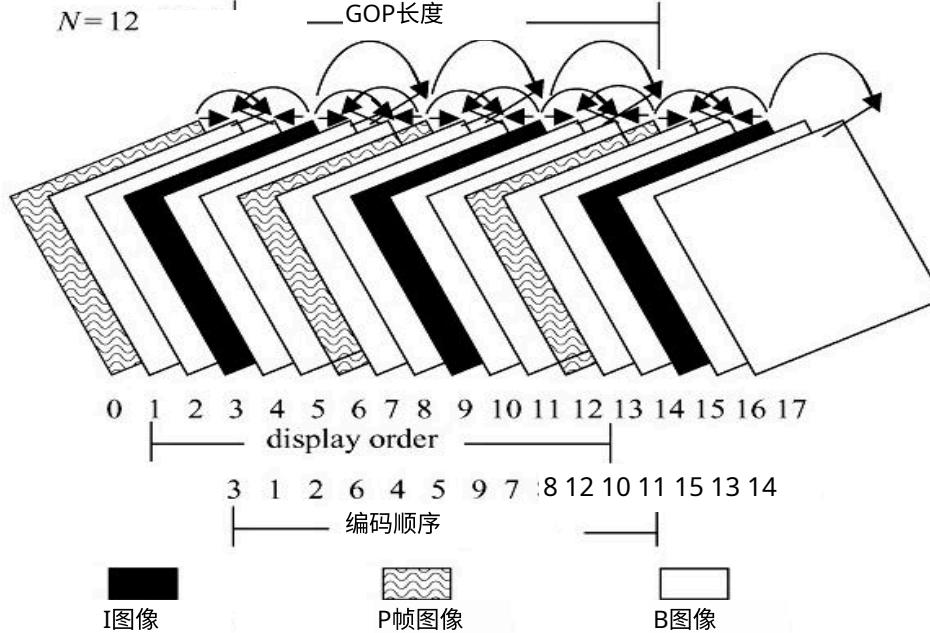


图11.3：MPEG帧序列。

1.1 MPEG - 1中的运动补偿

- B帧依赖于其后的P帧或I帧，因此播放顺序与编码顺序不同。



1.2 与H.261的其他主要区别

- H.261
 - 仅支持CIF (352×288)、QCIF (176×144)
- MPEG-1
 - 支持SIF (352×240 (适用于NTSC制), 352×288 (适用于PAL制))
 - 允许指定其他格式

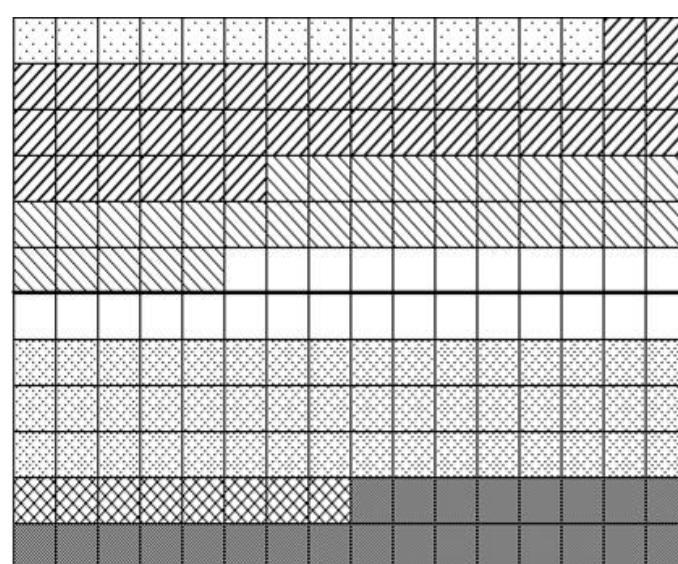
参数	值
图片的水平尺寸	≤ 768
图片的垂直尺寸	≤ 576
每张图片的兆字节数	≤ 396
每秒兆字节数	$\leq 9,900$
帧率	≤ 30 帧/秒
比特率	$\leq 1,856$ 千比特/秒

1.2 与 H. 261 的其他主要差异

- 与H.261中的GOB不同，MPEG - 1图像可划分为一个或多个片（图11.4）：
 - 单幅图像中可包含数量可变的宏块。
 - 只要能填满整幅图像，片的起始和结束位置可以任意设定。
 - 每个片独立编码——在比特率控制方面具有额外的灵活性。
 - 片的概念对于错误恢复很重要。

1.2 与H. 261的其他主要差异

- MPEG - 1图像可分为一个或多个片，每个片独立编码。



MPEG - 1图像中的片

1.2 与H. 261的其他主要差异

量化：

- MPEG - 1量化在其帧内和帧间编码中使用不同的量化表（表11.2和11.3）。

对于帧内模式下的DCT系数：

$$QDCT[i, j] = \text{round} \left(\frac{8 \times DCT[i, j]}{\text{step_size}[i, j]} \right) = \text{round} \left(\frac{8 \times DCT[i, j]}{Q_1[i, j] * \text{scale}} \right) \quad (11.1)$$

对于帧间模式下的DCT系数：

$$QDCT[i, j] = \left\lfloor \frac{8 \times DCT[i, j]}{\text{step_size}[i, j]} \right\rfloor = \left\lfloor \frac{8 \times DCT[i, j]}{Q_2[i, j] * \text{scale}} \right\rfloor \quad (11.2)$$

