





# 多媒体通信

# Multimedia Communications

## 第2章 多媒体数据压缩国际标准

### 关于视频压缩标准

 2019年10月12日 



# 课程内容

- ◆ 第1章 多媒体通信概述
- ◆ 第2章 多媒体数据压缩国际标准
- ◆ 第3章 多媒体同步机制
- ◆ 第4章 多媒体网络QoS
- ◆ 第5章 多媒体传输网络
- ◆ 第6章 多媒体通信终端、系统与国际标准
- ◆ 第7章 多媒体通信技术最新进展



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想： MPEG-4(1999)
  - 视频质量可伸缩性 & H.264 SVC
  - 多视角编码

# MC 视频编码的思路

## ◆ 现有编码技术可以分为三类

### □ 第一类方法基于图像统计特性

- 预测编码(例如帧内预测, 运动估计与补偿等)
- 变换编码(例如傅立叶变换, 哈达玛变换, KL变换, 离散余弦变换, 离散小波变换等)
- 熵编码(例如算术编码, 霍夫曼编码, 游程编码)

### □ 第二类方法基于人眼视觉特性

- 方向滤波编码
- 面向轮廓一纹理特征的编码

### □ 第三类方法基于图像内容特征

- 分形编码
- 基于对象的编码

◆ 后两类方法由于复杂度过高, 技术尚未成熟等原因, 仍处于研究阶段。而第一类方法则得到不断地发展和融合, 形成了**基于块的预测/变换混合编码**框架。



# 问题的提出：视频与图像编码的差异

## 帧间冗余是最大的冗余



Frame 1

Frame 2

Frame 3

Frame 4



# 视频编码器原理

## JPEG

e.g., DCT:  
spatial → frequency

common code words →  
shorter symbols  
Huffman, arithmetic coding

Frames of  
Digital Video

**Motion  
Estimation  
&  
Compensation**

**Transform,  
Quantization, Zig-  
Zag Scan & Run-  
Length Encoding**

**Symbol  
Encoder**

Bit Stream

**predict current  
frame from previous**

*Quantization* changes representation  
size for each symbol  
→ adjust rate/quality trade-off

*Run-length encoding:*  
long runs of zeros → run-length symbol

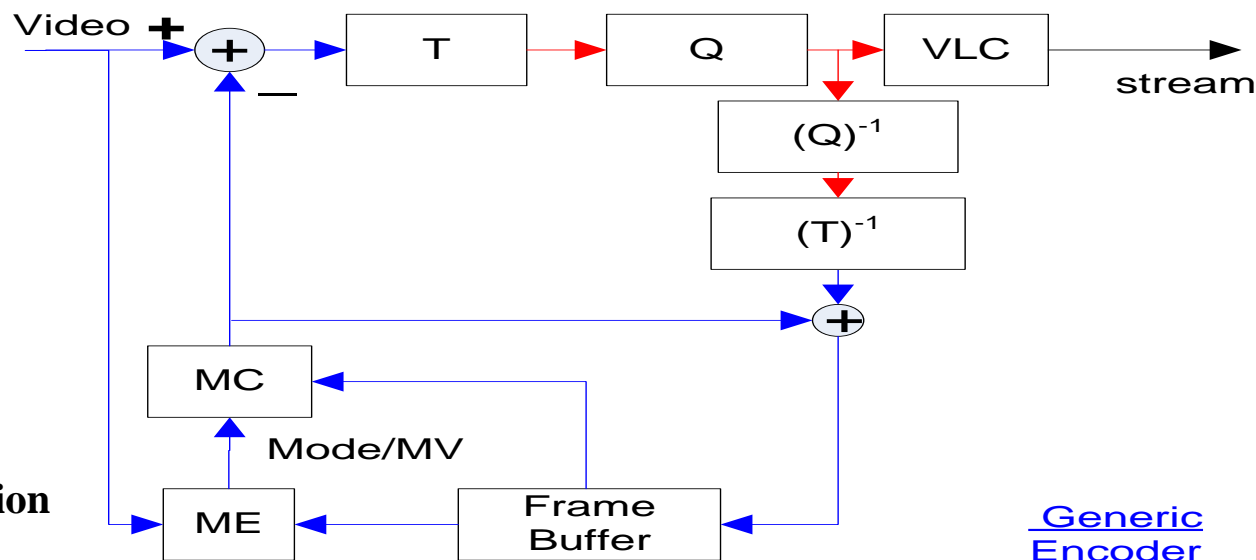
MPEG, H.26x  
6

# MC 视频编解码器框架

1979年国际图像编码学会上提出

◆ 自1988年H.261开始

- Temporal      ← Prediction
- Spatial        ← Transform
- Entropy        ← VLC



ME: Motion Estimation  
MC: Motion Compensation

Generic Encoder

基于块的预测/变换混合编码框架



# 20年来图像和视频压缩标准的进展

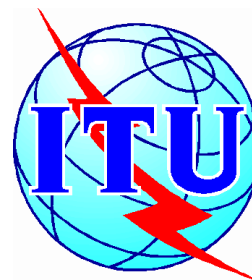
## ◆ 图像

- ◆ JPEG, 1992
- ◆ JPEG-LS(1998)
- ◆ JPEG 2000, 2000
- ◆ JPEG XR, 2009

## ◆ 视频

- ◆ H.261, 1988
- ◆ MPEG-1, 1993
- ◆ MPEG-2/H.262, 1996
- ◆ H.263, 1996
- ◆ MPEG-4, 1999
- ◆ **MPEG-4 AVC/H.264, 2003**
- ◆ **HEVC, 2013年4月13号**

## ◆ China AVS, 2003





# 图像和视频编码的标准化组织

- ◆ JPEG (Joint Photographic Expert Group, 1986)
  - Joint committee between ISO/IEC JTC1 and ITU-T
  - ISO/IEC Joint Technical Committee 1, Subcommittee 29, Working Group 1 (ISO/IEC JTC 1/SC 29/WG 1) - Coding of still pictures.
- ◆ MPEG (Moving Pictures Experts, 1988)
  - ISO/IEC JTC1/SC29 WG11 – Coding of moving pictures and audio (ISO/IEC Joint Technical Committee 1, Subcommittee 29, Working Group 11).





# 视频编码的标准化组织

- ◆ **ISO/IEC MPEG** (Moving Pictures Experts, 1988)
  - ISO/IEC JTC1/SC29 WG11 – Coding of moving pictures and audio (ISO/IEC Joint Technical Committee 1, Subcommittee 29, Working Group 11).
- ◆ **ITU-T VCEG** (Video Coding Experts Group)
  - (ITU-T SG16/Q6 = International Telecommunications Union – Telecommunications Standardization Sector (ITU-T, a United Nations Organization, formerly CCITT), Study Group 16, Working Party 3, Question 6)



# 视频编码的标准化组织(2000~)

## ◆ JVT (Joint Video Team, 2001)

- Joint project between ITU-T SG16/Q.6 (Study Group 16 / Question 6) – VCEG (Video Coding Experts Group) and ISO/IEC JTC1/SC29/WG11 – MPEG.

## ◆ JCT-VC (Joint Collaborative Team on Video Coding, 2010)

- A group of video coding experts from ITU-T Study Group 16 (VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (MPEG)
- It was created to develop **High Efficiency Video Coding**.

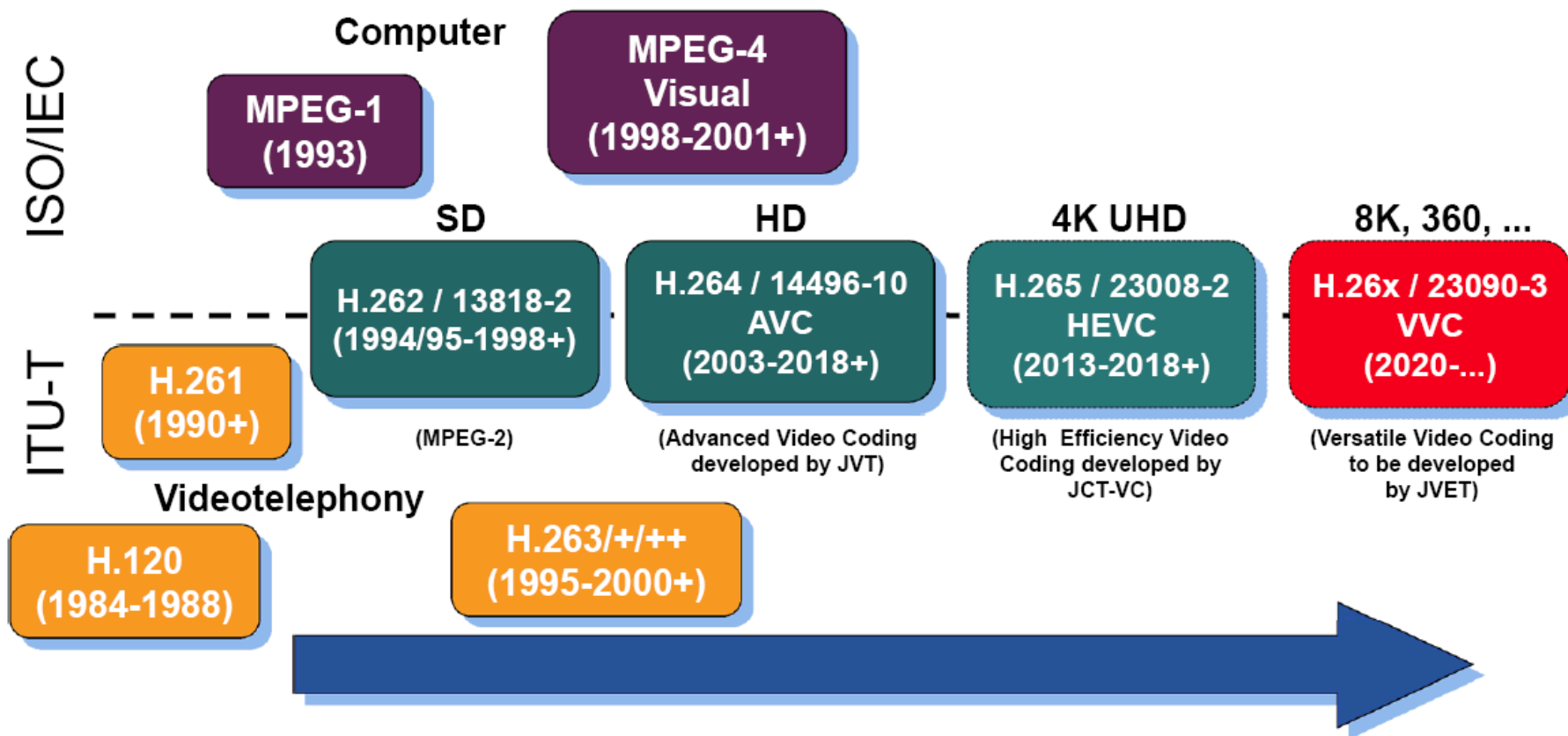
## ◆ JVET (Joint Video Experts Team, 2017)

- A new group between VCEG (Q6/16) and ISO/IEC JTC1 SC29/WG11 (MPEG)
- to develop **Versatile Video Codec (VVC)**



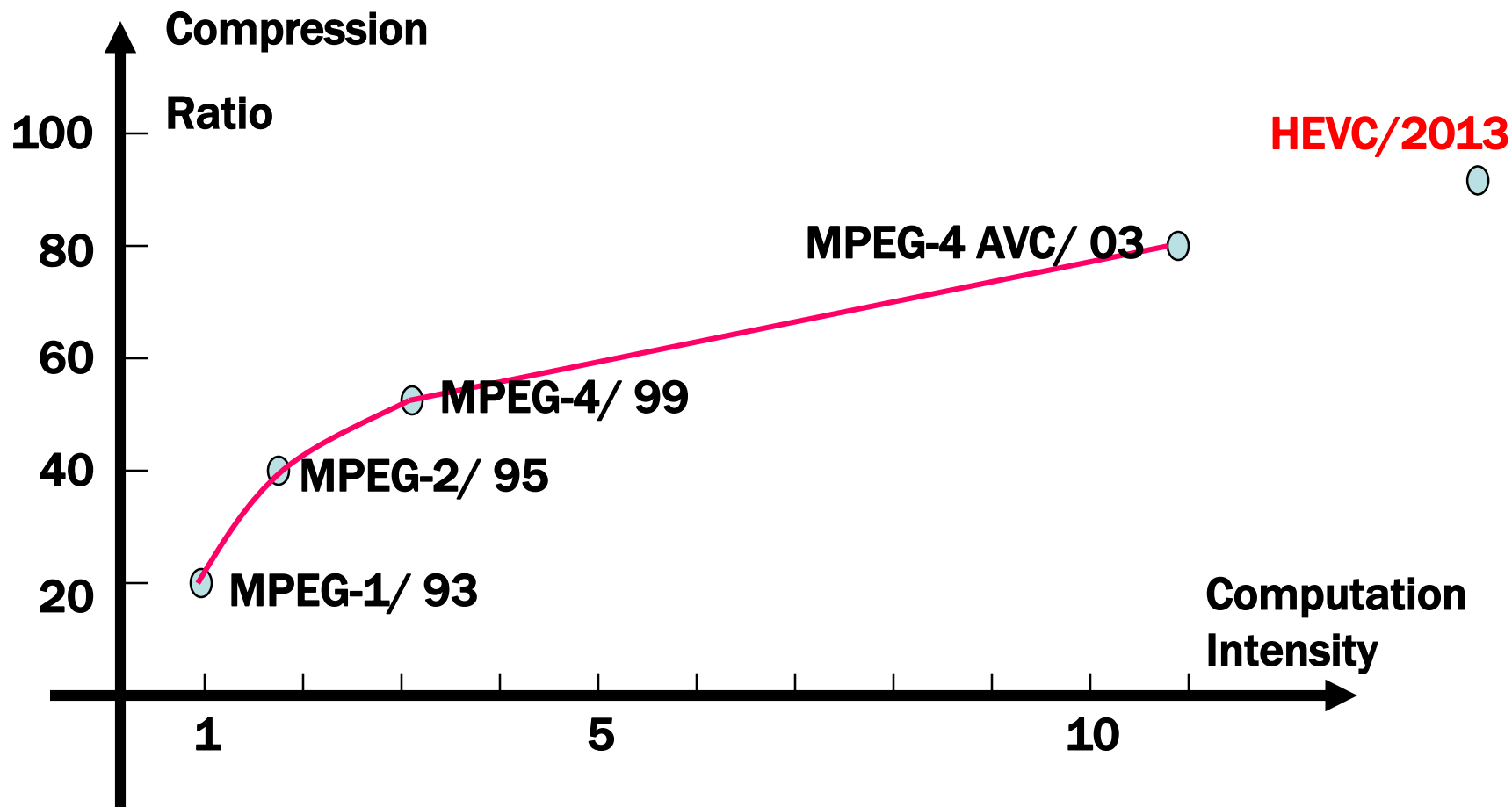
# 视频压缩标准(1984~2020)

## History of international video coding standardization (1985 ~ 2020)





# MPEG系列压缩算法性能(1993~)



# MC MPEG相关标准

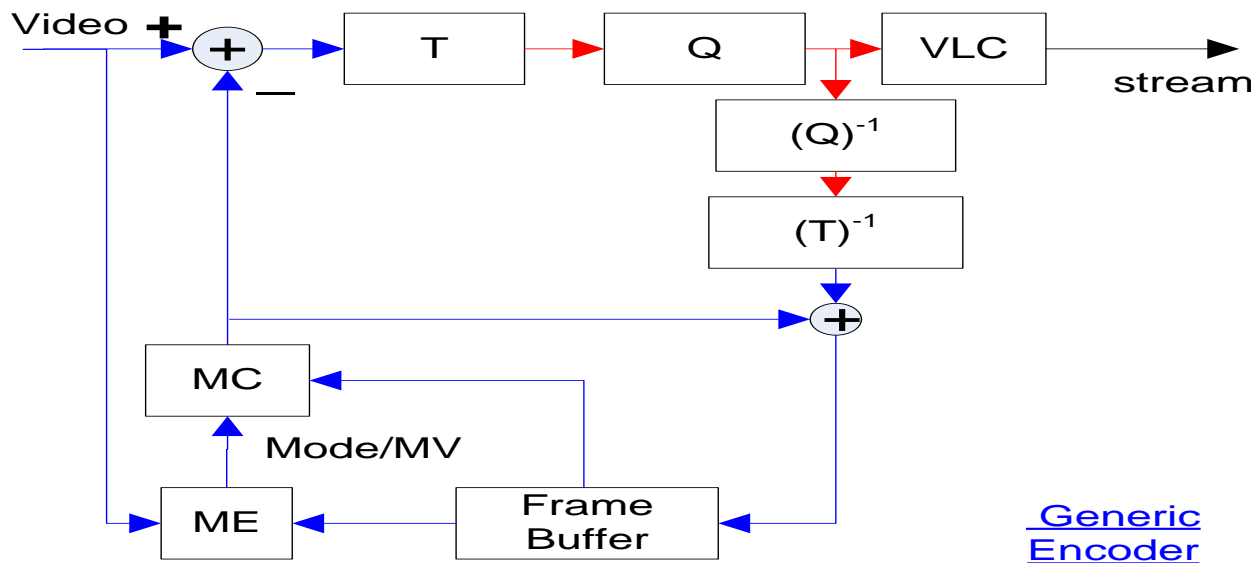
- ◆ MPEG-1: 在VCD中被采用，其中MPEG-1 Layer3简称MP3。
- ◆ MPEG-2: 广播质量的视频、音频和传输协议。被用于无线数码电视-ATSC、DVB以及ISDB、数字卫星电视、数字有线电视信号，以及DVD视频光盘技术中。
- ◆ MPEG-3: 原本目标是为HDTV设计，随后发现MPEG-2已足够HDTV应用，故MPEG-3的研发便中止。
- ◆ MPEG-4: 扩展MPEG-1、MPEG-2等标准以支持视频/音频物件（video/audio "objects"）的编码、3D内容、低位元率编码（low bitrate encoding）和数码版权管理。
- ◆ MPEG-7: MPEG-7并不是一个视频压缩标准，它是一个多媒体内容的描述标准。
- ◆ MPEG-21: MPEG-21是一个正在制定过程中的标准，它的目标是为未来多媒体的应用提供一个完整的平台。



# 小结：基于块的预测/变换混合编码框架

- ◆ MPEG-1(1993)
- ◆ MPEG-2/H.262(1996)
- ◆ MPEG-4 AVC/H.264(2003)
- ◆ HEVC(2013)

**Temporal** ← **Prediction**  
**Spatial** ← **Transform**  
**Entropy** ← **VLC**



Generic Encoder



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想： MPEG-4(1999)
  - 视频质量可伸缩性 & H.264 SVC
  - 多视角编码

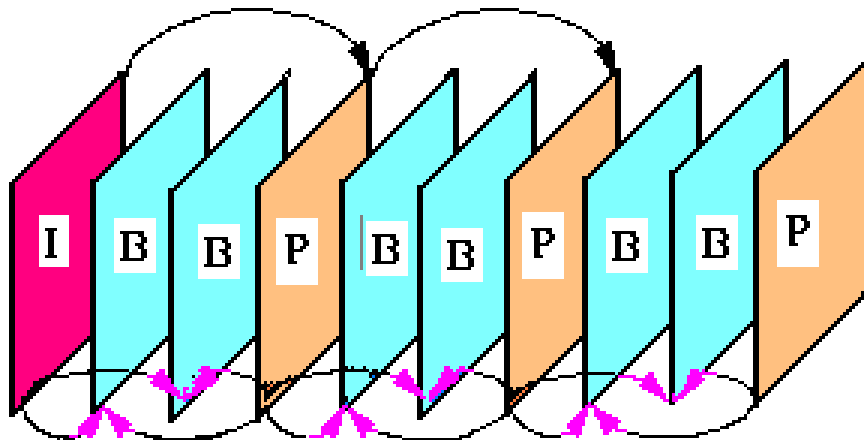


# 视频中帧间冗余是最大的冗余

## MPEG1、2中的三种图象帧

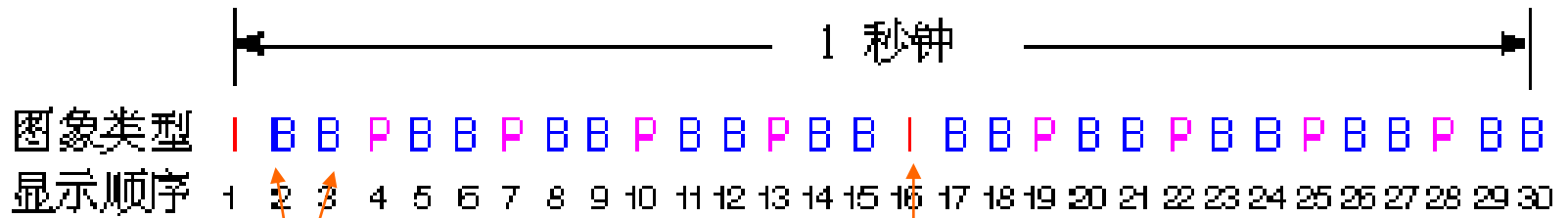
MPEG专家组定义了三种图像：

- ◆ 帧内图像I(intrapictures)
- ◆ 预测图像P(predicted pictures)
- ◆ 插补（双向预测）图像B(bidirectionally interpolated)



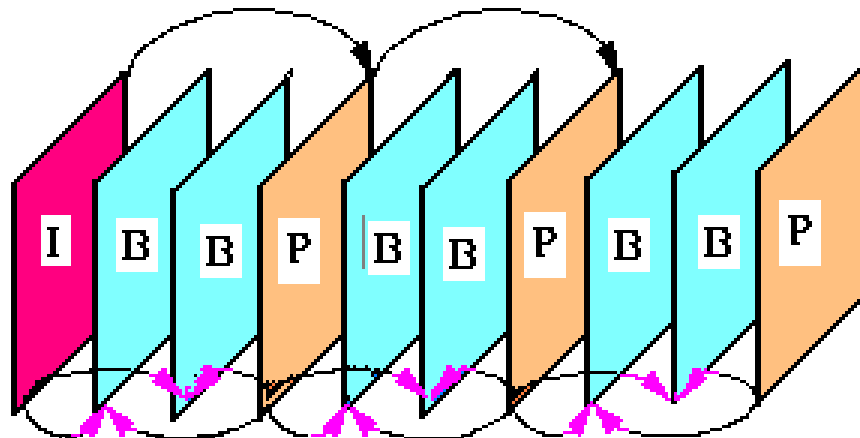


# MPEG Video 典型流结构



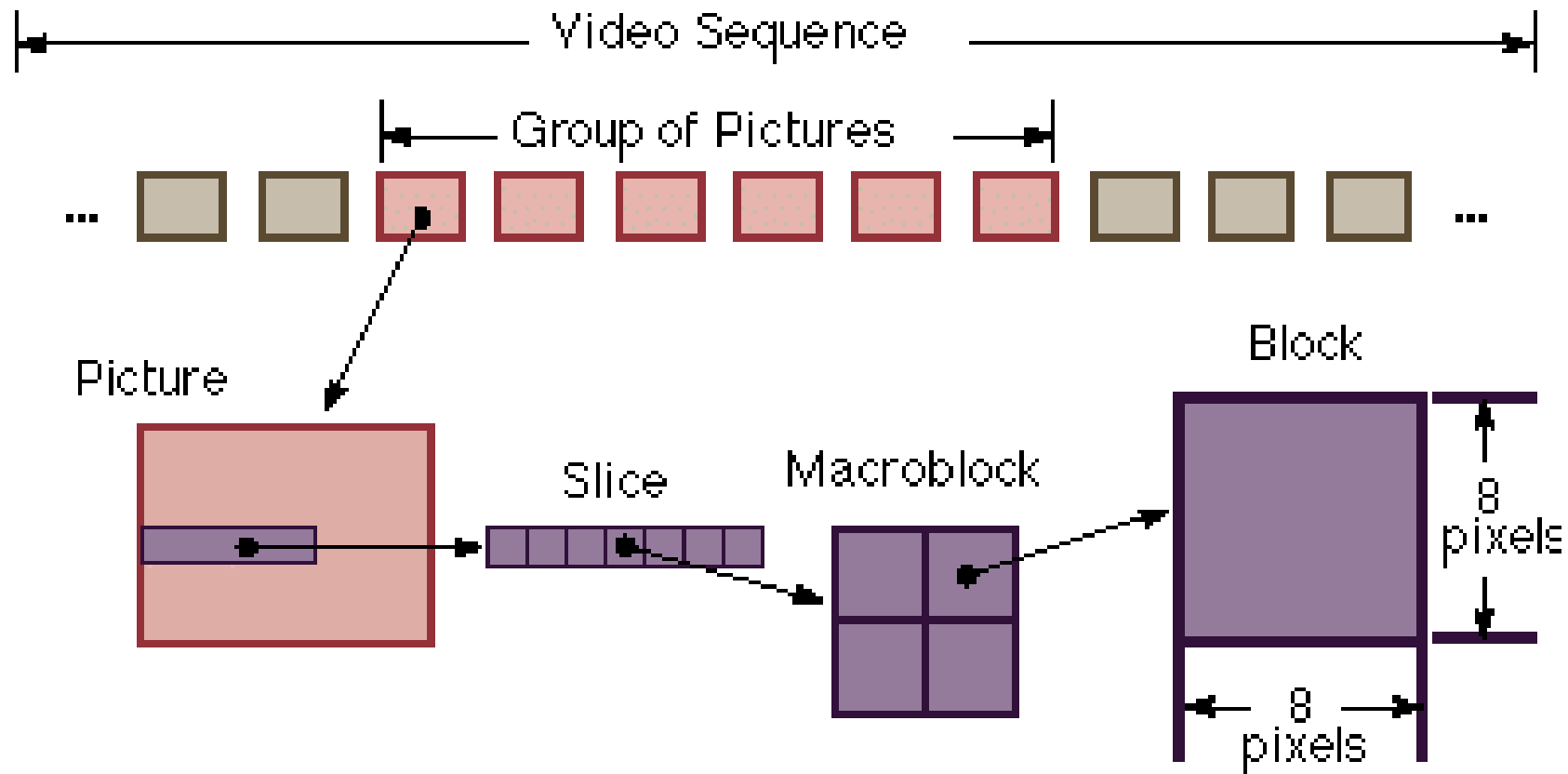
两幅B位于两个参考图像间

每15帧插入一幅I



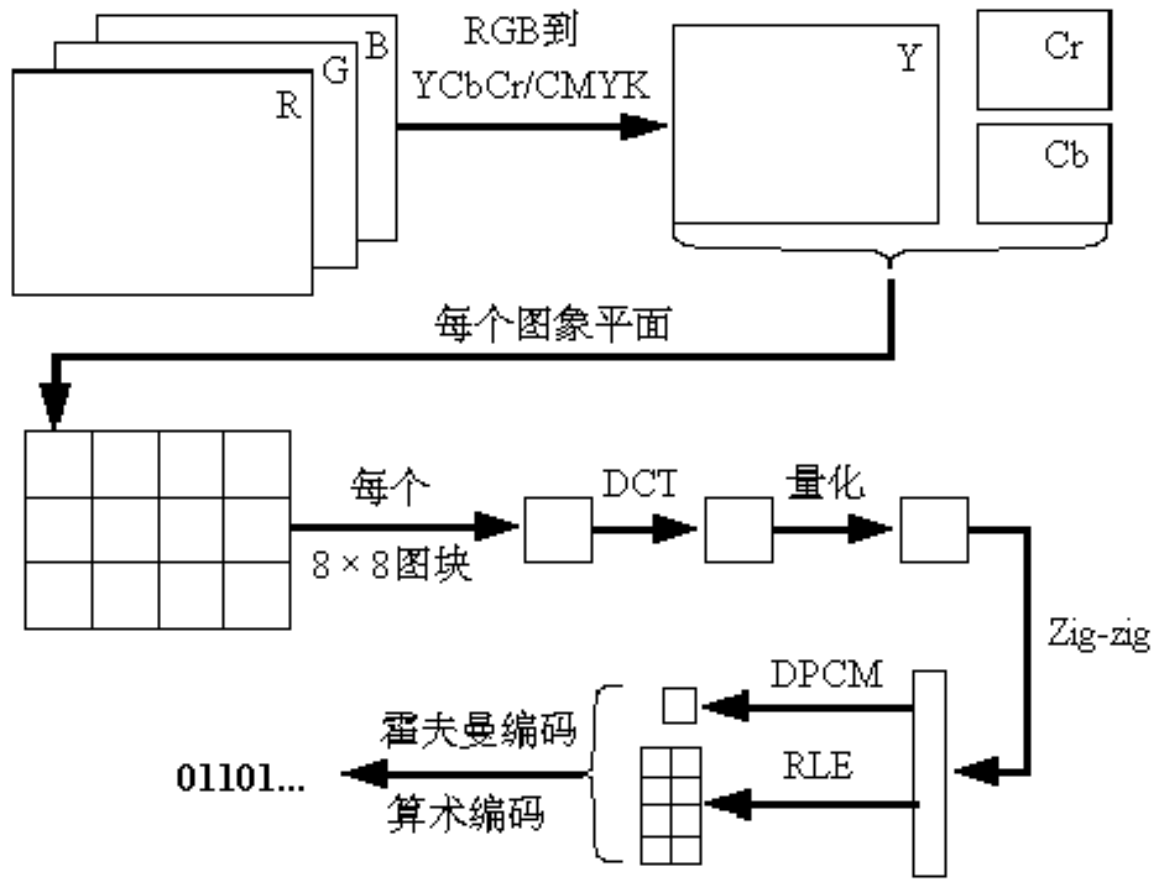
# MC 视频序列的组成

◆ GOP → Picture → Slice → Macroblock → Block





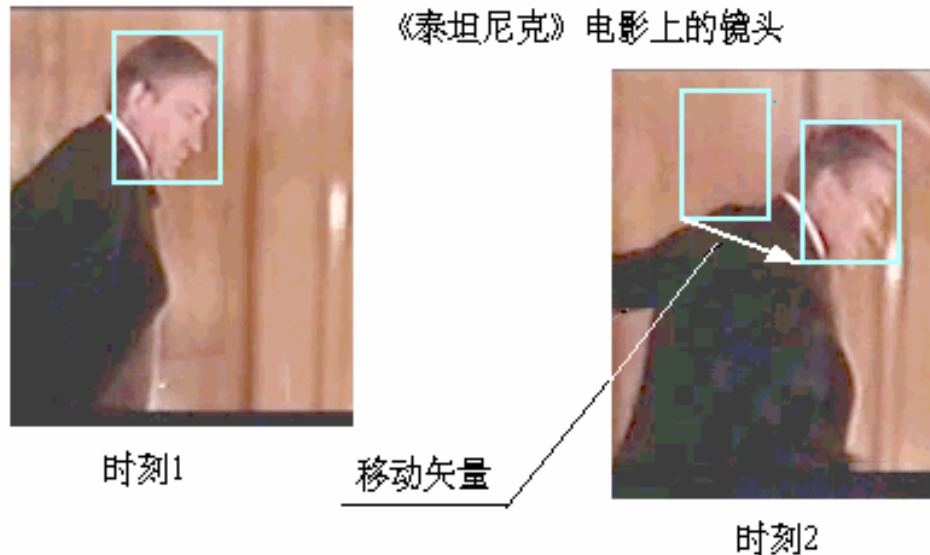
# 帧内图像I的压缩编码算法





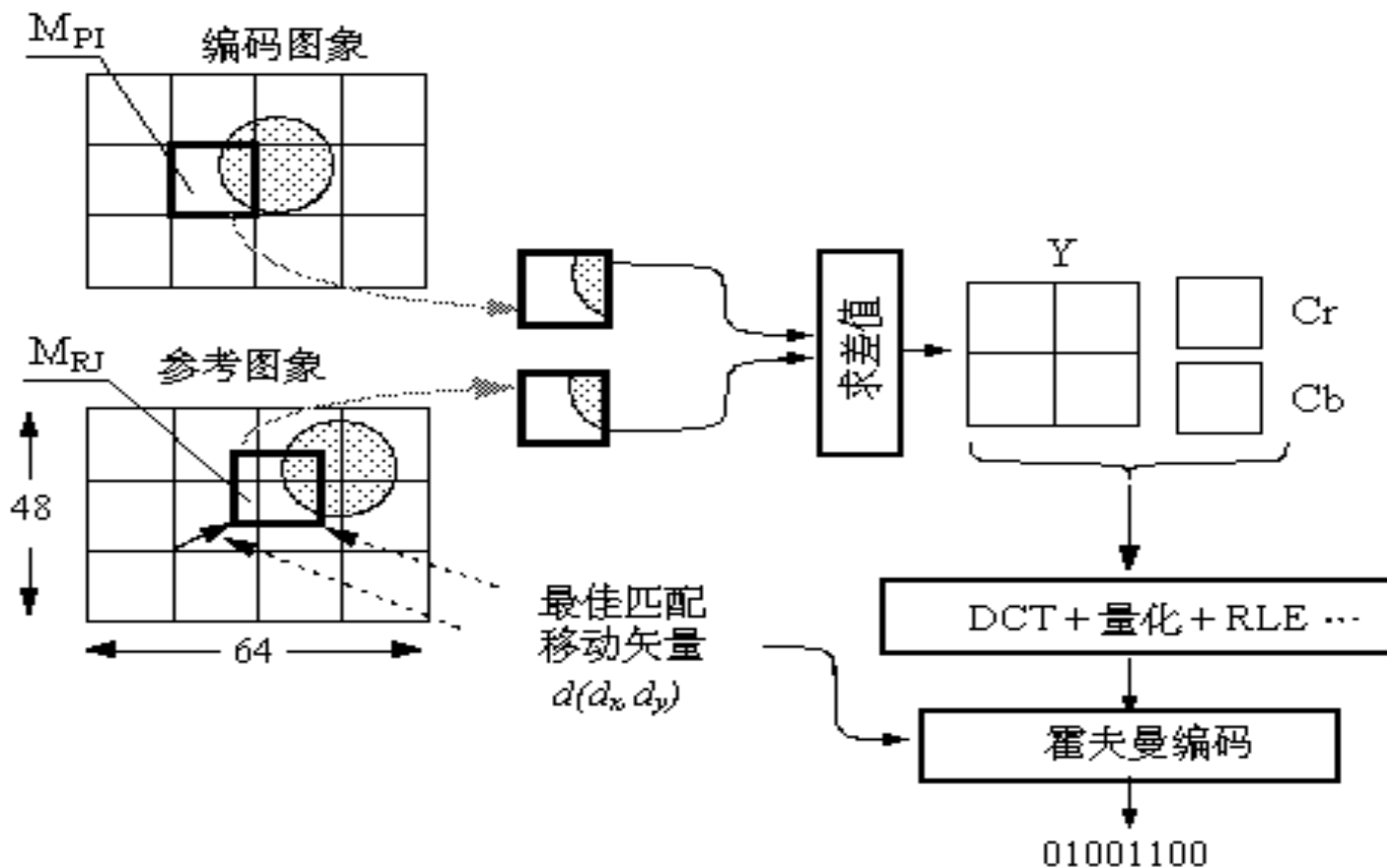
# 预测图像P的压缩编码算法

- ◆ 预测图像P使用两种类型的参数来表示：
  - 一种参数是当前要编码的图像宏块与参考图像的宏块之间的**差值**
  - 另一种参数是宏块的**移动矢量**

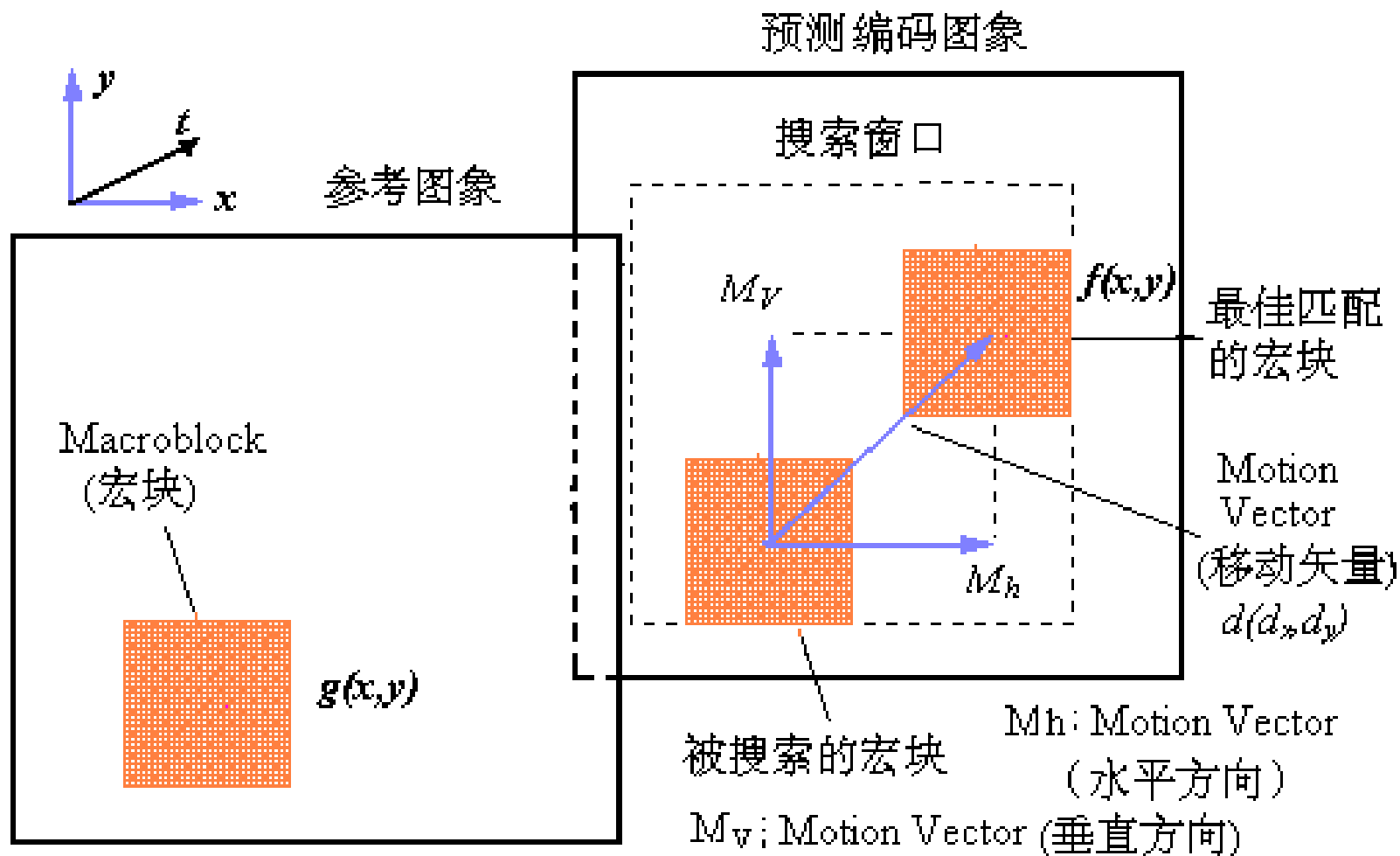


# MC 差值的记录

◆ 假设宏块 $M_{PI}$ 是参考宏块 $M_{RJ}$ 的最佳匹配块，它们的差值就是这两个宏块中相应像素值之差。



# MC 移动矢量算法原理



# MC 块匹配法

块匹配法（Block Matching Algorithm）需要解决两个关键问题：

## ◆ 匹配准则

- 绝对值最小
- 均方误差最小
- 平均绝对帧差最小
- .....

## ◆ 搜索算法

- 穷举式搜索
- 三步搜索法
- 二维对数搜索法
- orthogonal search
- 对偶搜索法
- .....

# MC 最佳匹配的原则

- ◆ 绝对值AE(absolute difference)最小

$$AE = \sum_{i=0}^{15} \sum_{j=0}^{15} |f(i, j) - g(i - d_x, j - d_y)|$$

- ◆ 均方误差MSE(mean-square error)最小

$$MSE = \frac{1}{I \times J} \sum_{|i| \leq \frac{I}{2}} \sum_{|j| \leq \frac{J}{2}} [f(i, j) - g(i - d_x, j - d_y)]^2$$

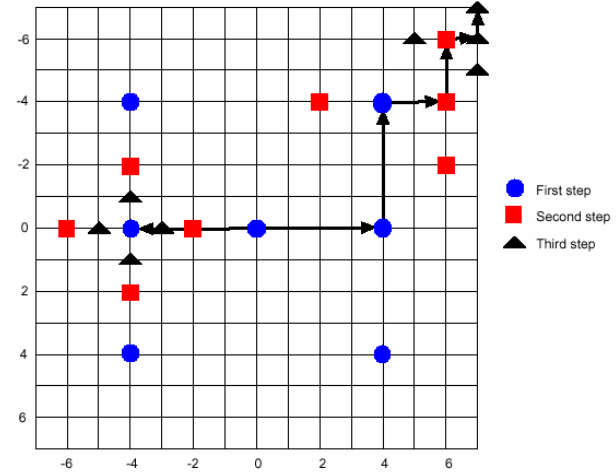
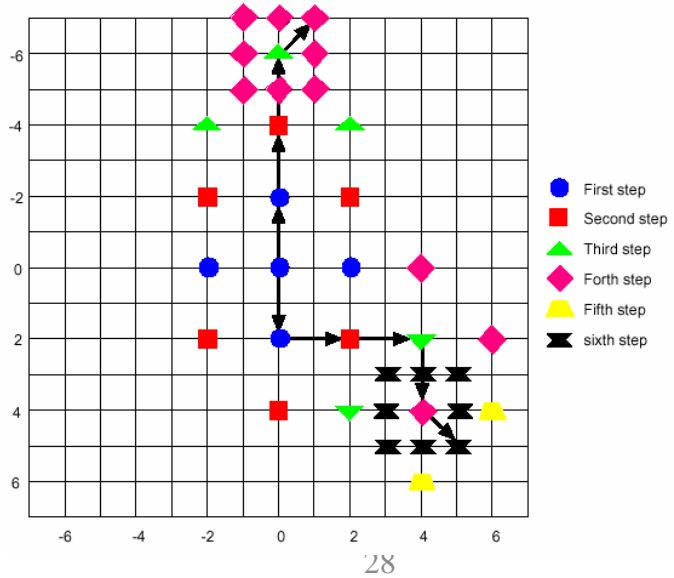
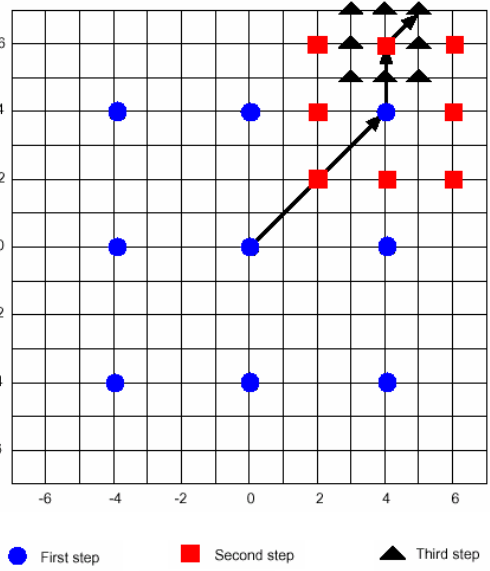
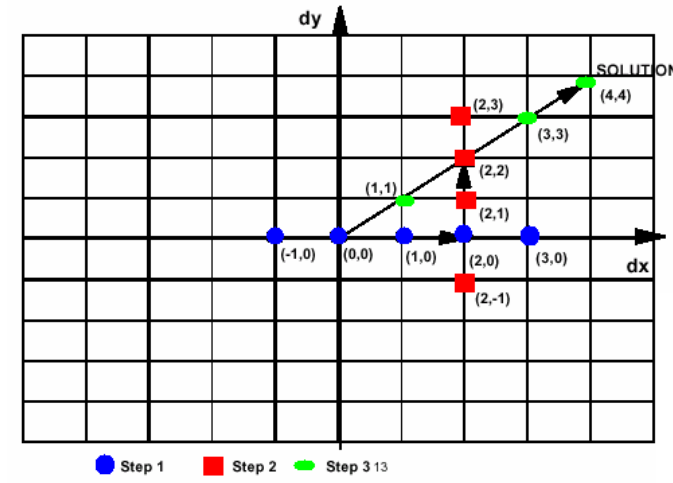
- ◆ 平均绝对帧差MAD(mean of the absolute frame difference)最小

$$MAD = \frac{1}{I \times J} \sum_{|i| \leq \frac{I}{2}} \sum_{|j| \leq \frac{J}{2}} |f(i, j) - g(i - d_x, j - d_y)|$$

dx和dy分别是参考宏块MRJ的移动矢量d(dx, dy)在X和Y方向上的矢量

# MC 移动矢量的估计

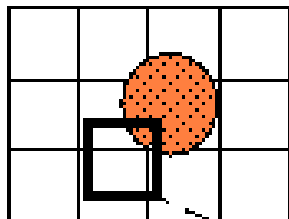
- ◆ 穷举式搜索(Exhaustive Search)
- ◆ 三步搜索法(three-step search)
- ◆ 二维对数搜索法(2D-logarithmic search)
- ◆ orthogonal search algorithm (OSA)
- ◆ 对偶搜索法(conjugate search)



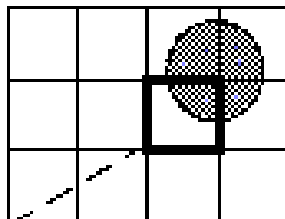


# 插补图像B的压缩编码算法

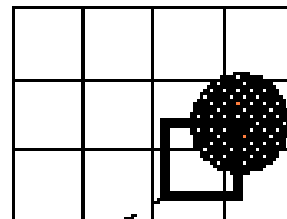
过去的参考图像



编码图像



将来的参考图像



$$\left[ \begin{array}{c} \text{过去的参考图像} \\ \text{编码图像} \end{array} \right] - 0.5 \times \left[ \begin{array}{c} \text{过去的参考图像} \\ \text{编码图像} \end{array} \right] + \left[ \begin{array}{c} \text{将来的参考图像} \\ \text{编码图像} \end{array} \right] = \left[ \begin{array}{c} \text{过去的参考图像} \\ \text{编码图像} \end{array} \right]$$

DCT + 量化 + RLE

移动矢量

霍夫曼编码

011010...