1.2 与H.261的其他主要差异

- 帧内编码的量化数在一个宏块内会发生变化，这与H.261不同

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

帧内量化表

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

帧间量化表

多媒体基础——MPEG视频编码I

17

1.2 与H.261的其他主要差异

- 压缩P帧的典型大小明显小于I帧 —— 因为帧间压缩利用了时间冗余。

- B帧甚至比P帧更小 —— 原因是 (a) 双向预测的优势，以及 (b) B帧的优先级最低。

表11.4: MPEG - 1帧的典型压缩性能

类型	大小	压缩
I	18kB	7:1
P	6kB	20:1
B	2.5kB	50:1
平均	4.8kB	27:1

多媒体基础——MPEG视频编码I

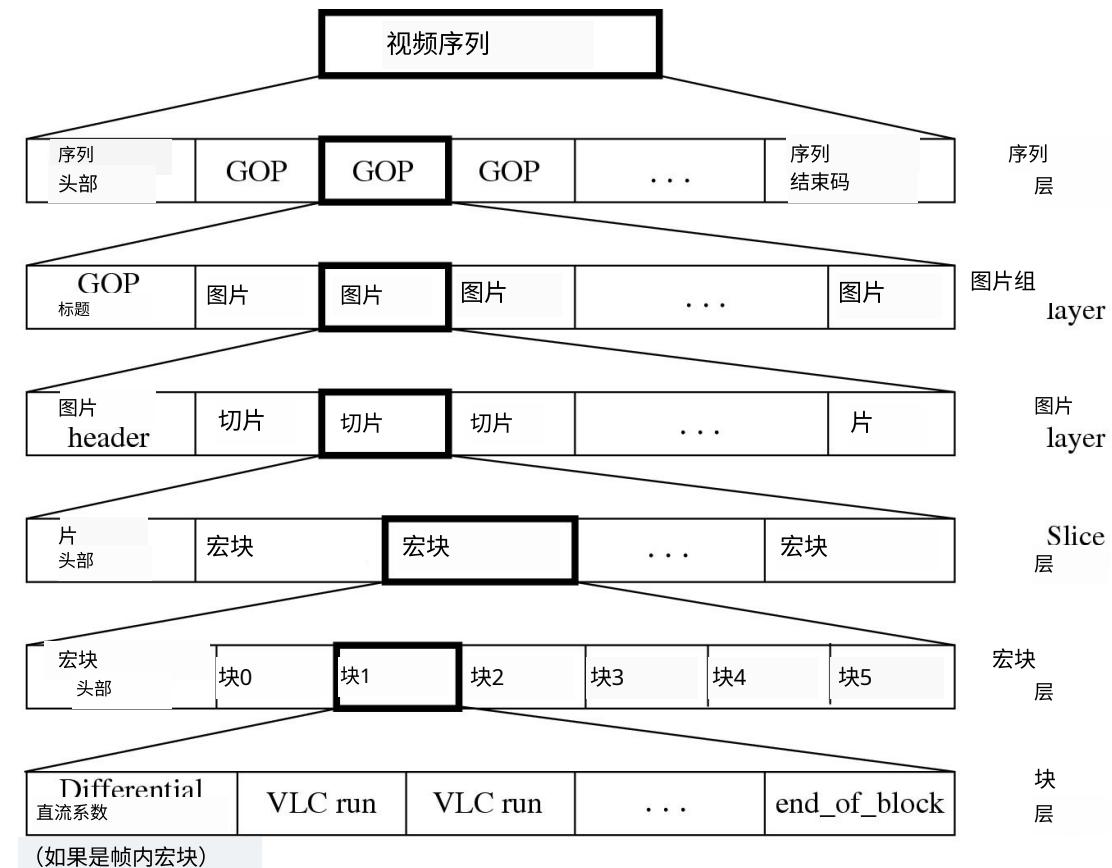
19

1.3 MPEG - 1视频比特流

多媒体基础 - MPEG视频编码I

18

1.3 MPEG - 1视频比特流



多媒体基础——MPEG视频编码I

20

- 图像组 (GOP) 层：一个或多个图片，其中一张必须是I帧图片

- 帧：可以是I帧、P帧或B帧

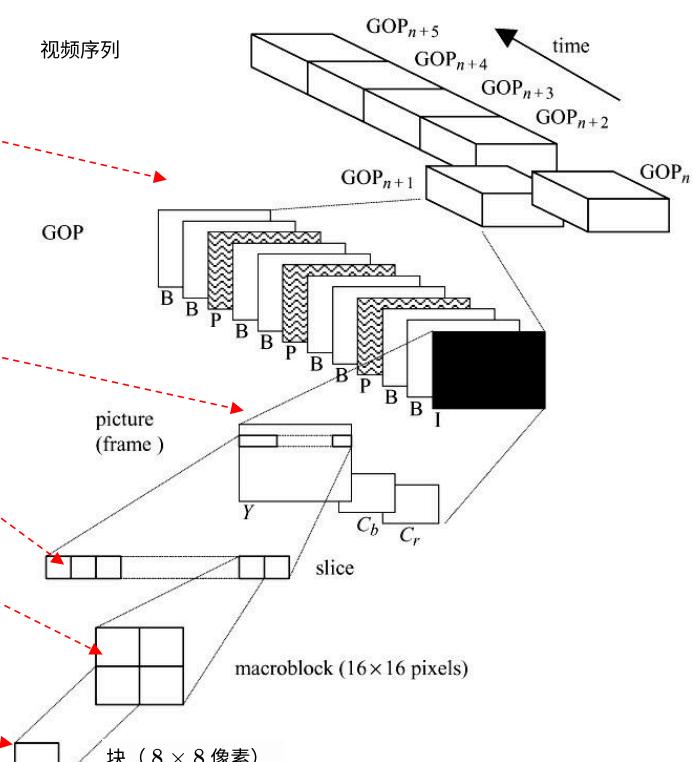
- 类型

- 片：GOB (宏块组)

- 宏块：16*16

宏块，以 4:2:0, 4个Y块，1个U块，1个V块

- 块：用于离散余弦变换 (DCT) 和量化



多媒体基础——MPEG视频编码I

21

2.1 概述

- 始于1990年，用于以超过4 Mbps的比特率提供更高质量的视频。

- 满足数字电视/高清电视的压缩和比特率要求

- 不同的分辨率，不同的压缩复杂度。

- 获得广泛认可：地面、卫星、有线电视网络

- 其他应用：交互式电视、DVD (数字视频光盘或数字多功能光盘)