# MC 小结:

## ◆ GOP(Group Of Picture)

- I帧、P帧、B帧

## ◆ I帧

- JPEG: DCT→Huffman→RLE

## ◆ 预测帧

- 宏块差值: 匹配的原则

- 宏块的移动矢量: 各类搜索方法、时域/频域



# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
  - H.264编码框架概述
  - H.264视频编码的技术创新点
  - H.264不同Profile的特点简介
- HEVC(2013)视频标准

## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想： MPEG-4(1999)
- 视频质量可伸缩性 & H.264 SVC
- 多视角编码



# 进一步视频/图像冗余及去除方法 H.264的精华?

- ◆ Temporal
  - Prediction
- ◆ Spatial
  - Transform: DCT/Wavelet
- ◆ Entropy
  - Huffman/VLC/...

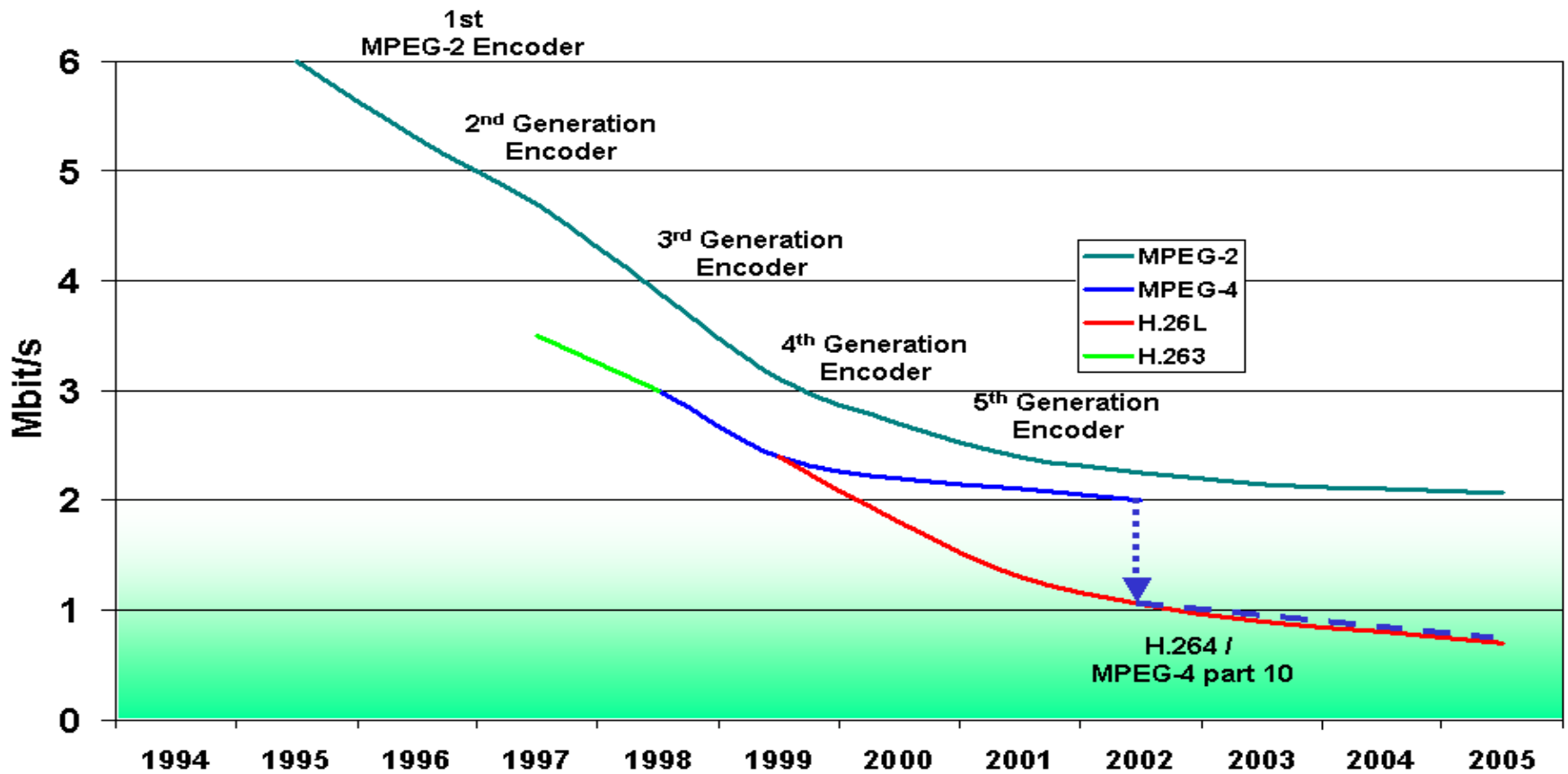


# H.264/MPEG-4 AVC

- ◆ H.264是ITU-T以H.26x系列为名称命名的标准之一，同时AVC是ISO/IEC MPEG一方的称呼。
- ◆ 这个标准通常被称之为H.264/AVC (或者 AVC/H.264 或者 H.264/MPEG-4 AVC or MPEG-4/H.264 AVC)而明确的说明它两方面的开发者。
- ◆ 该标准最早来自于ITU-T的称之为H.26L的项目的开发。有时候该标准也被称之为"JVT 编解码器"，这是由于该标准是由JVT组织并开发的。



# H.264的地位





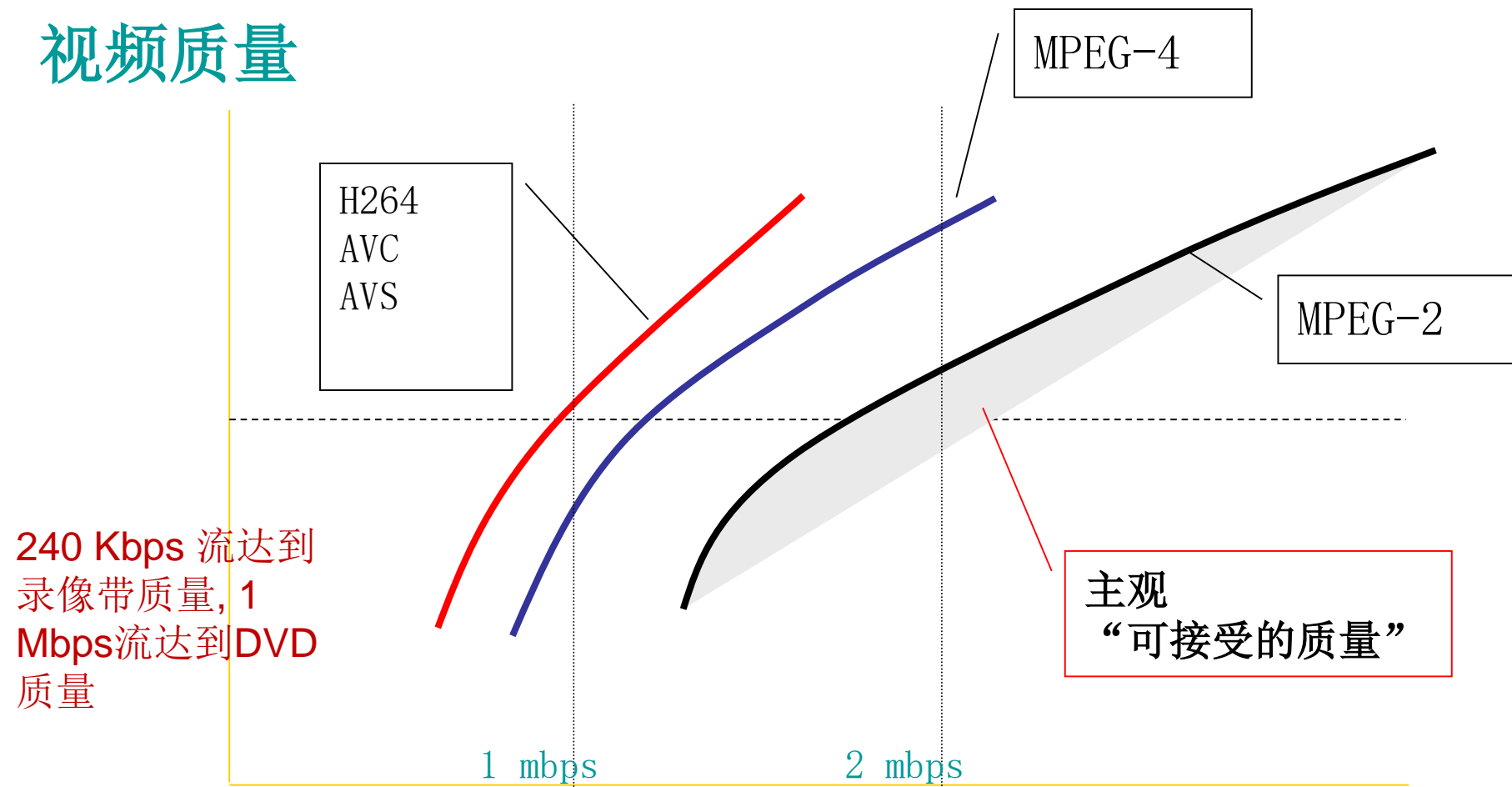
# H.264与MPEG-2/MPEG-4比较

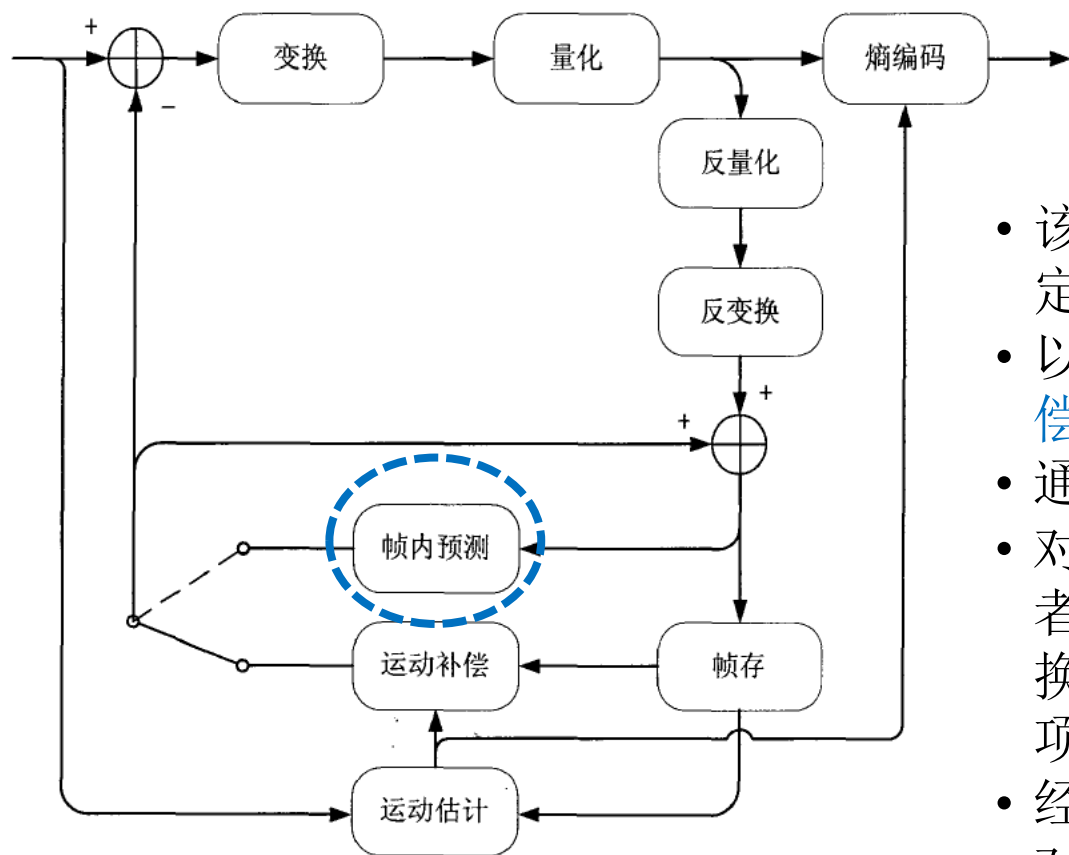
主要技术	H.264	MPEG-2	MPEG-4
帧内预测	4×4 像素以 9 种, 16×16 像素块 4 种预测模式	无	无
帧间预测	分 16×16 像素块 7 种模式, SDTV 图像最大预测 5 帧 (预测 1/4 像素)	以 16×16 像素块为单位从前面帧预测 (预测 1/2 像素)	以 16×16 和 8×8 像素块为单位从前面帧预测 (预测 1/4 像素)
变换	4×4 像素单位的整数变换	8×8 像素单位的离散余弦变换	
纠错编码	CAVLC 和 CABAC	霍夫曼编码	
环路滤波器	有	无	无



# H.264的视频质量

## 视频质量



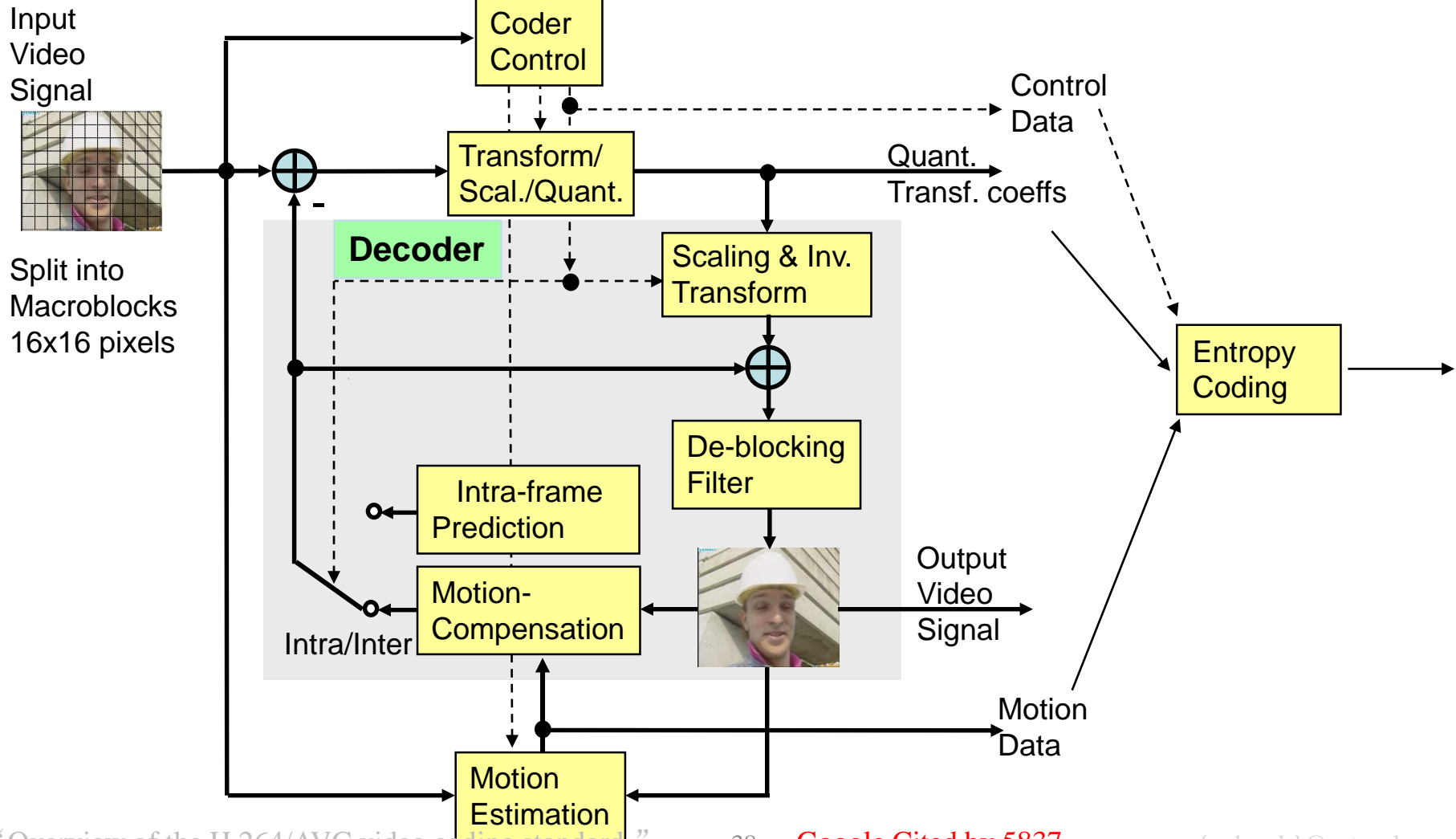


基于块运动补偿和变换编码的混合编码框架**1979年**提出

- 该编码框架中，将编码帧划分为固定大小的块，**以块为单位**编码。
- 以块为单位进行**运动估计**和**运动补偿**，去除视频序列内的时间冗余
- 通过**帧内预测**去除空间冗余
- 对运动估计，补偿后的残差数据或者帧内预测后的残差数据，进行变换（将能量集中到变换域的少数几项系数上）
- 经过变换编码的残差数据进行量化
- 对运动矢量(Motion Vector, MV)以及经过变换，量化后的残差数据进行熵编码



# H.264编码器框架





# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
  - H.264编码框架概述
  - H.264视频编码的技术创新点
  - H.264不同Profile的特点简介
- HEVC(2013)视频标准

## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想： MPEG-4(1999)
- 视频质量可伸缩性 & H.264 SVC
- 多视角编码

## ◆ Motion Compensation

- 1/4象素精度的运动估计
- 7种大小不同的块进行匹配
- 前向与后向多参考帧

## ◆ Intra-frame Prediction

- 包括9种 $4 \times 4$ 亮度块的预测、4种 $16 \times 16$ 亮度块的预测和4种色度块的预测

## ◆ De-blocking Filter

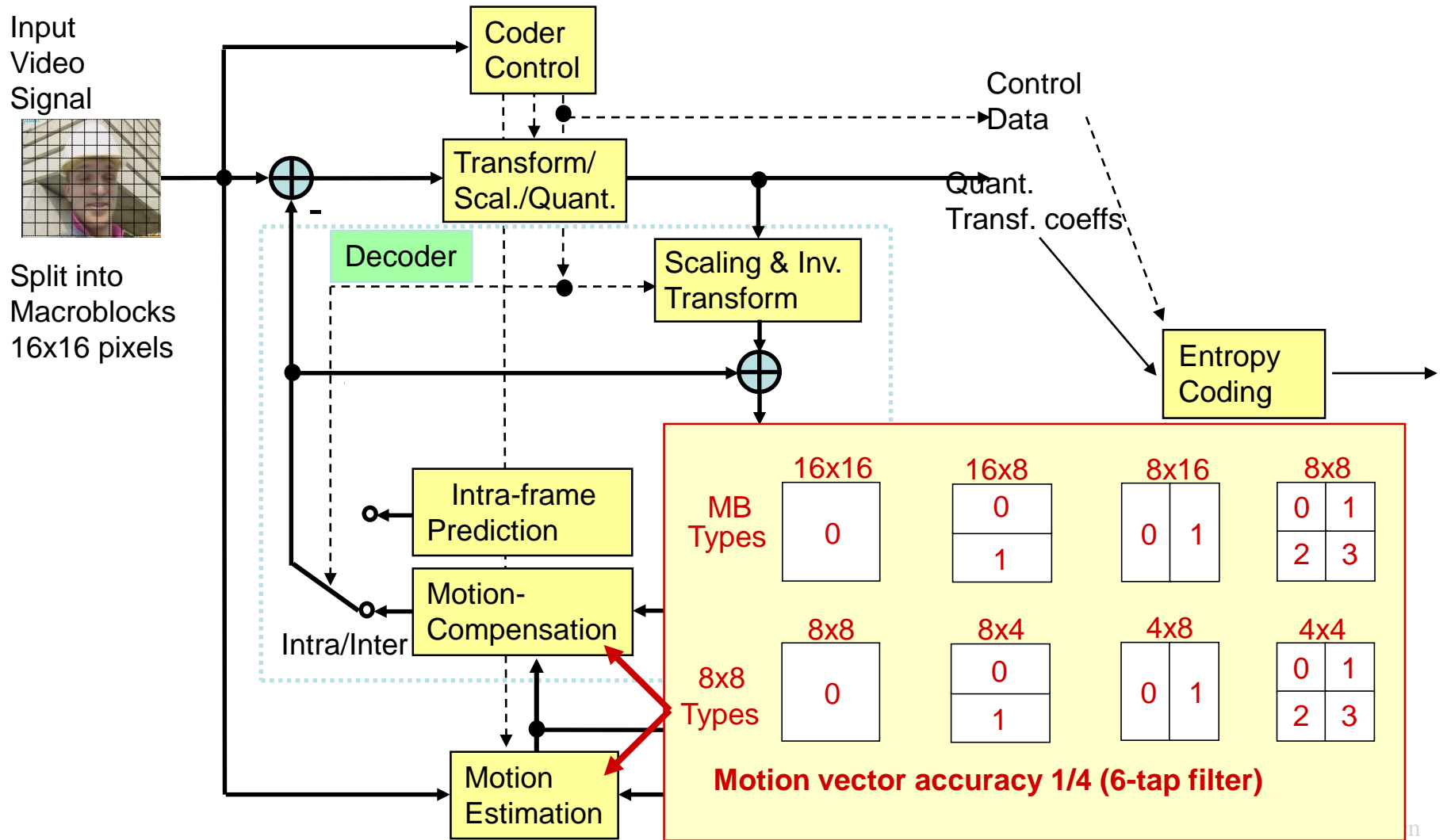
- 优秀的运动补偿滤波器

## ◆ Transform/Scal./Quant.

## ◆ Entropy Coding

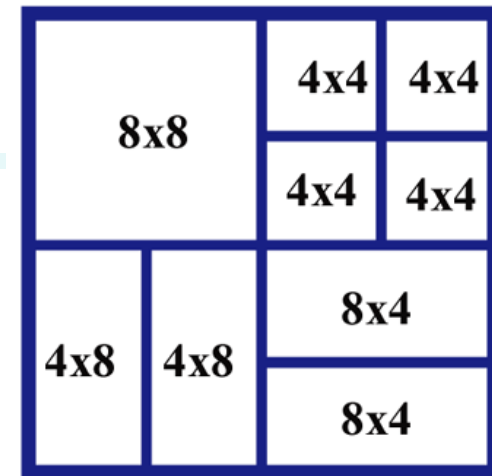


# Motion Compensation Accuracy

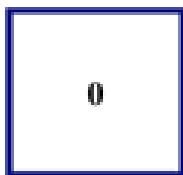


# MC 宏块大小

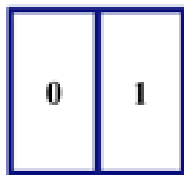
可以对不同大小和形状的宏块进行运动估计和补偿，从而提高了对运动矢量估计的精度



Mode 1  
One 16x16 blk  
One motion vector



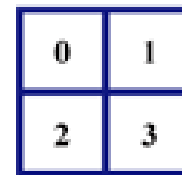
Mode 2  
Two 8x16 blks  
Two motion vectors



Mode 3  
Two 16x8 blks  
Two motion vectors



Mode 4  
Four 8x8 blks  
Four motion vectors



Mode 5  
Eight 4x8 blks  
Eight motion vectors



Mode 6  
Eight 8x4 blks  
Eight motion vectors

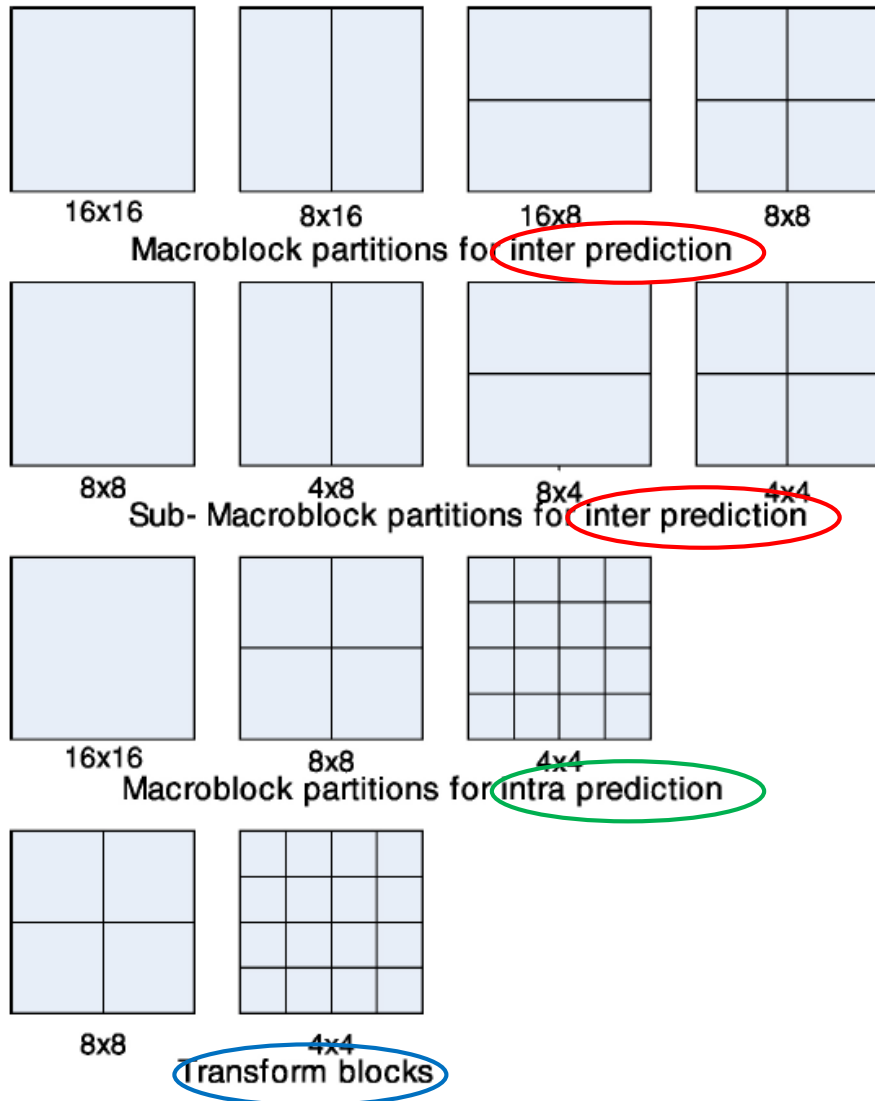


Mode 7  
Sixteen 4x4 blks  
Sixteen motion vectors





# H.264树状结构的运动补偿



- 宏块可进行进一步的分割，形成图示不同模式的小块。
- 对不同大小和形状（非正方形）的小块进行运动估计和补偿，从而提高了对运动矢量估计的精度。
- 对于图像中变化平缓的区域，采用较大的分块尺寸比较合适；对于细节丰富的区域，采用较小的尺寸比较合适。

预测块、变换块



# Why different partition sizes?

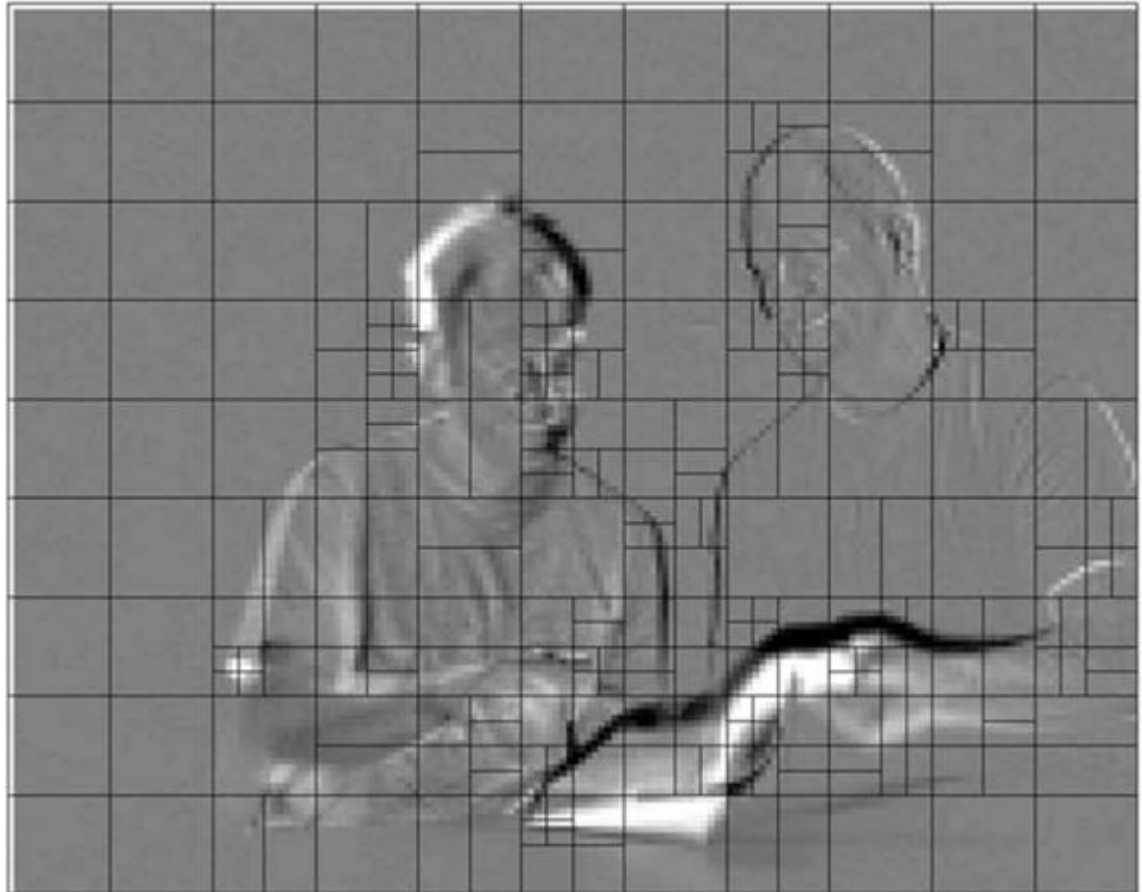
◆ MVs are expensive

◆ Smooth area

□ → large partition

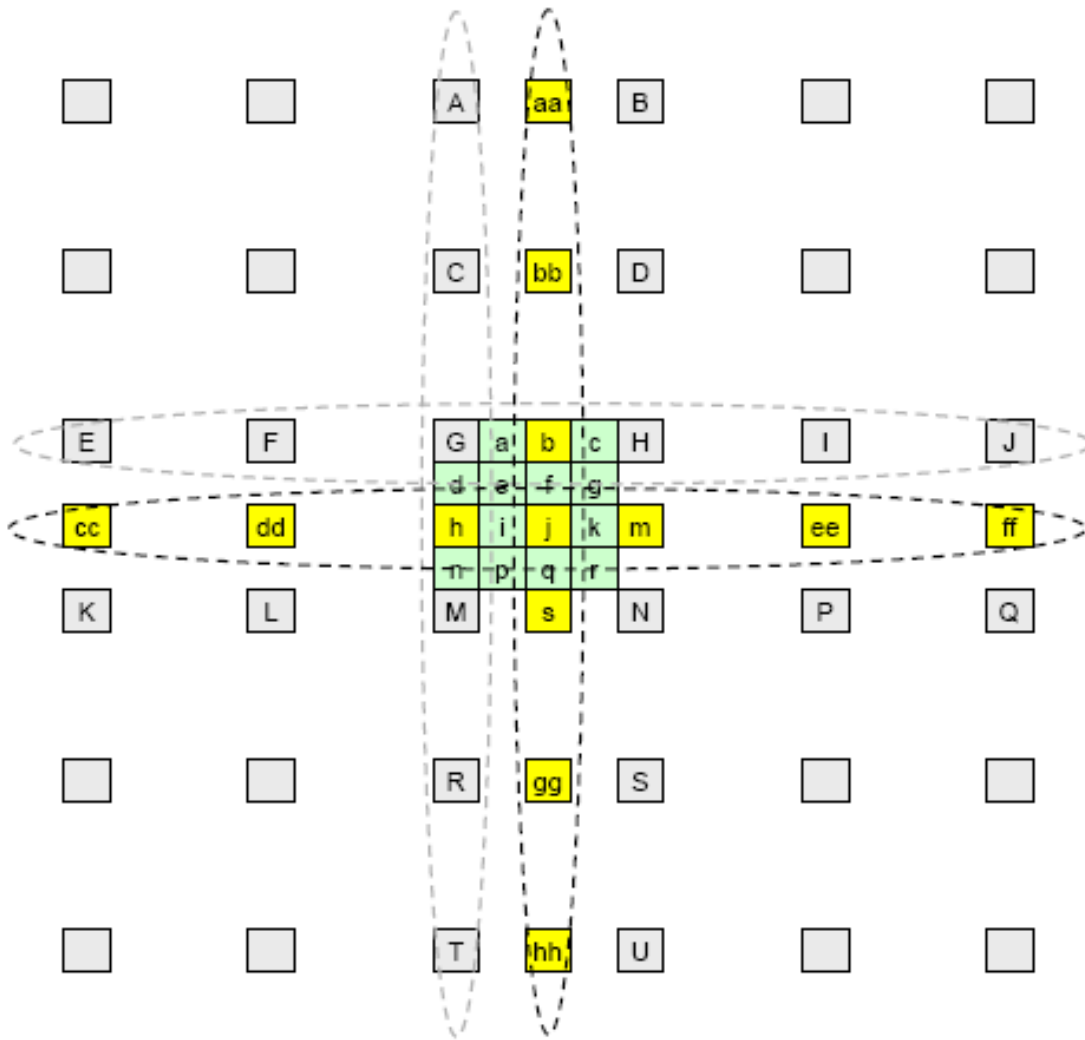
◆ Detailed area

□ → small partition



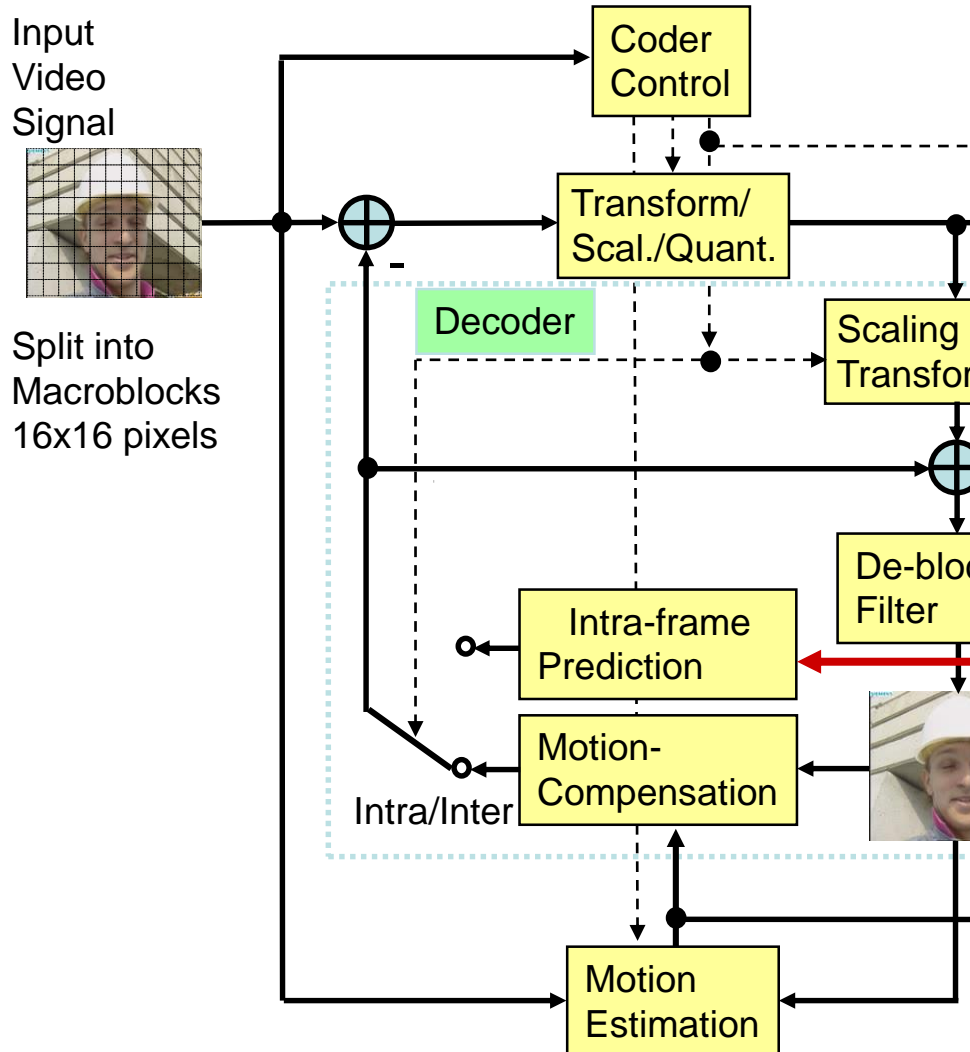


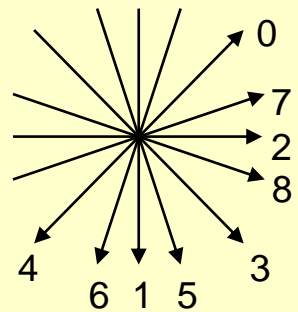
# 高精度的运动估计策略



- ◆  $\frac{1}{2}$ -pel pixel grid values are interpolated using a separable 6-tap filter
- ◆  $\frac{1}{4}$ -pel pixel grid values are interpolated using the average of two  $\frac{1}{2}$ -pel pixels

# MC Intra Prediction

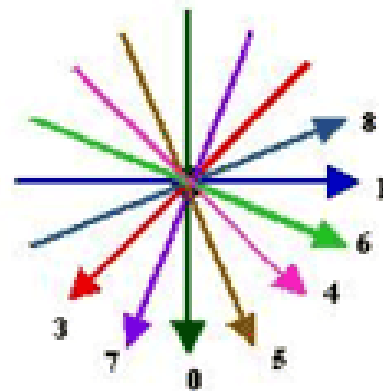


- Directional spatial prediction (9 types for luma, 1 chroma)
- |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|
| Q | A | B | C | D | E | F | G | H |
| I | a | b | c | d |   |   |   |   |
| J | e | f | g | h |   |   |   |   |
| K | i | j | k | l |   |   |   |   |
| L | m | n | o | p |   |   |   |   |
| M |   |   |   |   |   |   |   |   |
| N |   |   |   |   |   |   |   |   |
| O |   |   |   |   |   |   |   |   |
| P |   |   |   |   |   |   |   |   |
- 
- e.g., Mode 3: diagonal down/right prediction  
 $a, f, k, p$  are predicted by  $(A + 2Q + I + 2) \gg 2$