2.1 概述

- MPEG - 2定义了7种档次，针对不同应用，每个档次最多定义4个级别。

- 简单档次、主档次、信噪比可分级档次、空间可分级档次、高级档次、4:2:2档次和多视图档次

级别	最大值	最大帧率	最大值 像素 $\times 10^6$ /秒	最大编码数据速率 (Mbps)	应用
高	1920×1152	60	62.7	80	电影制作
高1440	1440×1152	60	47.0	60	消费级高清电视
主类	720×576	30	10.4	15	演播室电视
低级	352×288	30	3.0	4	消费级磁带等效格式

MPEG - 2主类中的四个级别

2.2 支持隔行视频

•MPEG - 2也必须支持隔行视频，因为这是数字广播电视和高清电视的可选方案之一。

•在隔行视频中，每一帧由两个场组成，分别称为顶场和底场。

•在帧图像中，两个场的所有扫描线交织形成单帧，然后划分为 16×16 个宏块并使用运动补偿（MC）进行编码。

•如果每个场被视为单独的图像，则称为场图像。

2.2 支持隔行视频

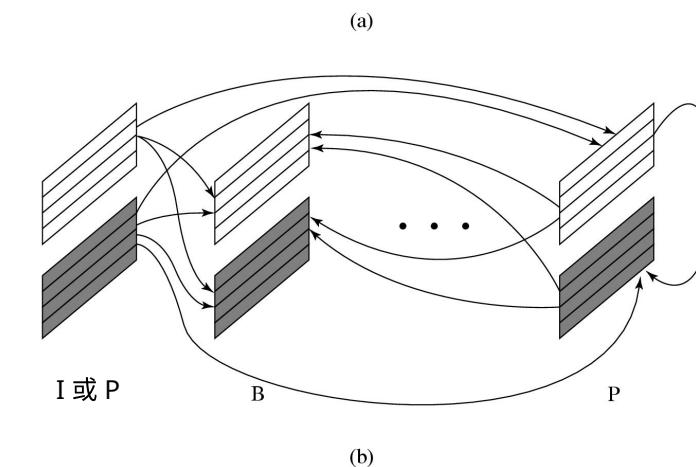
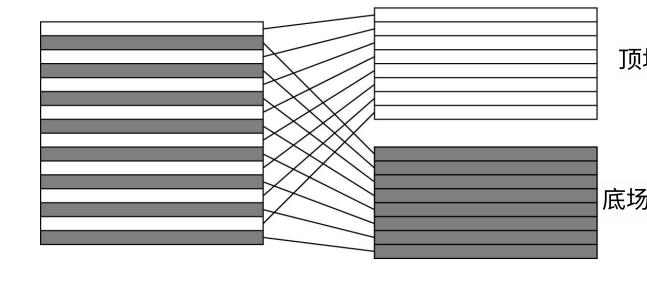


图 11.6：MPEG - 2 中的场图像和场图像的场预测。(a) 帧图像与场图像，(b) 场图像的场预测

2.2 支持隔行视频

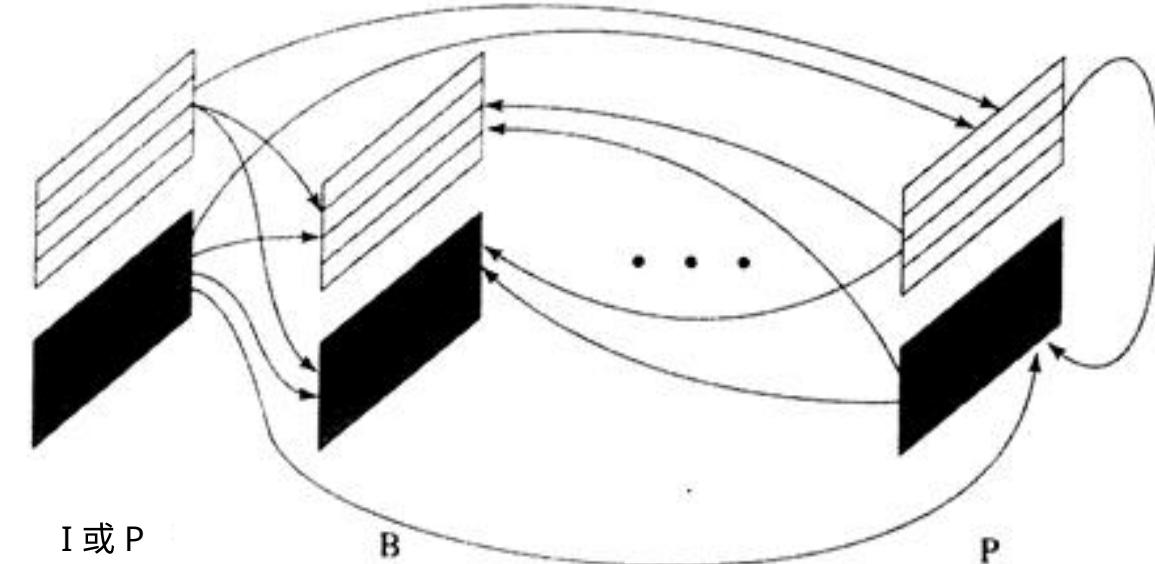
•五种预测模式

•帧图像的帧预测——与 Mpeg1 运动

补偿相同

•适用于包含缓慢和中等速度物体的视频 - 场画面的场预测 - 见下一页的图 - 帧画面的场预测 - 对于场画面，分别处理顶场和底场 - 16×8 MC - 适用于快速且不规则的运动 - P 画面的双基预测 - 运动矢量 (MV) 用于推导计算运动矢量 (CV)

•场图像的场预测



2.2 支持隔行视频

交替扫描和场 DCT

•旨在提高 DCT 对预测误差处理效果的技术，仅适用于隔行视频中的帧画面：

•由于隔行视频的特性， 8×8 块中的连续行来自不同场，它们之间的相关性低于隔行行之间的相关性。 - 交替扫描认识到在隔行视频中，垂直方向上较高的空间频率分量可能具有更大的幅度，因此允许在序列中更早地扫描这些分量。 - 在MPEG - 2中，场DCT也可用于解决相同的问题。