# MC 帧内预测的本质

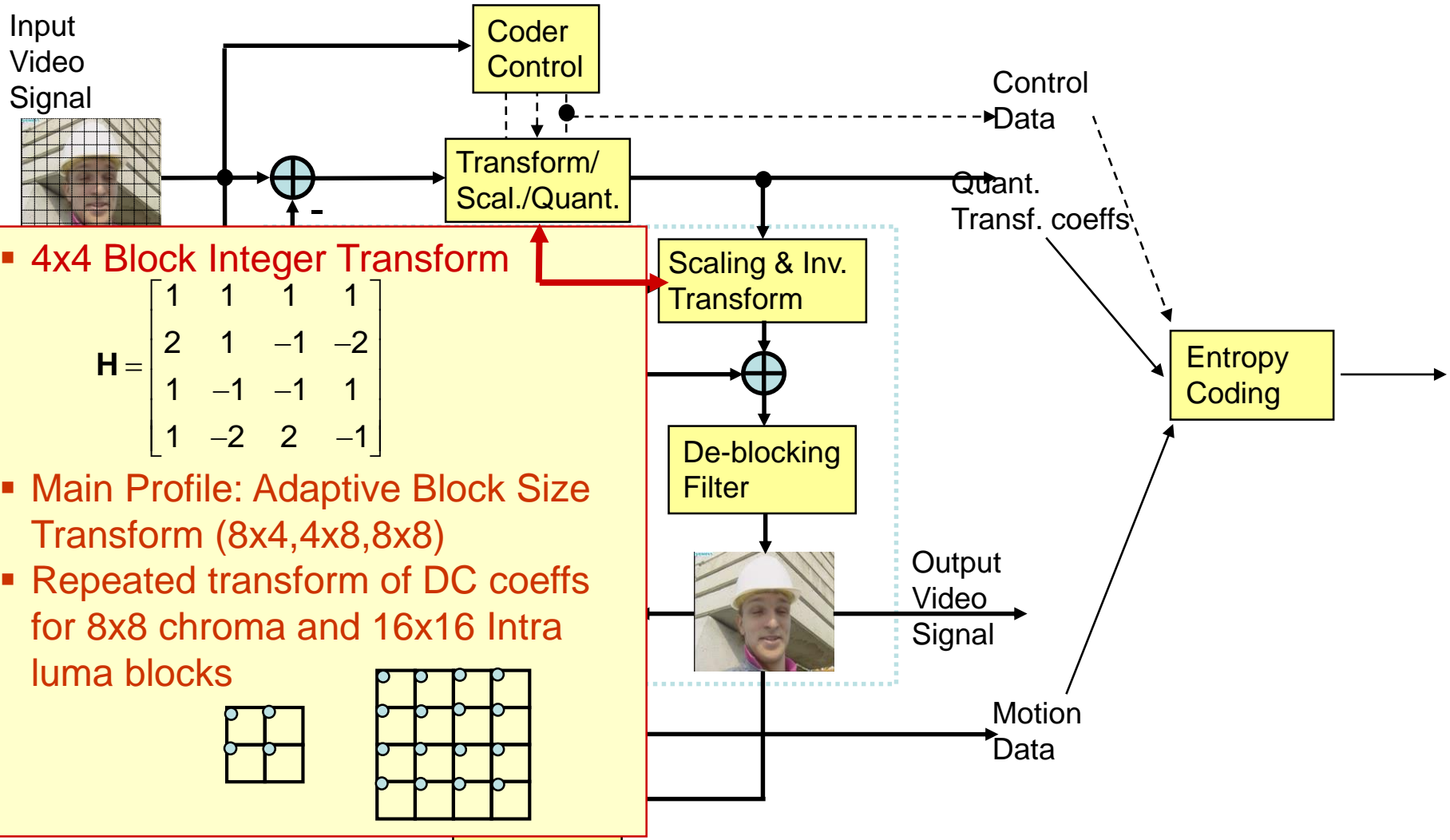
- ◆ 相邻像素之间通常具有很强的相关性
- ◆ 每个子块用其相邻的上方和左方已编码重建后的像素点来进行预测
- ◆ 后编码的子块若和已编码过的子块相似，则只需要记录预测值和实际值之间的差值

	M	A	B	C	D	E	F	G	H
I		a	b	c	d				
J		e	f	g	h				
K		i	j	k	l				
L		m	n	o	p				



4×4 亮度块的 9 种帧内预测模式示意图

# MC Transform Coding



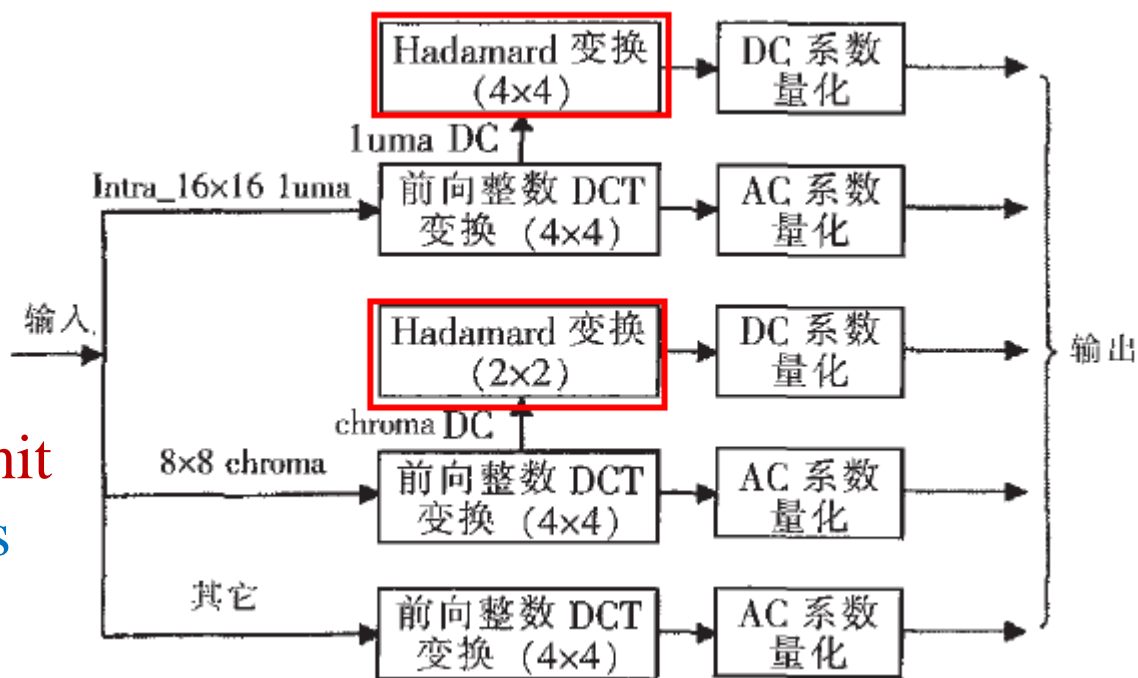


# Transform Coding

## 整数DCT变换、离散哈达玛变换

Intra\_16x16的亮度块，通过16(4x4)个4x4的前向DCT变换，得到16个DC系数再进行4x4的Hadamard变换，然后对16个DC系数和240个AC系数进行量化。

8x8的色度块（帧内、帧间），进行4x4的DCT变换后，得到4个DC和60个AC系数，同样对于DC系数先进行2x2的Hadamard变换后，分别对DC和AC系数进行量化。



Basic syntax & processing unit  
contains 16x16 luma samples  
and 2 x 8x8 chroma samples

Why?



# 52级步长的量化器

H.264草案规定亮度的QP最大值是51，色度QP的最大值是39。QP是量化参数，是量化步长的序号，QP每增加6，Qstep增加一倍。

一般量化器的原理为： $FQ = \text{round}(y/Q_{\text{step}})$ 。其中， $y$ 为输入样本点编码， $Q_{\text{step}}$ 为量化步长， $FQ$ 为 $y$ 的量化值。其相反过程即反量化为： $y' = FQ \cdot Q_{\text{step}}$ 。

QP	Qstep	QP	Qstep	QP	Qstep	QP	Qstep	QP	Qstep
0	0.625	12	2.5	24	10	36	40	48	160
1	0.6875	13	2.75	25	11	37	44	49	176
2	0.8125	14	3.25	26	13	38	52	50	208
3	0.875	15	3.5	27	14	39	56	51	224
4	1	16	4	28	16	40	64		
5	1.125	17	4.5	29	18	41	72		
6	1.25	18	5	30	20	42	80		
7	1.375	19	5.5	31	22	43	88		
8	1.625	20	6.5	32	26	44	104		
9	1.75	21	7	33	28	45	112		
10	2	22	8	34	32	46	128		
11	2.25	23	9	35	36	47	144		

H.264量化使用了52级步长的量化器，而H.263标准只有31级。



# 整数变换和量化的协同处理： 量化过程完成变换的部分运算

$$A = \begin{bmatrix} \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) & \frac{1}{2} \cos(0) \\ \sqrt{\frac{1}{2}} \cos(\frac{\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{3\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{5\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{7\pi}{8}) \\ \sqrt{\frac{1}{2}} \cos(\frac{2\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{6\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{10\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{14\pi}{8}) \\ \sqrt{\frac{1}{2}} \cos(\frac{3\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{9\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{15\pi}{8}) & \sqrt{\frac{1}{2}} \cos(\frac{21\pi}{8}) \end{bmatrix}$$

设  $a = \frac{1}{2}$ ,  $b = \sqrt{\frac{1}{2}} \cos(\frac{\pi}{8})$  及  $c = \sqrt{\frac{1}{2}} \cos(\frac{3\pi}{8})$

$$A = \begin{bmatrix} a & a & a & a \\ b & c & -c & b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

对  $4 \times 4$  DCT 中的 A 改造，  
⊗ 表示左右矩阵对应元素相乘

$$Y = (C_f X C_f^T) \otimes E_f$$

$$= \left( \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} X \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \right) \otimes \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix}$$

$$Z_{ij} = \text{round} \left( W_{ij} \frac{PF}{Qstep} \right)$$

$$PF = \begin{cases} a^2 & (0,0), (2,0), (0,2) \text{ 或 } (2,2) \\ b^2/4 & (1,1), (1,3), (3,1) \text{ 或 } (3,3) \\ ab/2 & \text{其它情况} \end{cases}$$

$$FQ = \text{round} \left( \frac{y}{QP} \right)$$

实际 DCT 输出  $W = CXC^T$



# 整数变换和量化的协同处理： 避免量化过程的除法

H.264草案规定亮度的QP最大值是51，色度QP的最大值是39。QP是量化参数，是量化步长的序号，QP每增加6，Qstep增加一倍。

量化过程为整数运算，可避免使用除法

$$FQ = \text{round}\left(\frac{y}{QP}\right)$$



$$PF = \begin{cases} a^2 & (0,0), (2,0), (0,2) \text{ 或 } (2,2) \\ b^2/4 & (1,1), (1,3), (3,1) \text{ 或 } (3,3) \\ ab/2 & \text{其它情况} \end{cases}$$

$$Z_{ij} = \text{round}\left(W_{ij} \frac{PF}{Qstep}\right)$$



设:  $qbits = 15 + \text{floor}(QP/6)$

$$MF = \frac{PF}{Qstep} 2^{qbits}$$

$$Z_{ij} = \text{round}\left(W_{ij} \frac{MF}{2^{qbits}}\right)$$

$$\begin{aligned} |Z_{ij}| &= (|W_{ij}| \cdot MF + f) \gg qbits \\ \text{sign}(Z_{ij}) &= \text{sign}(W_{ij}) \end{aligned}$$



表中列出对应QP 值为  
0 到5 的MF值（整数）

样点位置 QP	(0, 0), (2, 0), (2, 2), (0, 2)	(1, 1), (1, 3), (3, 1), (3, 3)	其它样 点位置
0	13107	5243	8066
1	11916	4660	7490
2	10082	4194	6554
3	9362	3647	5825
4	8192	3355	5243
5	7282	2893	4559



# MC 整数变换和量化

## ◆ Transform

- H.264/MPEG-4 AVC使用整数DCT4X4变换；而MPEG-2和MPEG-4使用浮点DCT8X8变换。更小块减少了块效应和明显的人工痕迹。整数系数消除了MPEG-2和MPEG-4中进行浮点系数运算时导致的精度损失。

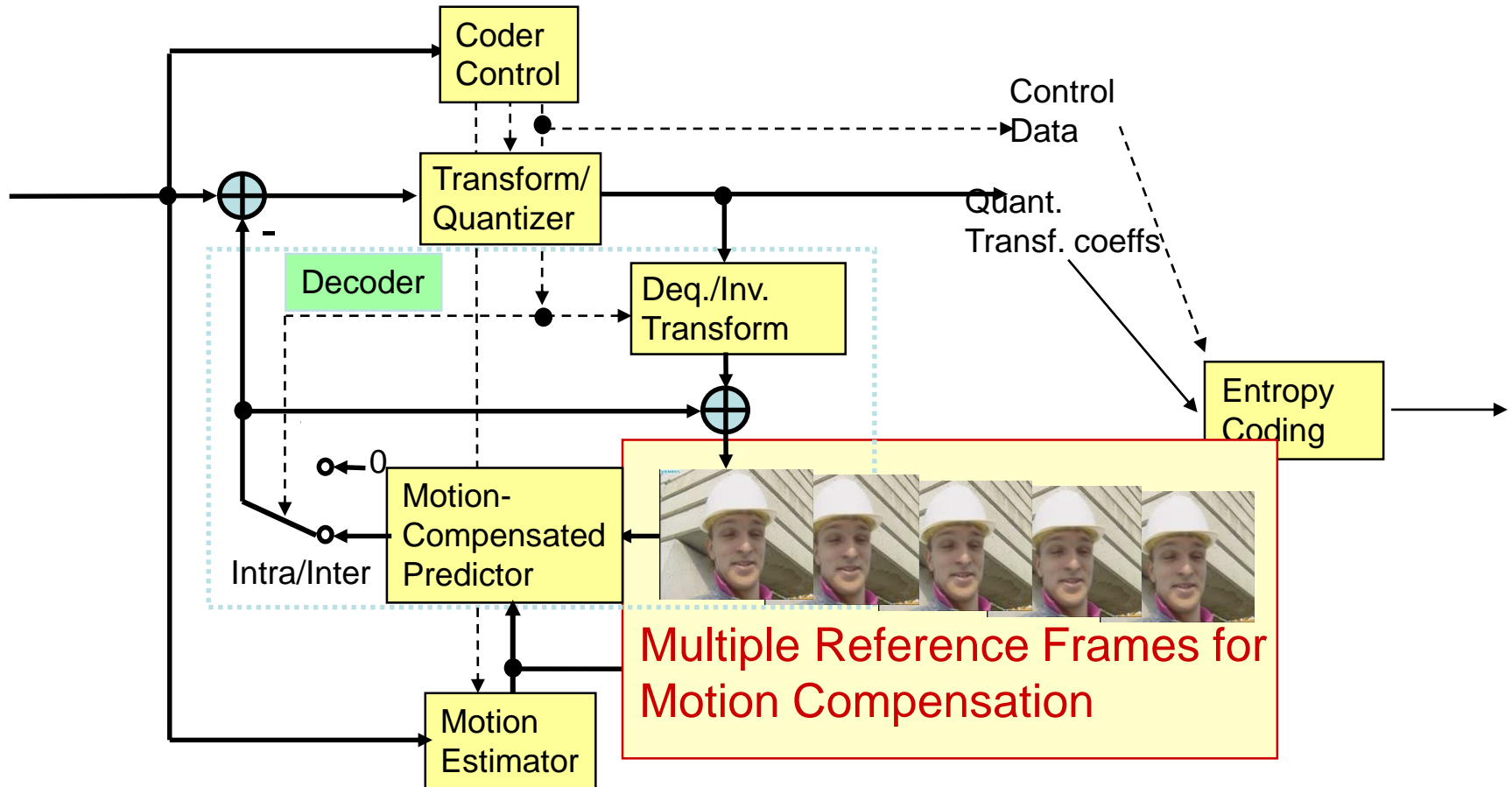
## ◆ Quantization

- 量化使用了52级步长的量化器，而H.263标准只有31级。量化步长以12.5%递增，量化步长范围的扩大使得编码器能够更灵活、精确地进行控制，在比特率和图像质量之间达到折中。

## ◆ Transform和Quantization是协同处理的

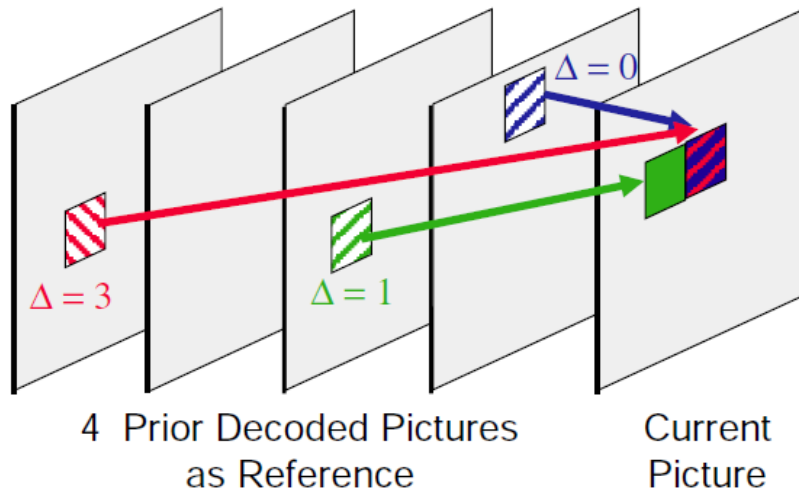


# Multiple Reference Frames

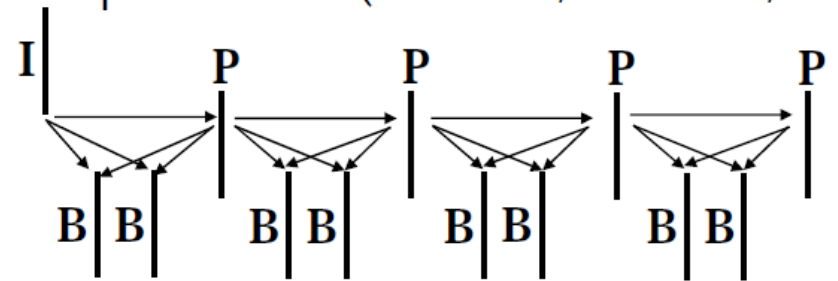




# Multiple Reference Frames and Generalized Bi-Predictive Frames

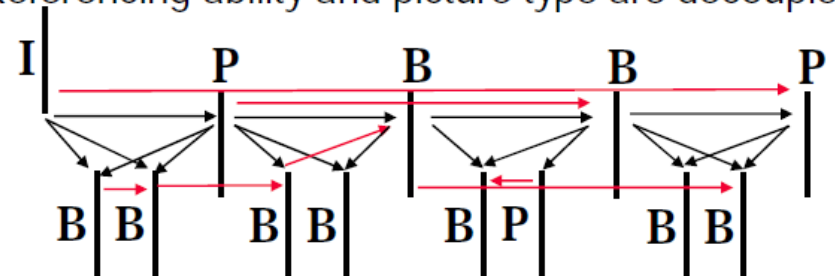


- Known dependencies (MPEG-1, MPEG-2, etc.)



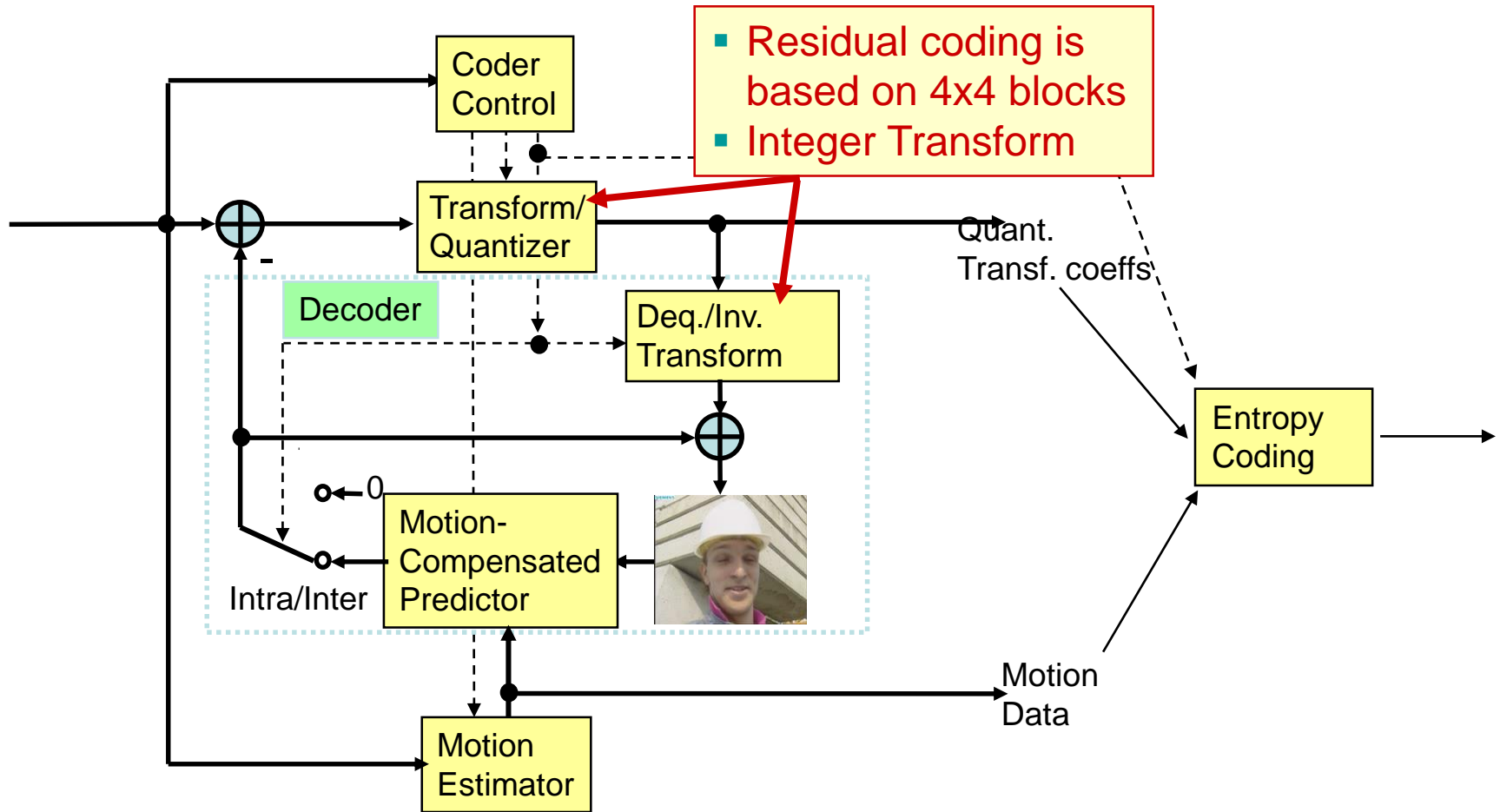
- New types of dependencies:

- Referencing order and display order are decoupled
- Referencing ability and picture type are decoupled