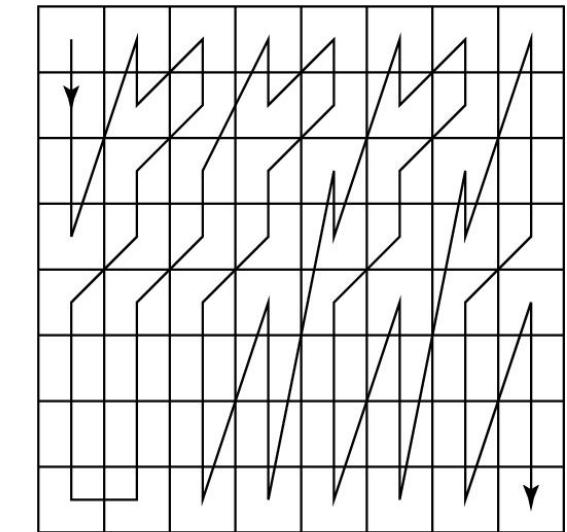
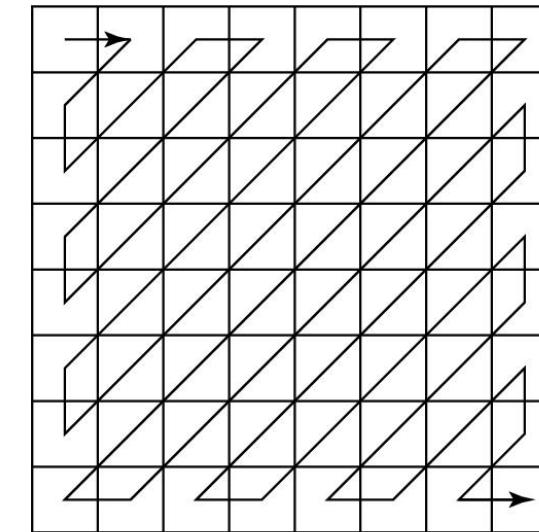


图11.7：MPEG - 2中逐行和隔行视频的DCT系数之字形和交替扫描。

2.3 MPEG - 2的可扩展性

•MPEG - 2可伸缩编码：可以定义一个基础层和一个或多个增强层 —— 也称为分层编码。

•基础层可以独立进行编码、传输和解码，以获得基本的视频质量。

•增强层的编码和解码依赖于基础层或前一个增强层。

•可伸缩编码对于通过具有以下特征的网络传输的MPEG - 2视频特别有用：

•比特率差异很大的网络。

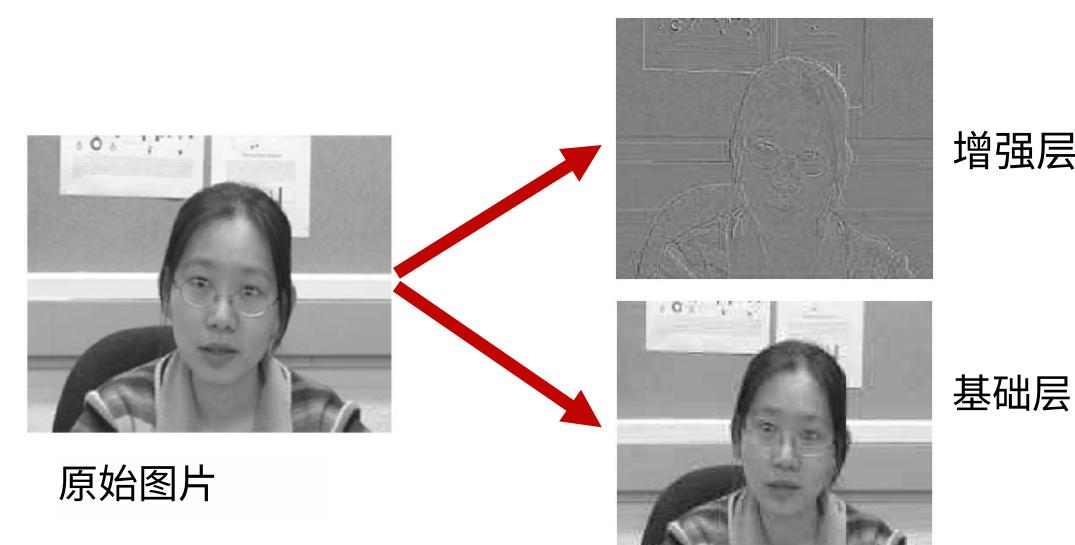
- 具有可变比特率 (VBR) 信道的网络。

•连接存在噪声的网络。多媒体基础

2.3 MPEG - 2可扩展性

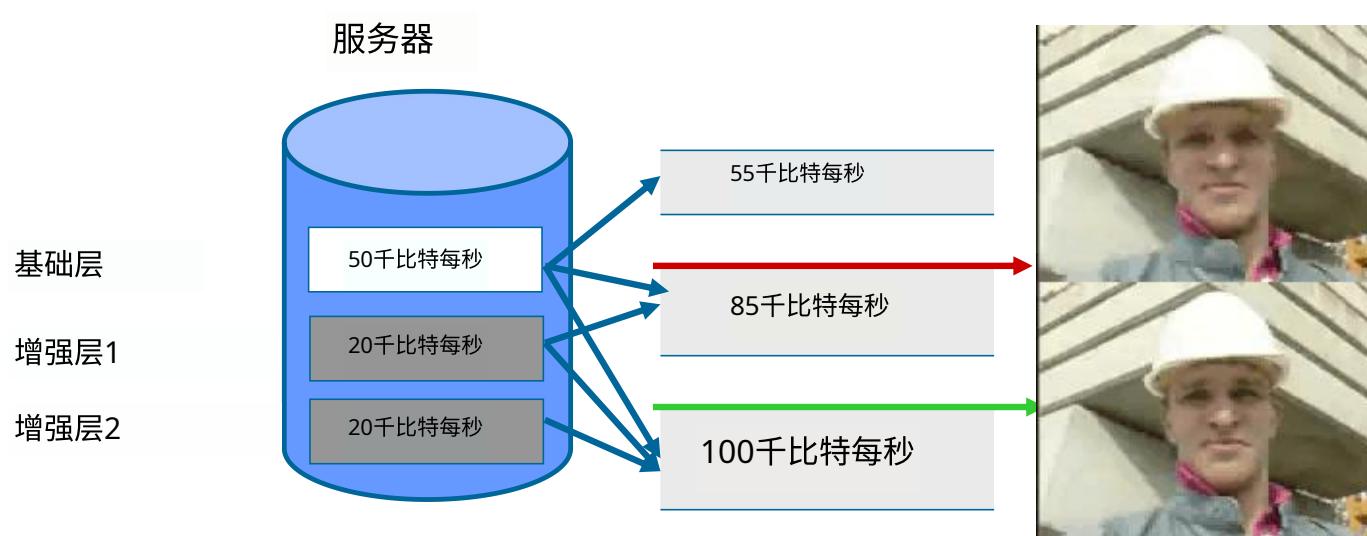
•可伸缩编码：基础层提供基本视频质量，增强层提供更高质量。

•基础层独立编码，增强层依赖于基础层或前一个增强层。



2.3 MPEG - 2可伸缩性

- 在带宽不足时仅发送基本层
- 为宽带提供基本层和增强层，以便接收器能获得更好的质量。



2.3 MPEG - 2可伸缩性

- MPEG - 2支持以下可伸缩性：
 1. 信噪比可伸缩性——增强层提供更高的信噪比。
 2. 空间可伸缩性——增强层提供更高的空间分辨率。
 3. 时间可扩展性增强层有助于实现更高的帧率。
 4. 混合可扩展性 - 上述三种可扩展性中任意两种的组合。
 5. 数据分区 - 量化后的离散余弦变换 (DCT) 系数被划分为多个分区。

2.3 MPEG - 2可扩展性

- 信噪比可扩展性：指在基础层之上进行增强/细化以提高信噪比 (SNR)。
- MPEG - 2信噪比可扩展编码器将在两层生成输出比特流：基础比特流 (Bits_base) 和增强比特流 (Bits_enhance)。
 1. 在基础层，采用对离散余弦变换 (DCT) 系数进行粗量化的方
法，这会使所需比特数减少，但视频质量相对较低。
 2. 然后对粗量化后的离散余弦变换 (DCT) 系数进行 b_0 反量化，并将其输入增强层，与原始的离散余弦变换 (DCT) 系数进行比较。
 3. 对它们的差值进行细量化，以生成离散余弦变换 (DCT) 系数的细化值，该细化值经过可变长度编码 (VLC) 后，成为名为“Bits_enhance”的比特流。