# Residual Coding

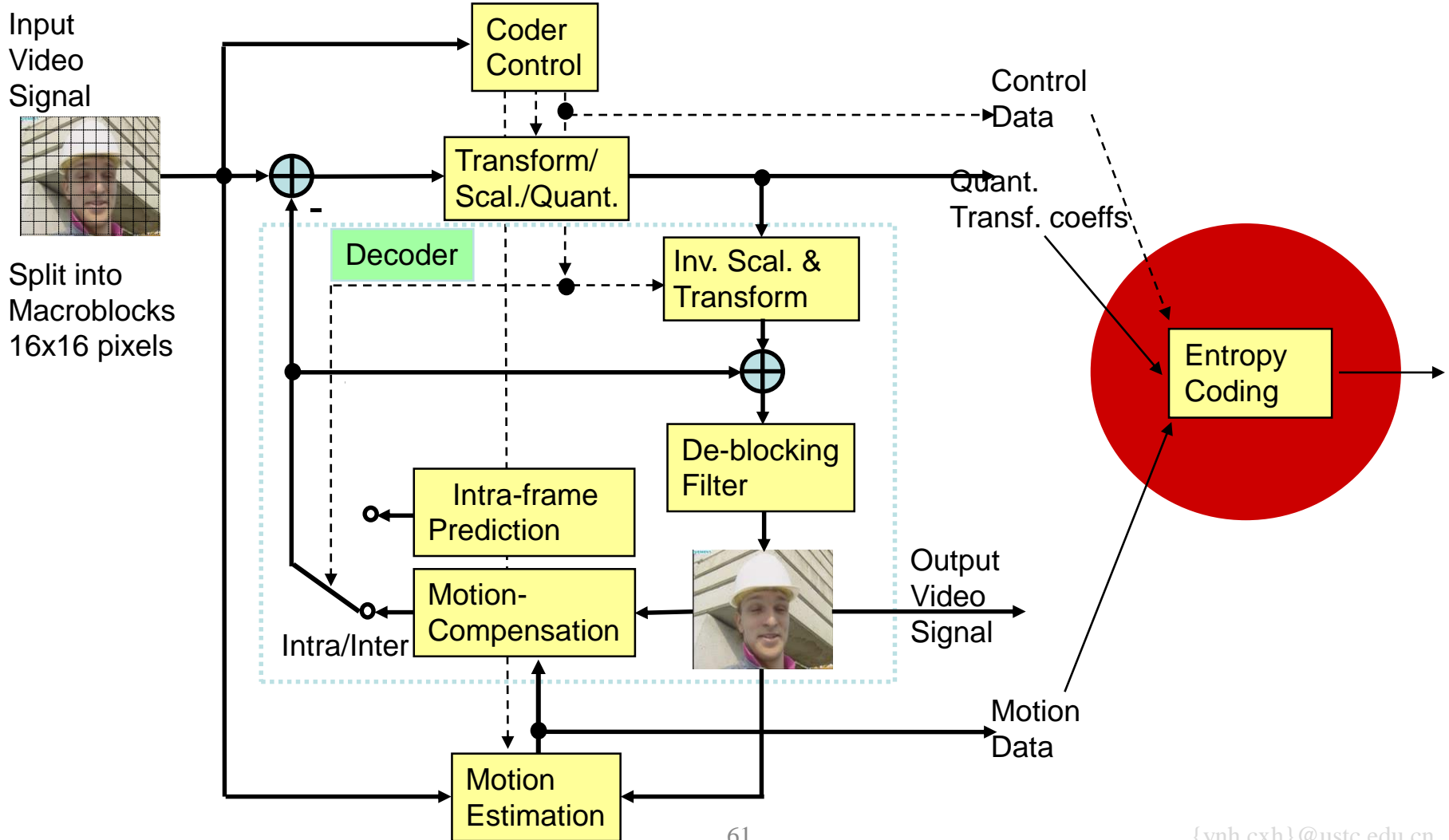


◆ 消除由于相邻宏块由不同的运动估计或者量化导致的人工痕迹。

- Whether or not to filter a particular block boundary
- Whether to apply strong or normal deblocking



# MC Entropy Coding



# MC 熵编码的改进

- ◆ VLC(Variable Length Coding)
  - Huffman Coding
  - Golomb Coding
- ◆ CAVLC
  - Context-Adaptive Variable Length Coding
- ◆ CABAC
  - Context-Adaptive Binary Arithmetic Coding

# MC Context-Adaptive

- ◆ In H.264, all symbols are encoded in a context, and the probabilities of each symbol adapt to the context in which they are coded.
- ◆ 例如，在英文中字母“u”出现的概率是3%，但是当前一个字母是“q”的时候，“u”出现的概率几乎是100%
- ◆ H.264 and AIC use several contexts. The **probability of a DCT coefficient being 0** is greater for **coefficients** in the **lower right corner** of the 8x8 coefficient matrix as for coefficients in the upper left corner.
- ◆ CAVLC与CABAC根据相邻块的情况进行当前块的编码
- ◆ CABAC比CAVLC压缩效率高，但要复杂一些



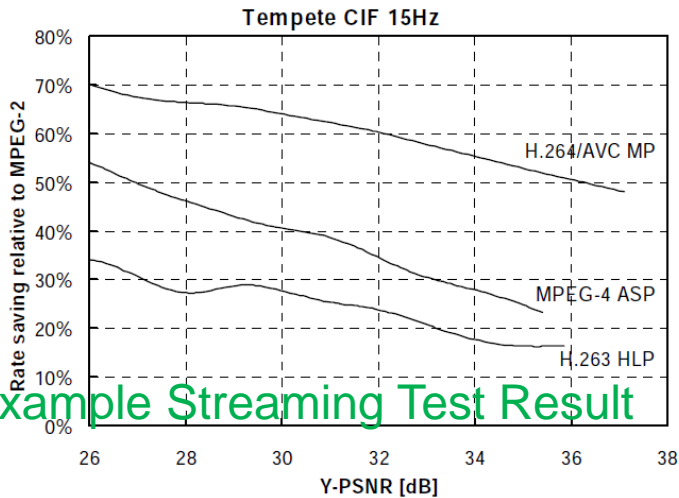
# H.264解码器个单元的运算量

H.264的编码处理计算量有MPEG-2的十多倍[参考文献?]

- ◆ The intent of the H.264/AVC project was to create a standard capable of providing good video quality at substantially lower bit rates than previous standards (i.e., **half or less the bit rate of MPEG-2**, H.263, or MPEG-4 Part 2), **without increasing the complexity** of design so much that it would be impractical or excessively expensive to implement.
- ◆ **Loop filtering (33%)** and **interpolation (25%)** are the largest components, followed by bitstream parsing and entropy decoding (13%), and inverse transforms and reconstruction (13%).
  - Comparison is based on two similarly optimized “C” only decoder implementations running on a 600-MHz P3 PC.

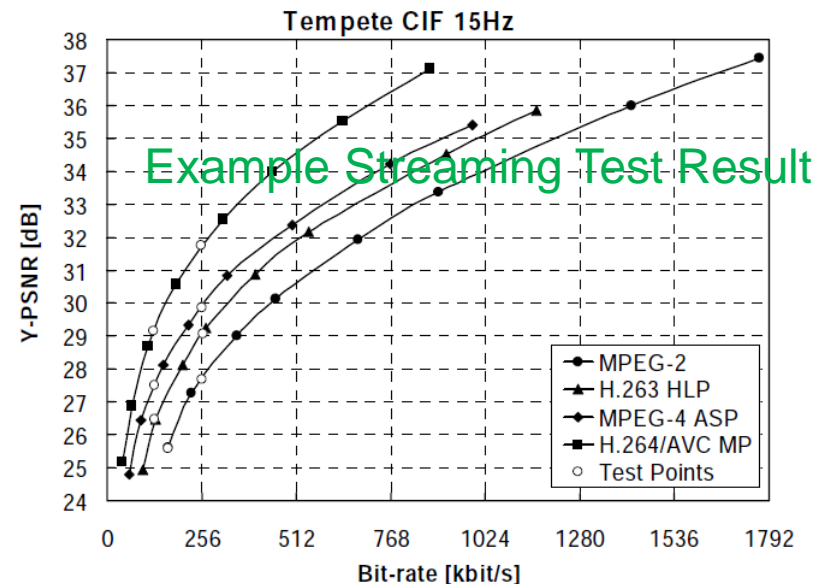
# MC Bit-rate saving

## Test Results for Streaming Application



Coder	Average bit-rate savings relative to:		
	MPEG-4 ASP	H.263 HLP	MPEG-2
H.264/AVC MP	37.44%	47.58%	63.57%
MPEG-4 ASP	-	16.65%	42.95%
H.263 HLP	-	-	30.61%

Name	Resolution	Duration	Characteristics
Foreman	QCIF	10 sec.	Fast camera and content motion with pan at the end
Container Ship	QCIF	10 sec.	Still camera on slow moving scene
News	QCIF	10 sec.	Still camera on human subjects with synthetic background
Tempete	QCIF	8.67 sec.	Camera zoom; spatial detail; fast random motion
Bus	CIF	5 sec.	Fast translational motion and camera panning; moderate spatial detail
Flower Garden	CIF	8.33 sec.	Slow and steady camera panning over landscape; spatial and color detail
Mobile & Calendar	CIF	8.33 sec.	Slow panning and zooming; complex motion; high spatial and color detail
Tempete	CIF	8.67 sec.	Camera zoom; spatial detail; fast random motion



## Test Set for Streaming Applications

# MC 小结：H.264优点

- ◆ 运动估计（时间冗余）
  - 1/4象素精度的运动估计
  - 7种大小不同的块进行匹配
  - 前向与后向多个参考帧
- ◆ 帧内预测（空间冗余）
  - 去除相临块之间的相关性
- ◆ 去块效应滤波器，对块的边界进行滤波
- ◆ 空间域到频率域的变换
  - DCT  $\rightarrow$   $4 \times 4$ 整数变换
  - 52级步长的量化器，量化步长以12.5%递增
  - 整数变换和量化协同处理
- ◆ 熵编码
  - CAVLC/CABAC

◆ AVS是由我国自主制定的音/视频编码技术标准，主要面向高清晰度电视、高密度光存储媒体等应用。AVS标准以AVC框架为基础，强调自主知识产权，同时充分考虑了实现的复杂度。主要特点有：

1.  $8 \times 8$ 的整数变换与64级量化；
2. 亮度和色度帧内预测都是以 $8 \times 8$ 块为单位，亮度块采用5种预测模式，色度块采用4种预测模式；
3.  $16 \times 16$ 、 $16 \times 8$ 、 $8 \times 16$ 和 $8 \times 8$  4种块模式的运动补偿；
4. 在 $1/4$ 象素运动估计方面，采用不同的四抽头滤波器进行半象素插值和 $1/4$ 象素插值；
5. P帧可以利用最多2帧的前向参考帧，而B帧采用前后各一个参考帧。

**2016年5月17日**，我国第二代数字视频编码标准**AVS2**颁布成为广播电视行业标准，**AVS2**标准颁布标志着我国视频技术和产业正式进入"超高清"和"超高效"的"双超时代"。根据国家新闻出版广电总局广播电视规划院进行的严格测试，**AVS2**编码效率比上一代标准**AVS+**提高了一倍以上，压缩效率超越最新国际标准**HEVC(H.265)**。



# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
  - H.264编码框架概述
  - H.264视频编码的技术创新点
  - H.264不同Profile的特点简介
- HEVC(2013)视频标准

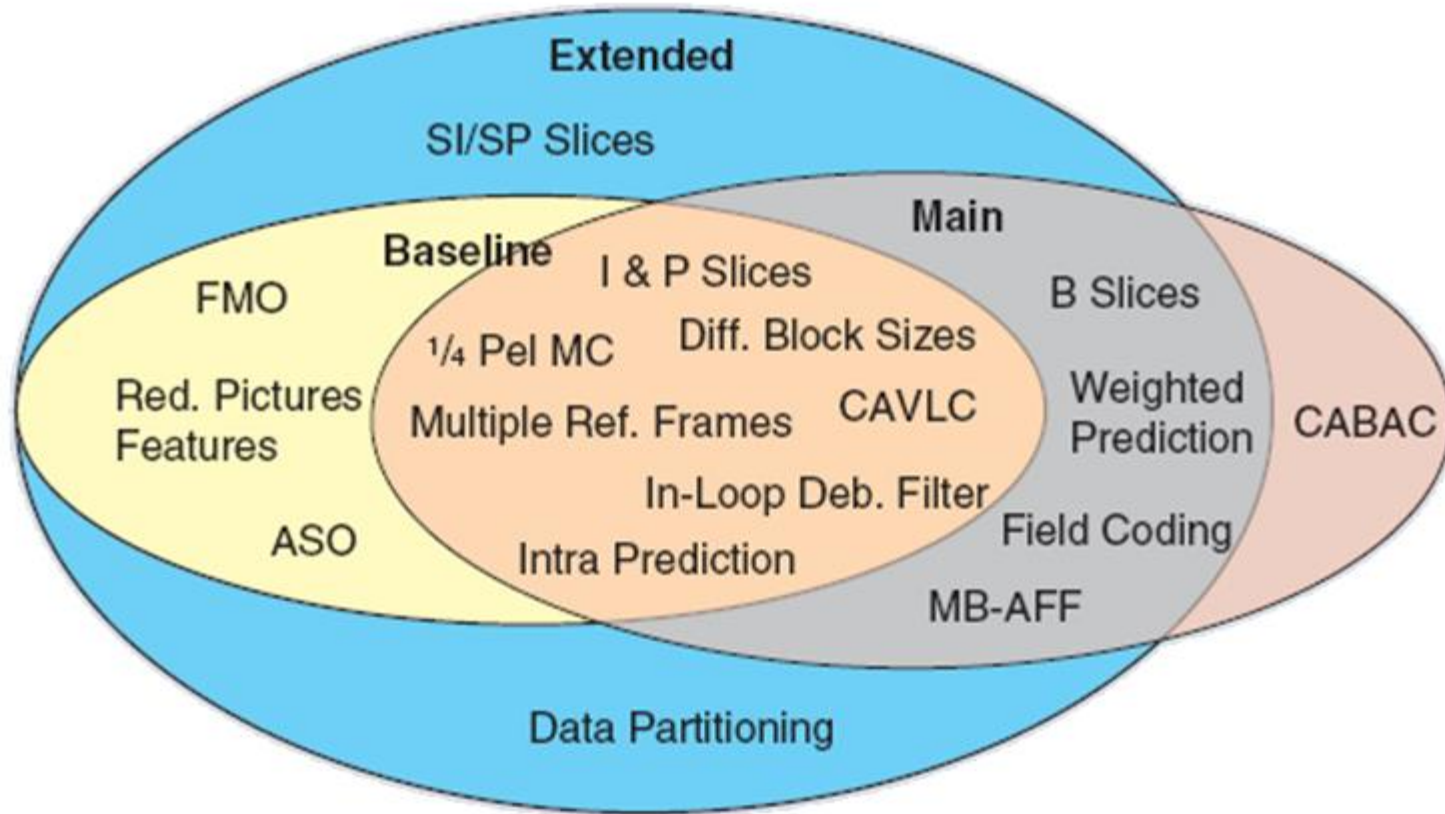
## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想： MPEG-4(1999)
- 视频质量可伸缩性 & H.264 SVC
- 多视角编码



# H.264的三种 Profile

◆不同的配置(profile)复杂度不同



FMO, AMO, SP/SI, Data partition



# H.264 Slice(≠帧)类型

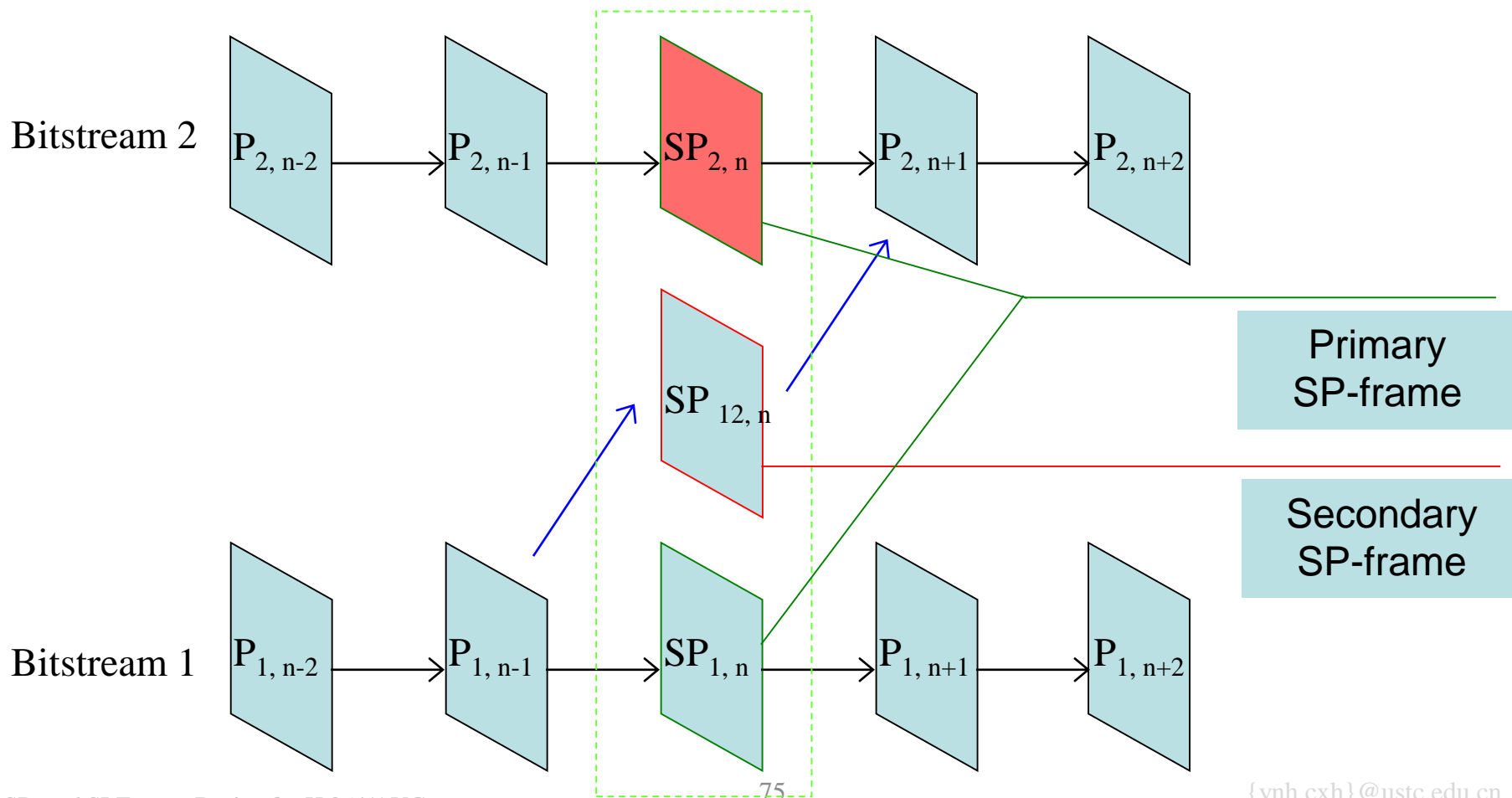
Slice Type	Description	Profile(s)
I (Intra)	Contains only I macroblocks (each block or MB is predicted from previously coded data within the same slice).	All
P (Predicted)	Contains P macroblocks (each MB or MB partition is predicted from one list 0 reference picture) and/or I MBs.	All
B (Bi-predictive)	Contains B macroblocks (each MB or MB partition is predicted from a list 0 and/or a list 1 reference picture) and/or I macroblocks.	Extended and Main
<b>SP</b> <b>(Switching P)</b>	Facilitates switching between coded streams; contains P and/or I macroblocks.	Extended
<b>SI</b> <b>(Switching I)</b>	Facilitates switching between coded streams; contains SI macroblocks (a special type of intra coded MB).	Extended

在使用不同参考帧的情况下重构相同的图像帧



# 基于SP(Switching P)帧的流切换

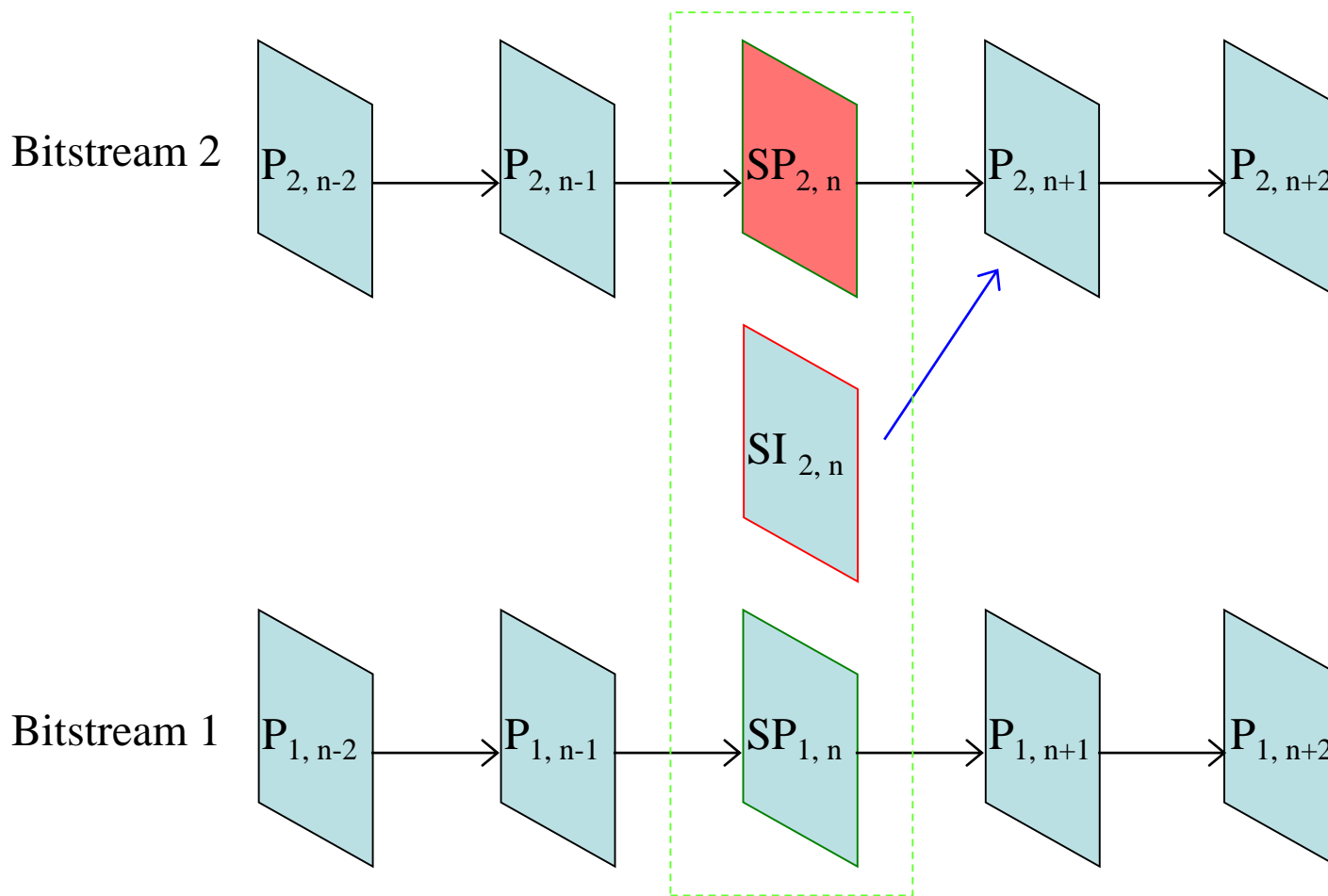
应用场景示例：两个数据流对应同一个视频序列，如分级编码情况下，发送端可以发送多个流，接收到收到所有的数据流则达到最高质量





# SP(Switching P)和SI(Switching I)用于流的拼接

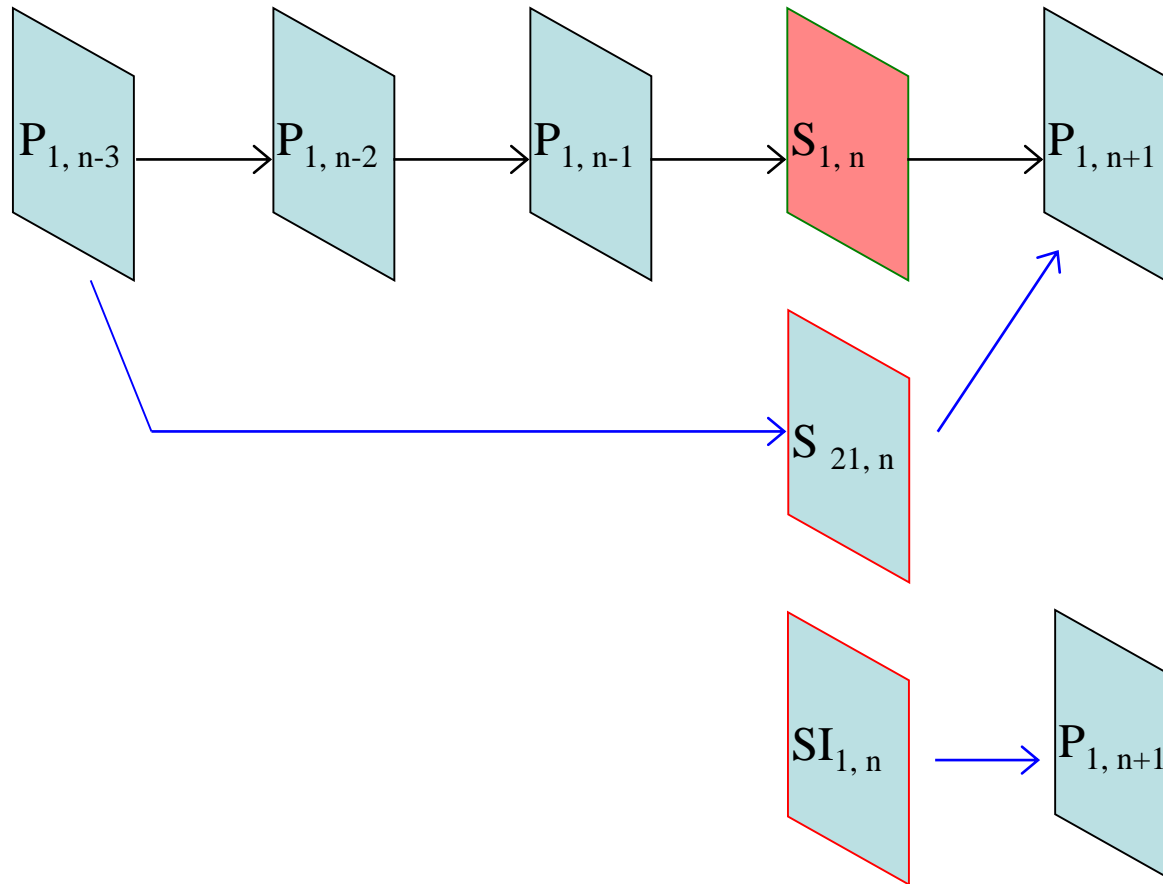
应用场景示例：两个数据流不是同一个视频序列，如多视角拍摄时不同视角摄像机影像





# SP和SI用于错误恢复

应用场景示例：SP和SI帧携带了冗余信息可以用于错误恢复

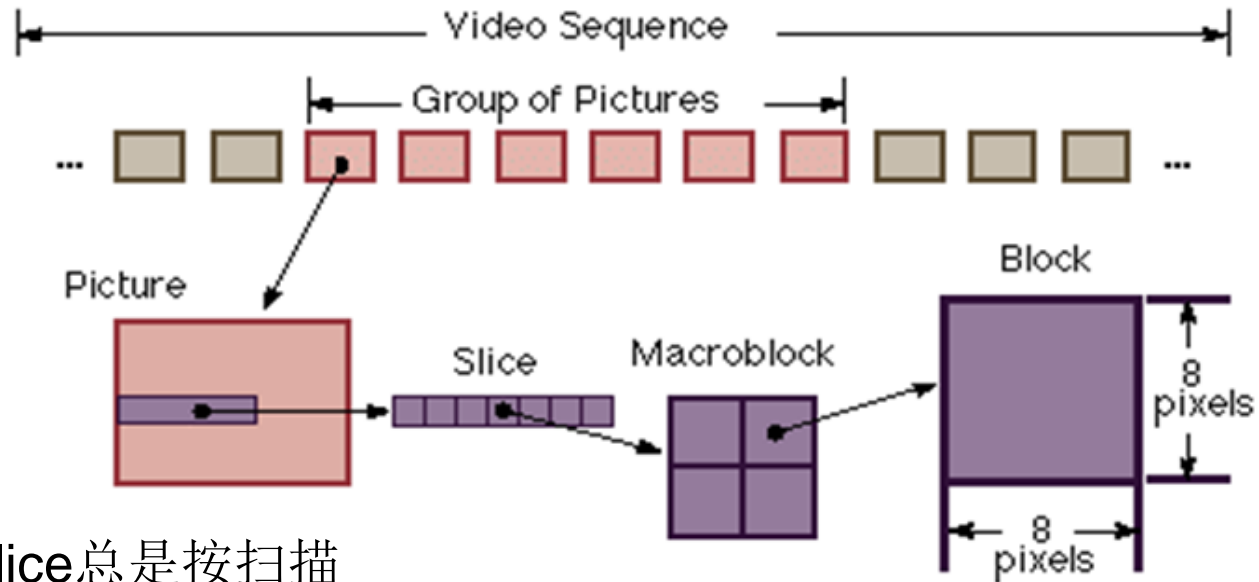


# MC Data Partitioning

◆ Since some coded information for representation of each region (e.g., motion vectors and other prediction information) is more important or more valuable than other information for purposes of representing the video content, H.264/AVC allows the syntax of each slice to be separated into up to **three different partitions** for **transmission**, depending on a categorization of syntax elements.



# ASO(Arbitrary Slice Order)



如果不使用ASO，那么slice总是按扫描顺序编解码的；如果使用了ASO，序号靠后的slice可以先于前面的slice编解码。

**A slice is a data structure that can be decoded independently from other slices of the same picture, in terms of entropy coding, signal prediction, and residual signal reconstruction.**

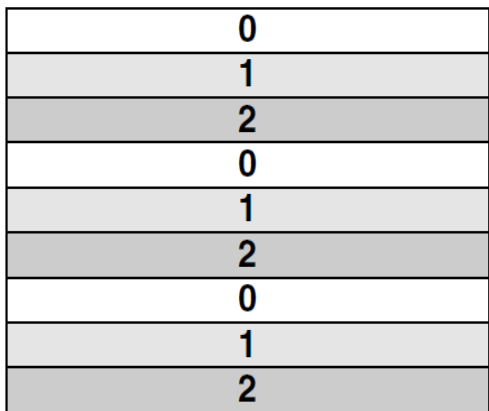




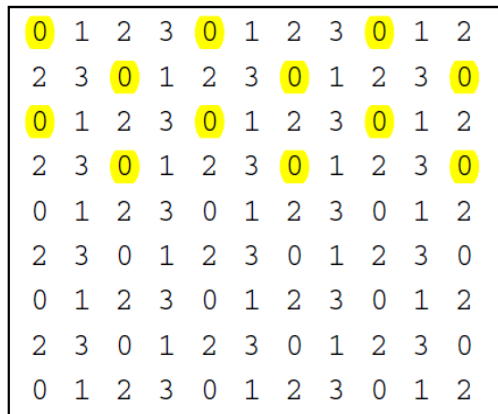
# FMO(Flexible Macroblock Ordering)

◆FMO映射表（macroblock to slice group map）决定具体每个宏块属于哪个slice groups，标准定义了6种映射

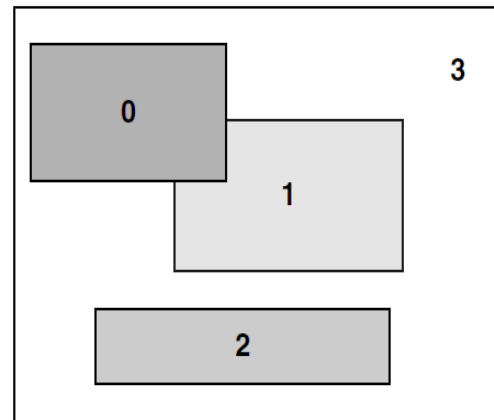
Interleaved



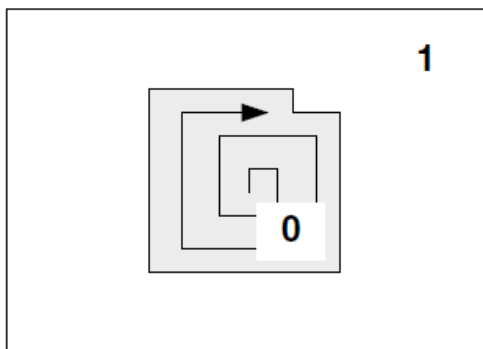
Dispersed



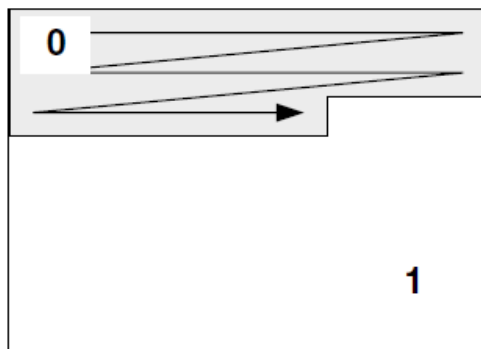
Foreground and Background



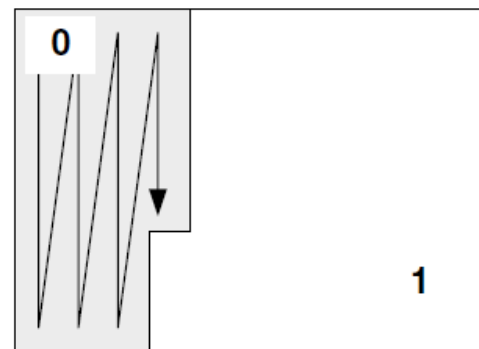
Box-out



Raster



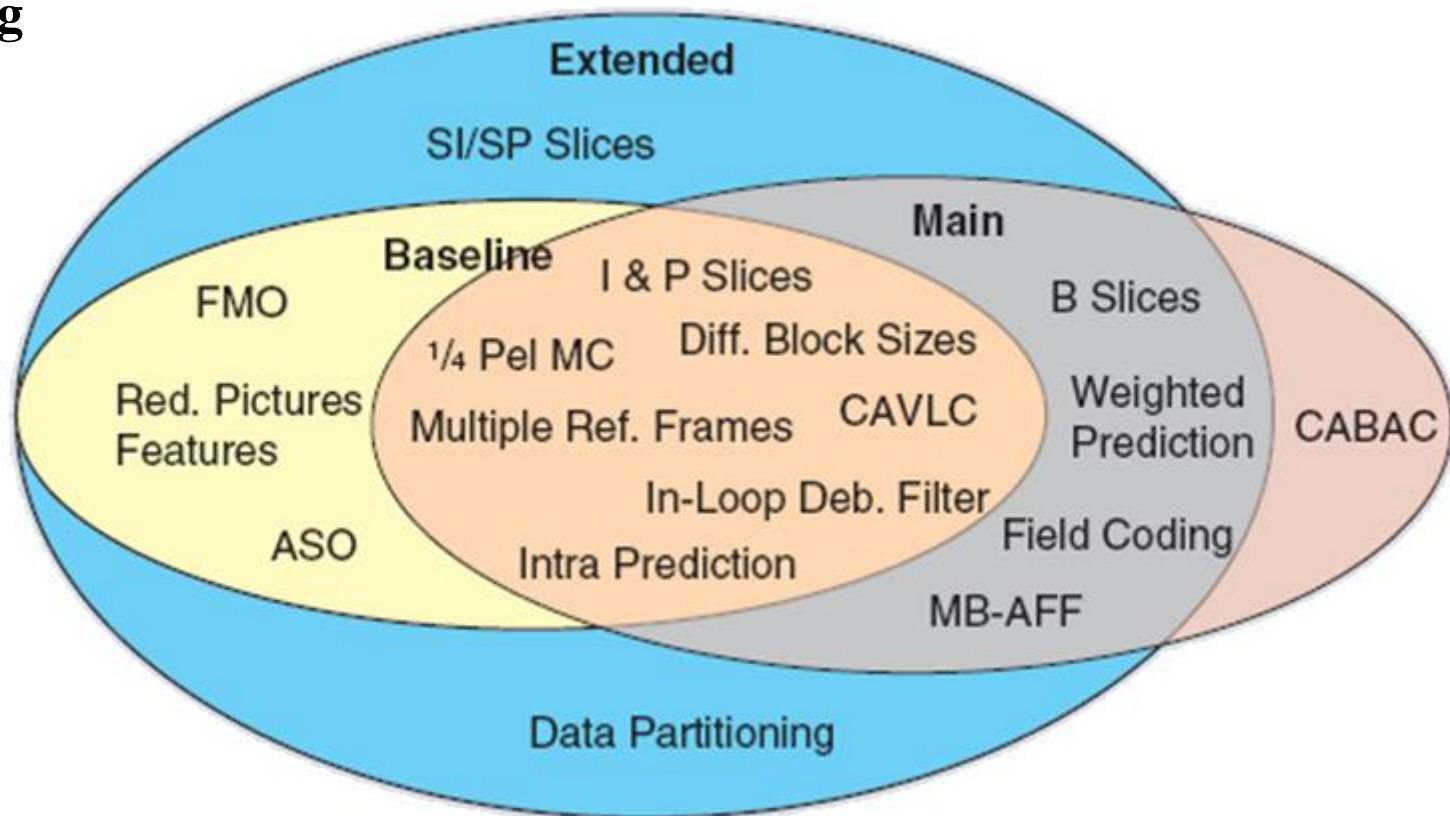
Wipe



# MC 小结: H.264 profile

ASO(Arbitrary Slice Order)  
FMO(Flexible Macroblock Ordering)  
Switching P Slice  
Switching I Slice  
Data Partitioning

Baseline Profile  
Main Profile  
Extended Profile



# MC 小结: H.264 细节决定成败

- ◆ Temporal?
  - Prediction
- ◆ Spatial?
  - Transform
  - 帧内预测
- ◆ Entropy?

H.265/HEVC (High Efficiency Video Coding)  
仍属于预测加变换的混合编码框架



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想： MPEG-4(1999)
  - 视频编码可伸缩性 & H.264 SVC
  - 多视角编码



# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
- HEVC(2013)视频标准
  - HEVC(H.265)产生的背景：4K时代来临
  - HEVC性能与编解码框架概述
  - Quadtree-based coding structure
  - Deblocking & SAO Filter
  - HEVC各模块运算量

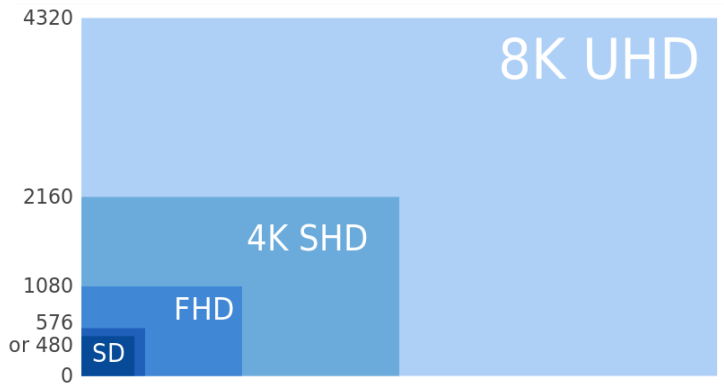
## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想：MPEG-4(1999)
- 视频编码可伸缩性 & H.264 SVC
- 多视角编码

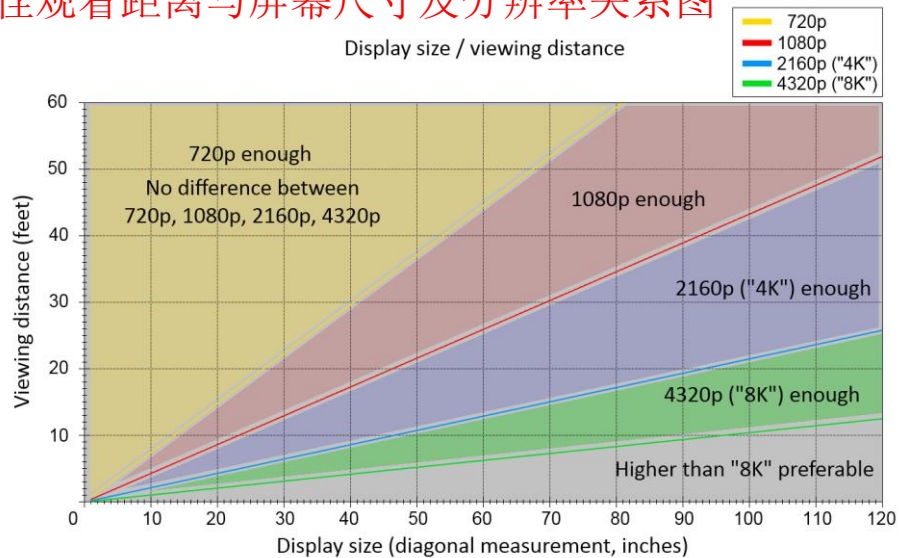


# HEVC(H.265)产生的背景：4K时代来临

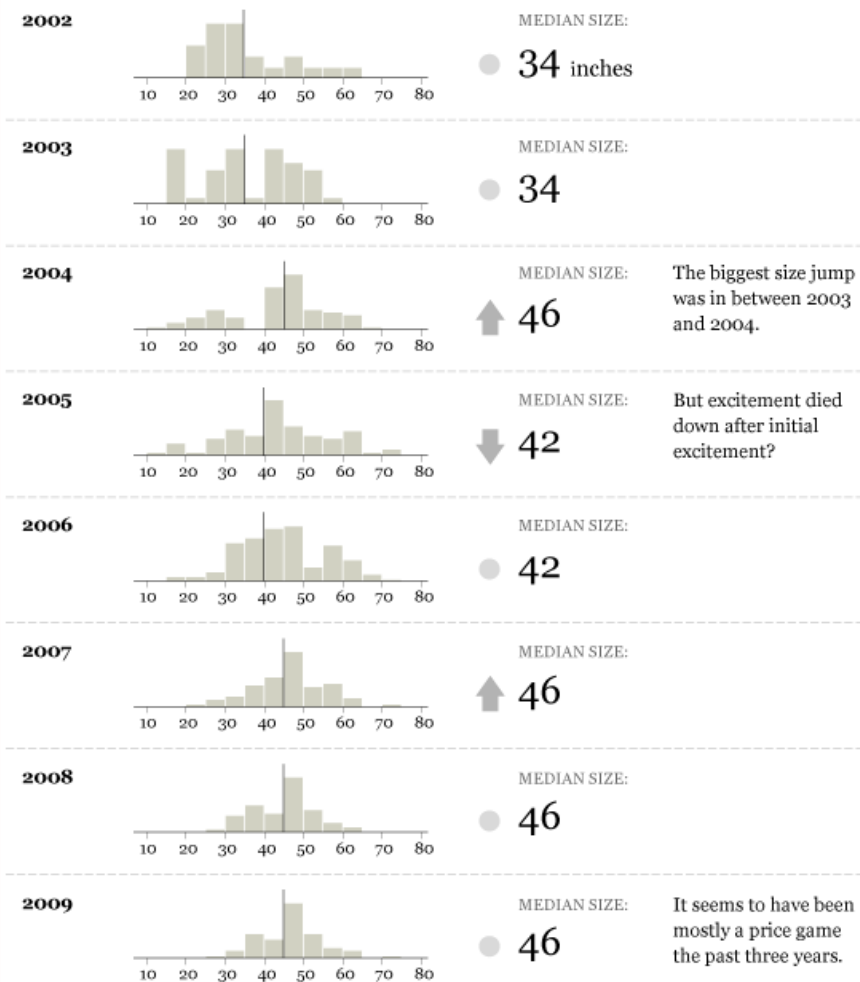
## UHDTV (Ultra high definition television)