2.3 MPEG - 2的可扩展性

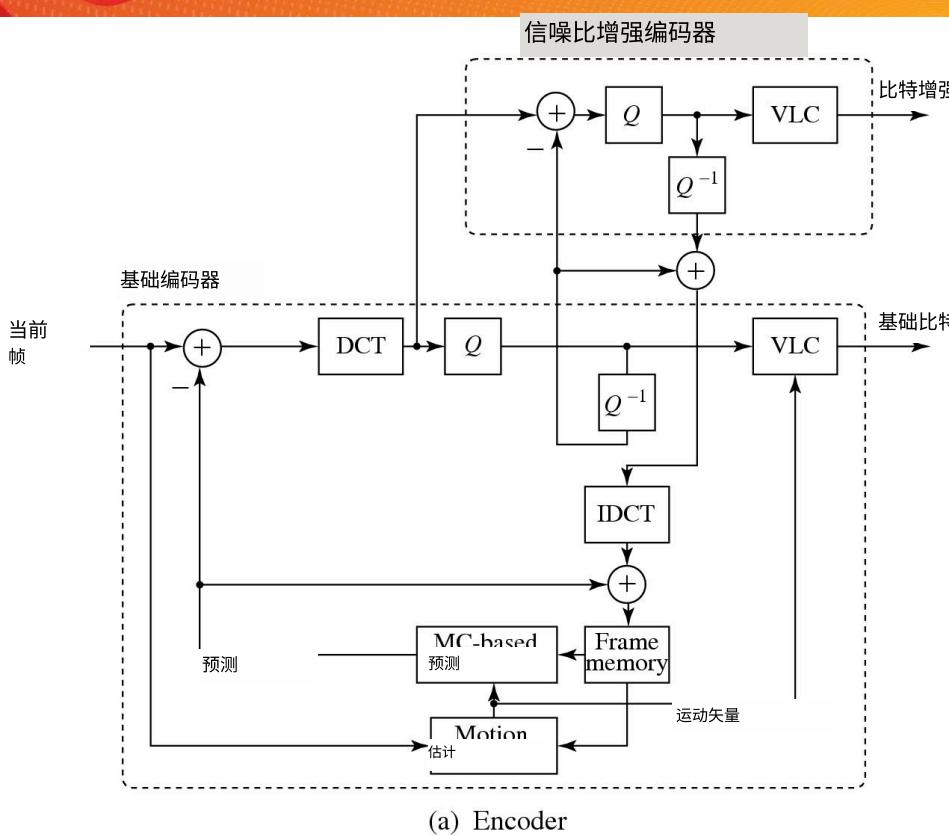


图11.8 (a): MPEG - 2信噪比可伸缩性 (编码器)。多媒体基础——MPEG视频编码I

2.3 MPEG - 2可伸缩性

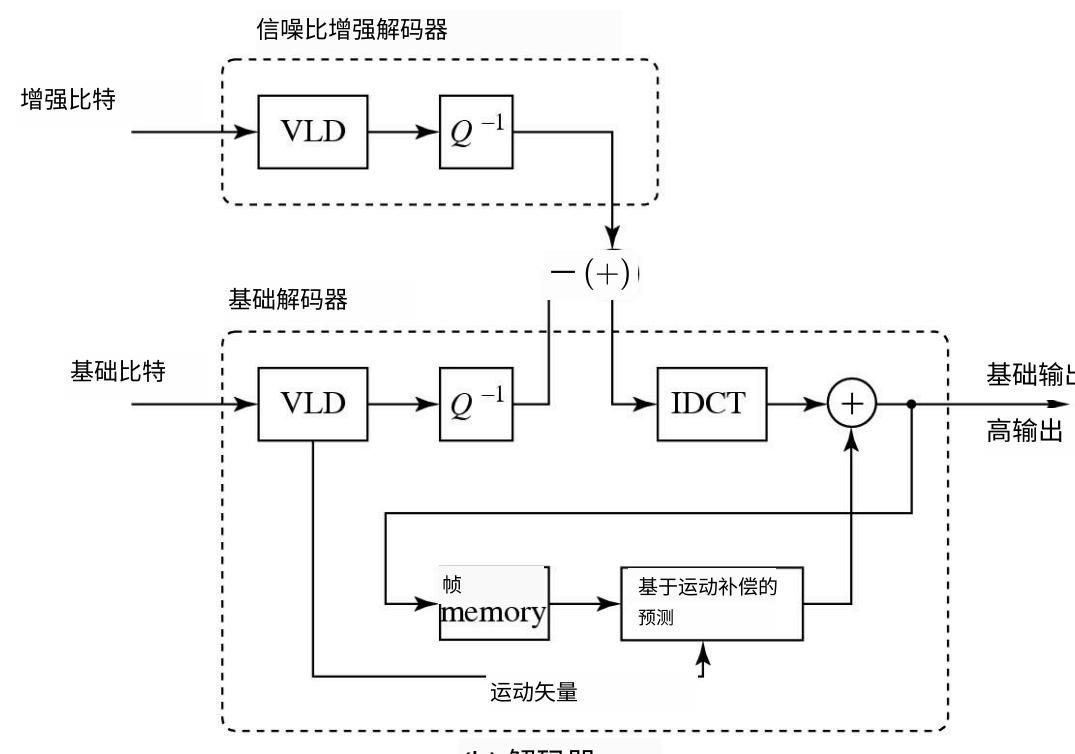
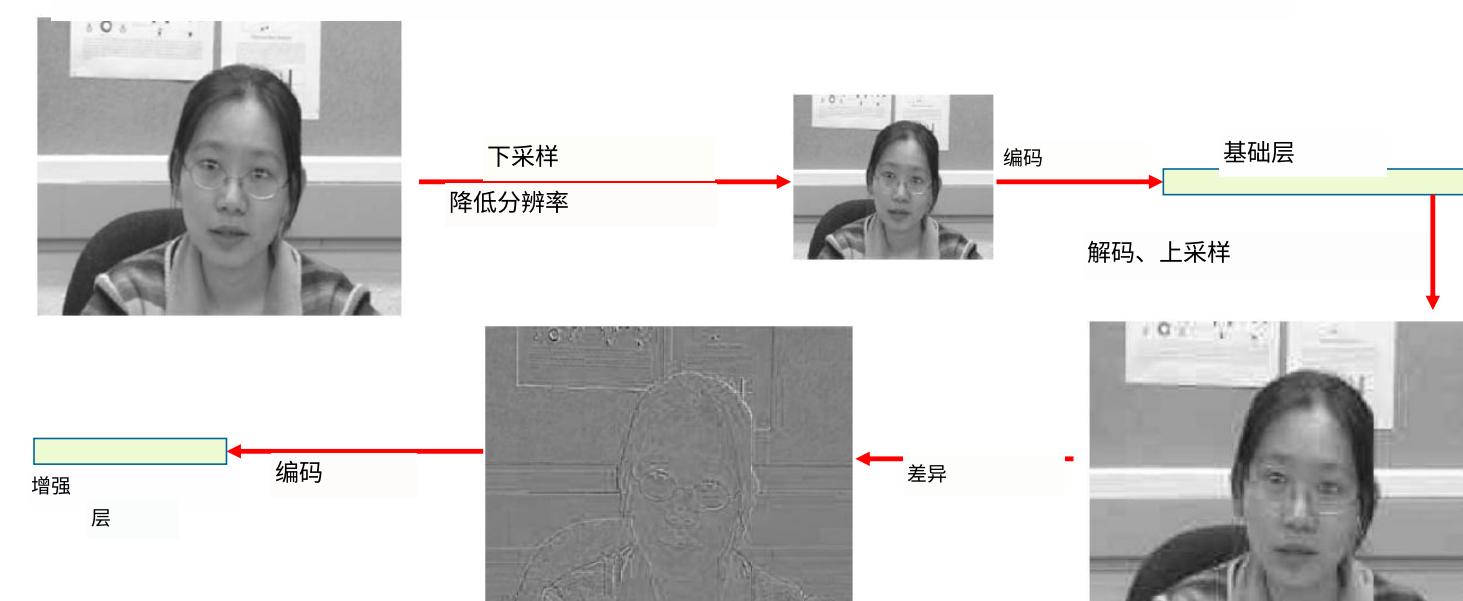


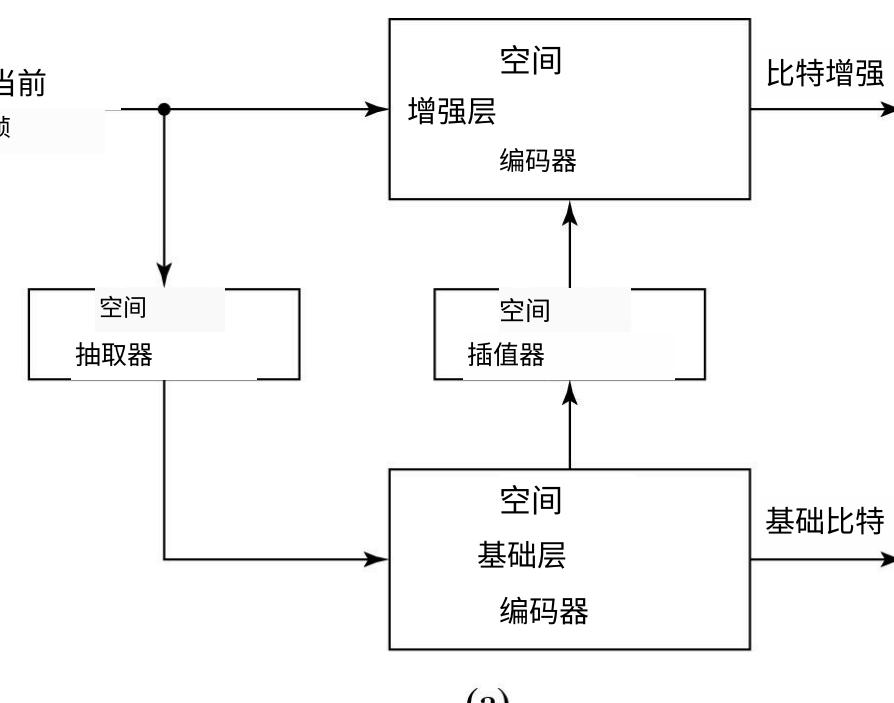
图11.8 (b): MPEG - 2信噪比可伸缩性 (解码器)。多媒体基础——MPEG视频编码I

空间可伸缩性

基础层旨在生成低分辨率图片的比特流。与增强层结合时，可生成原始分辨率的图片。



2.3 MPEG - 2可扩展性



MPEG - 2空间可伸缩性编码器

时间可伸缩性

• 输入视频在时间上被解复用为两部分，每部分承载原始帧率的一半。

• 基层编码器对其自身的输入视频执行正常的单层编码程序，并生成输出比特流Bits_base。

• 增强层匹配宏块的预测可以通过两种方式获得：

- 层间运动补偿 (MC) 预测 (图11.10(b))

• 组合运动补偿预测与层间运动补偿预测 (图11.10(c))

2.3 MPEG - 2可伸缩性

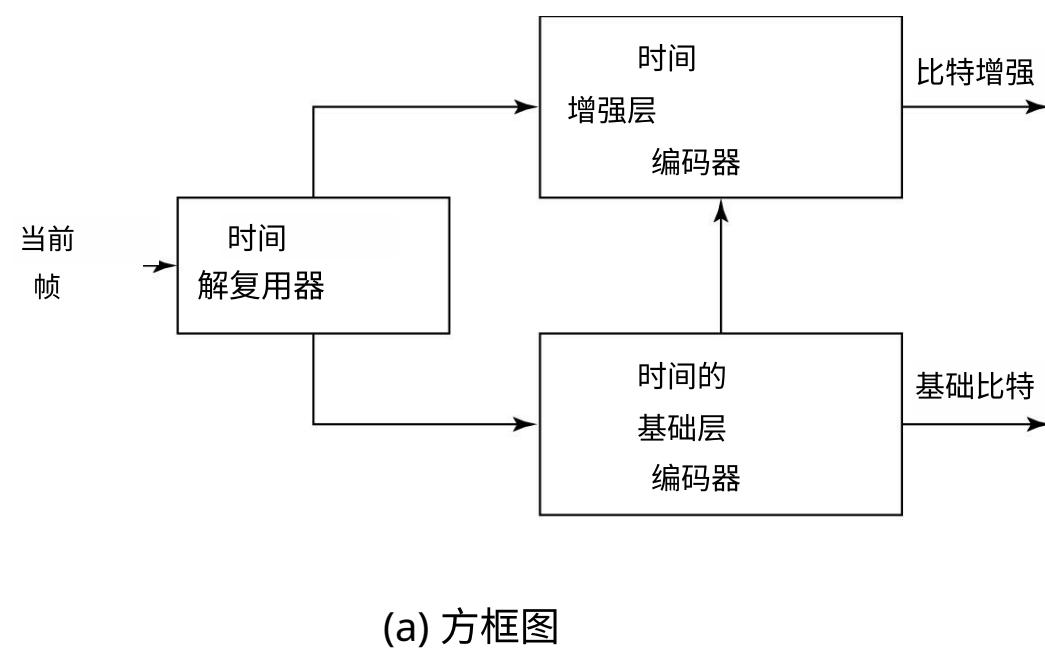
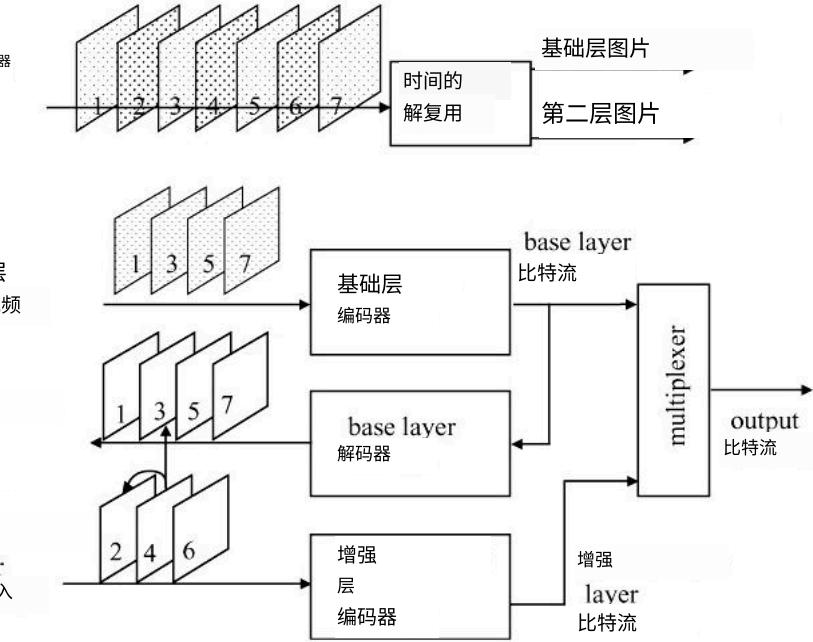


图11.10: MPEG - 2时间可伸缩性编码器。

2.3 MPEG - 2可伸缩性

- 基础层和增强层的图片与输入具有相同的空间分辨率。时间可伸缩性

1、3、5、7等，奇数帧作为基础层，偶数帧作为增强层。基础层按普通视频编码，通常有I帧和P帧类型。增强层的帧可以以基础层或其自身层的帧作为参考



2.3 MPEG - 2 可扩展性

混合可伸缩性

- 上述三种可伸缩性中的任意两种可以组合形成混合可伸缩性：

- 空间和时间混合可伸缩性。
- 信噪比和空间混合可伸缩性。
- 信噪比和时间混合可伸缩性。

- 通常，将采用由基础层、增强层 1 和增强层 2 组成的三层混合编码器。

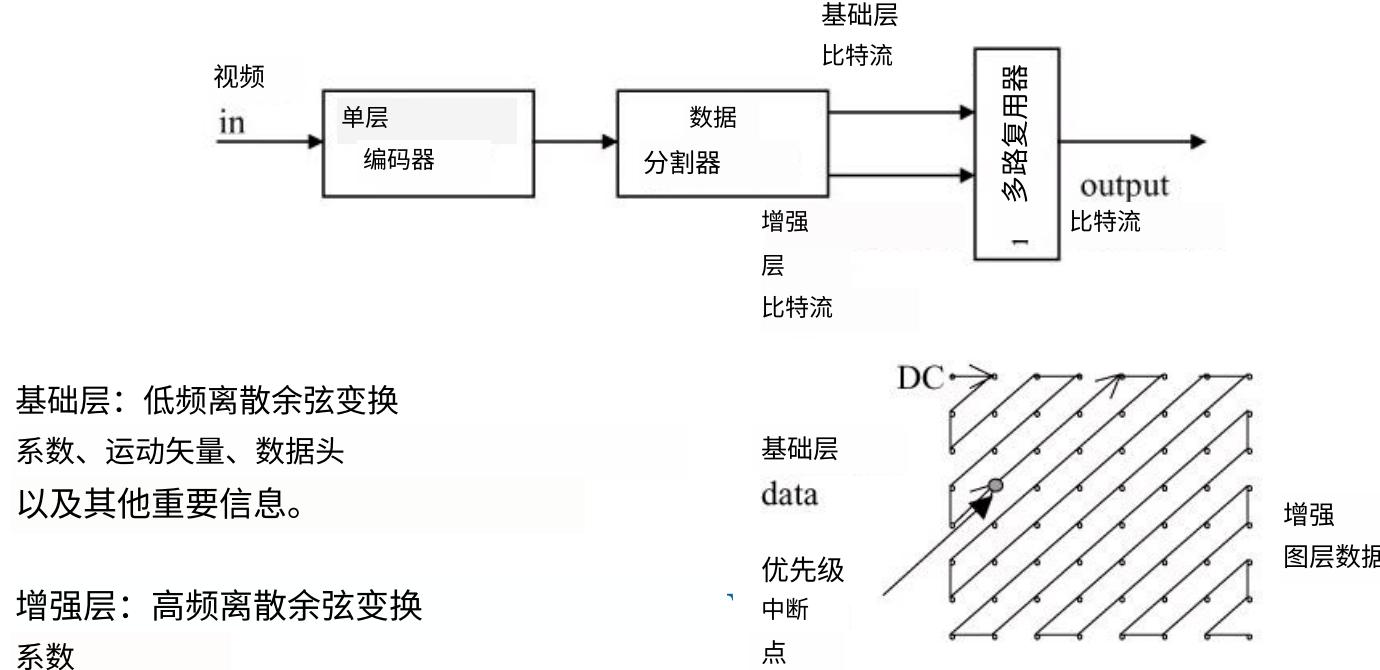
2.3 MPEG - 2可伸缩性

数据分割

- 基础分区包含低频DCT系数，增强分区包含高频DCT系数。
- 严格来说，数据分割并非分层编码，因为单一视频数据流只是被简单分割，且在生成增强分区时对基础分区没有进一步依赖。
- 适用于在噪声信道上传输以及渐进式传输。

2.3 MPEG - 2可伸缩性

- 它根据重要性将视频分为两个部分。



2.3 MPEG - 2可伸缩性

数据分区说明



2.4 与MPEG - 1的其他主要差异

- 更好的抗比特错误能力：除节目流外，MPEG - 2比特流中增加了传输流。
- 支持4:2:2和4:4:4色度子采样。
- 更严格的片结构：MPEG - 2的片必须在同一宏块行开始和结束。换句话说，图像的左边缘总是开始一个新的片，并且MPEG - 2中最长的片只能有一行宏块。
- 更灵活的视频格式：它支持DVD、ATV和HDTV所定义的各种图像分辨率。

2.4 与MPEG - 1的其他主要区别

- 非线性量化——允许两种类型的量化尺度：

- 对于第一种类型，量化尺度与MPEG - 1中的相同，即它是一个介于 $[1, 31]$ 和 $i = i$ 之间的整数。
- 对于第二种类型，存在一种非线性关系，即尺度 $i \neq i$ 。第 i 个尺度值可以从表11.7中查找。

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$scale_i$	1	2	3	4	5	6	7	8	10	12	14	16	18	20	22	24
i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
$scale_i$	28	32	36	40	44	48	52	56	64	72	80	88	96	104	112	

结束

谢谢！

邮箱：junx@cs.zju.edu.cn