依据 NHK 人类视觉系统测试制作的最佳观看距离与屏幕尺寸及分辨率关系图



## Television Size Over the Years





# ITU-R BT系列的演进

<http://www.itu.int/rec/R-REC-BT.2020/>, **2012-08**

BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange

编码格式: 每个分量为**10**或**12**比特 (线性/非线性)

像素数: 水平 × 垂直            7680 × 4320            3840 × 2160

<http://www.itu.int/rec/R-REC-BT.709/>, **1990**

BT.709: Parameter values for the HDTV standards for production and international programme exchange

编码格式: 每个分量**8**或**10**比特

像素数: 水平 × 垂直            1920 × 1035            1920 × 1152

<http://www.itu.int/rec/R-REC-BT.709/>, **1982**

BT.601: Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios

编码格式: 每个分量为**8**或**10**比特

像素数: 水平 × 垂直            720 × 480



# Rec. 2020 (UHDTV) Color Space

UHDTV是CIE1931的75.8%  
数码相机是53.6%  
HDTV 是 35.9%

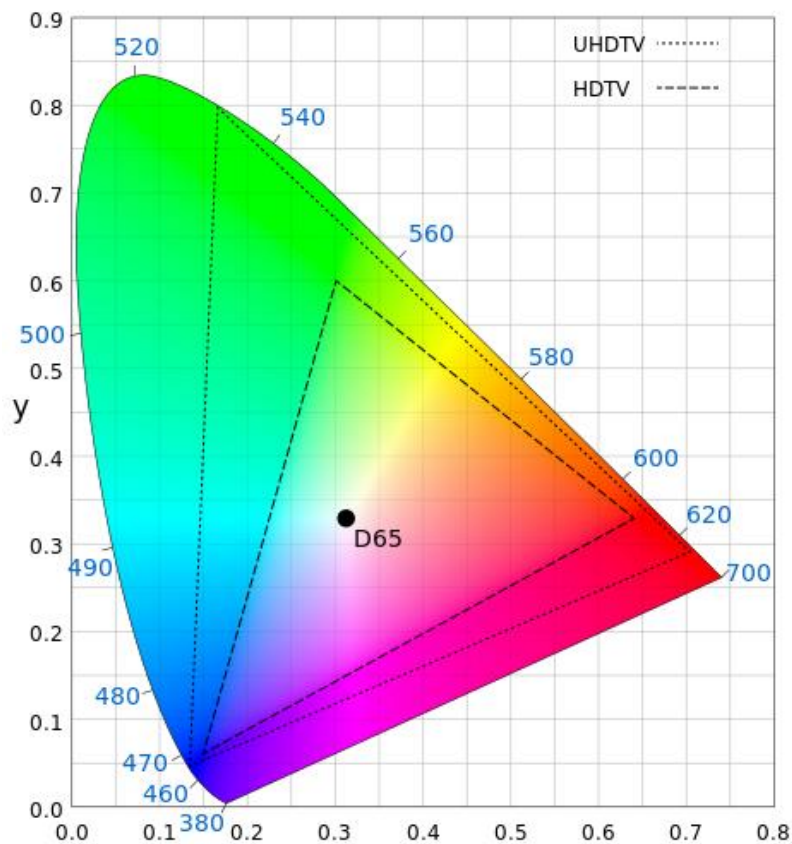
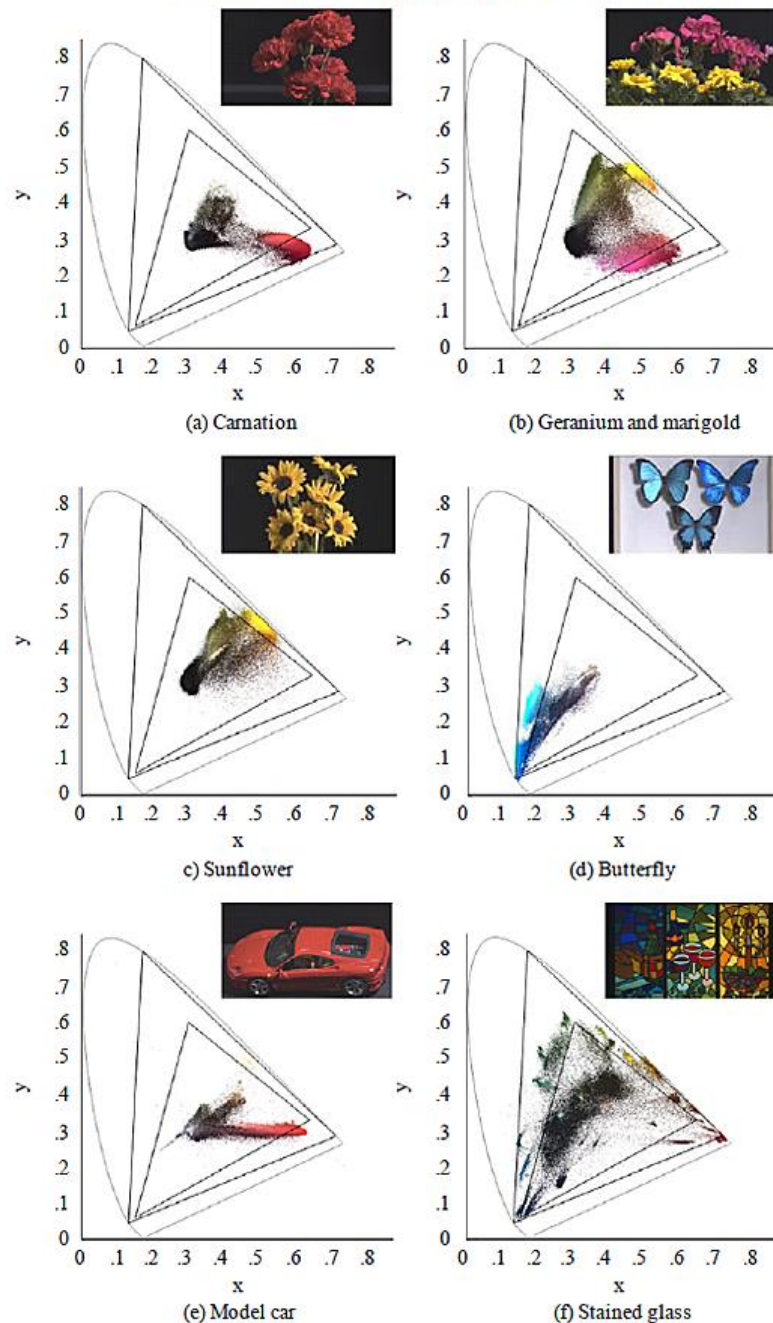


FIGURE 24

Colour distribution of objects on the x-y chromaticity coordinates  
(Inner triangle: HDTV primaries, Outer triangle: UHDTV primaries)





# HEVC(H.265)产生的背景：4K时代来临

## H.264在处理巨幅画面时的问题

- ◆ 宏块个数的爆发式增长，会导致用于编码宏块的预测模式、运动矢量、参考帧索引和量化级等宏块级**参数信息所占用的码字过多**，用于编码残差部分的码字明显减少。
- ◆ 由于分辨率的大大增加，**单个宏块**所表示的图像内容的信息大大减少，这将导致相邻的4 x 4或8 x 8块**变换后的低频系数相似程度也大大提高**，导致出现大量的冗余。
- ◆ 由于分辨率的大大增加，表示同一个运动的运动矢量的幅值将大大增加，H.264中采用一个运动矢量预测值，对运动矢量差编码使用的是哥伦布指数编码，该编码方式的特点是数值越小使用的比特数越少。因此，随着**运动矢量幅值的大幅增加**，H.264中用来对运动矢量进行预测以及编码的方法压缩率将逐渐降低。
- ◆ H.264的一些关键算法如CAVLC和CABAC两种熵编码方法、deblock滤波等都要求串行编码，**并行度比较低**。

目标：在H.264标准2~4倍的复杂度基础上，将压缩效率提升一倍以上。



# 第2章 多媒体数据压缩国际标准

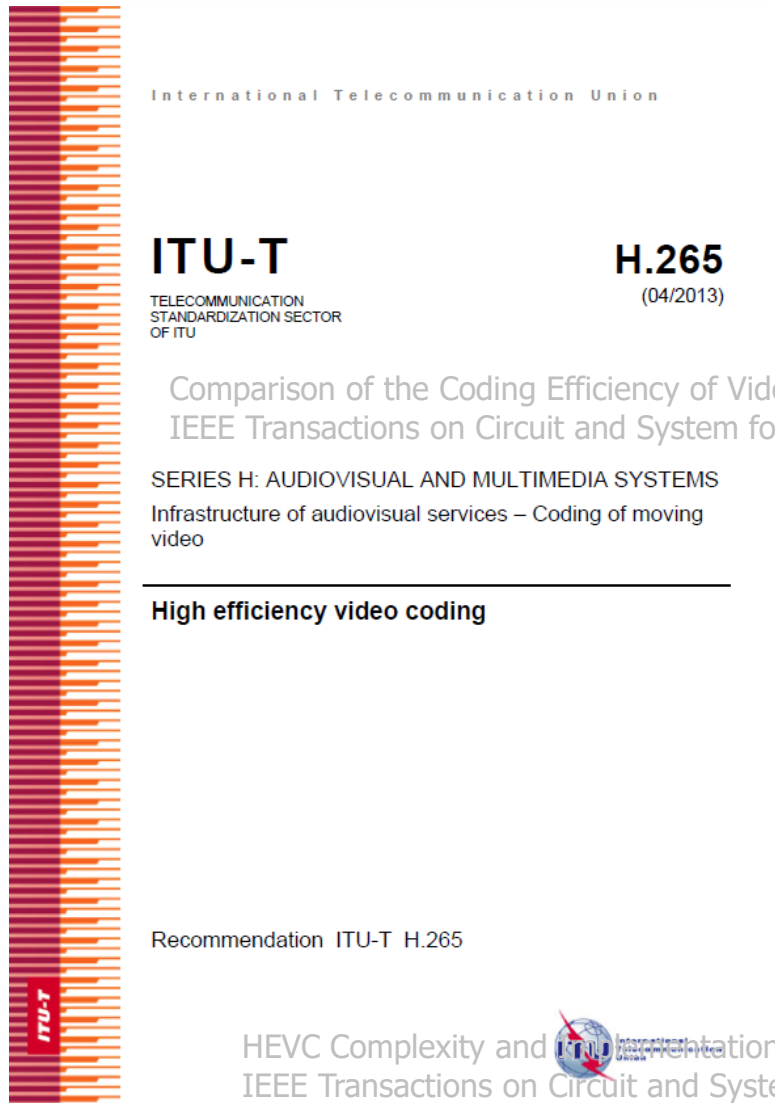
## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
- HEVC(2013)视频标准
  - HEVC(H.265)产生的背景：4K时代来临
  - HEVC性能与编解码框架概述
  - Quadtree-based coding structure
  - Deblocking & SAO Filter
  - HEVC各模块运算量

## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想：MPEG-4(1999)
- 视频编码可伸缩性 & H.264 SVC
- 多视角编码

# MC 在H.264十年之后



HEVC encoders can achieve **equivalent** subjective reproduction **quality** as encoders that conform to H.264/MPEG-4 AVC when using approximately **50% less bit rate on average**.

Overall the complexity of HEVC **decoders does not appear to be significantly different** from that of H.264/AVC decoders. HEVC **encoders** are expected to be **several times more complex** than H.264/AVC encoders, and a subject of research in years to come.



# HEVC的性能

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR ENTERTAINMENT APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/H.262 MP
HEVC MP	35.4%	63.7%	65.1%	70.8%
H.264/MPEG-4 AVC HP	–	44.5%	46.6%	55.4%
MPEG-4 ASP	–	–	3.9%	19.7%
H.263 HLP	–	–	–	16.2%

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR INTERACTIVE APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	H.263 CHC	MPEG-4 ASP	MPEG-2/H.262 MP
HEVC MP	40.3%	67.9%	72.3%	80.1%
H.264/MPEG-4 AVC HP	–	46.8%	54.1%	67.0%
H.263 CHC	–	–	13.2%	37.4%
MPEG-4 ASP	–	–	–	27.8%

➤ MP: Main Profile

➤ HP: High Profile

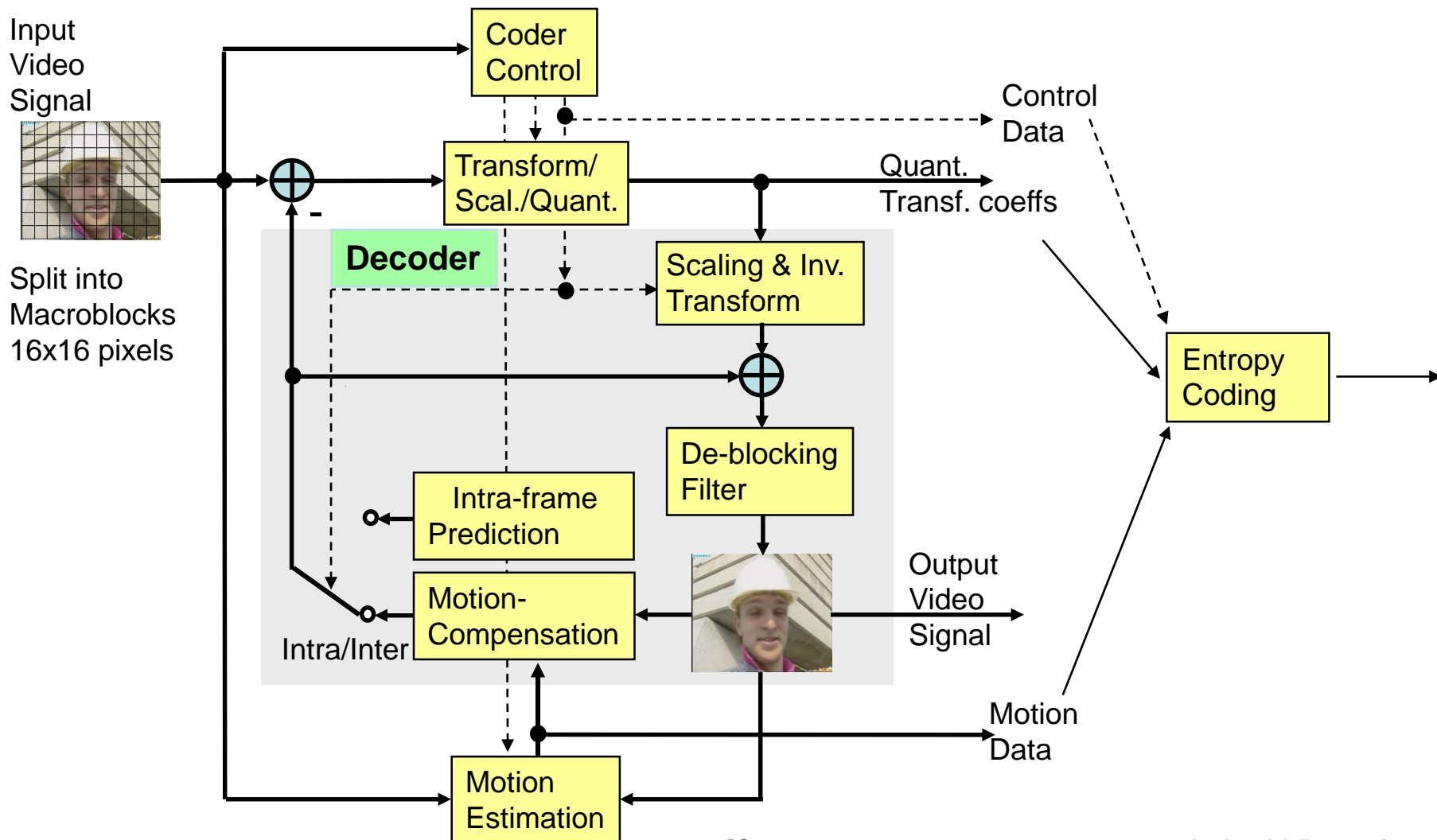
➤ ASP: Advanced Simple Profile

➤ HLP: High Latency Profile



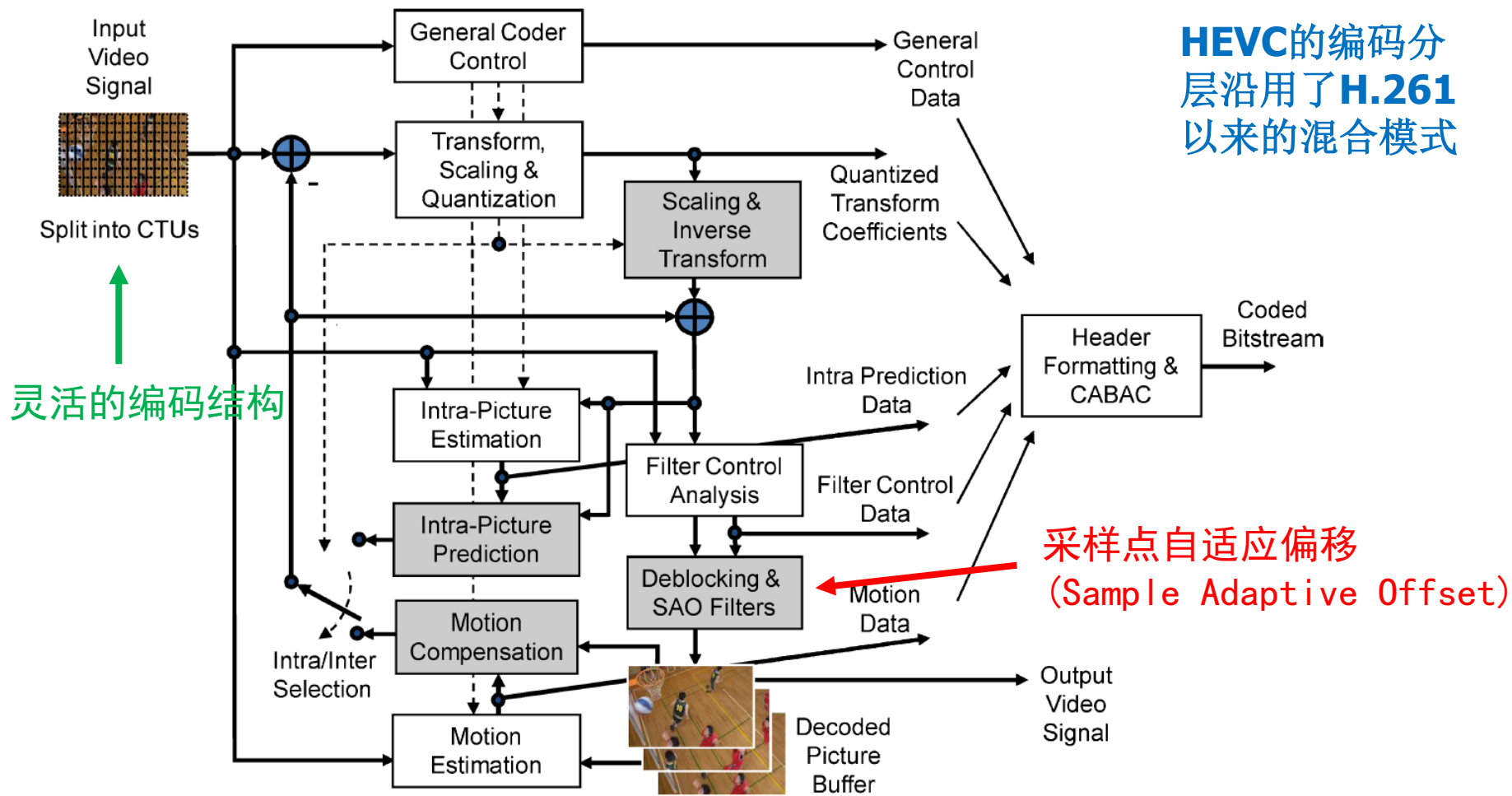
# H.264编码器框架

## Review





# HEVC(H.265)编码器框架



# MC 推荐阅读的文献

- ◆ 朱秀昌等，新一代视频编码标准——HEVC, 《南京邮电大学学报》，2013年6月
- ◆ IEEE Transactions on Circuit and System for Video Technology, vol. 22, N0. 12, Dec 2012
  - Overview of the High Efficiency Video Coding (HEVC) Standard
  - Sample Adaptive Offset in the HEVC Standard
  - Block Partitioning Structure in the HEVC Standard
  - Comparison of the Coding Efficiency of Video Coding Standards—Including HEVC
  - HEVC Complexity and Implementation Analysis



# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
- HEVC(2013)视频标准
  - HEVC(H.265)产生的背景：4K时代来临
  - HEVC性能与编解码框架概述
  - **Quadtree-based coding structure**
  - Deblocking & SAO Filter
  - HEVC各模块运算量

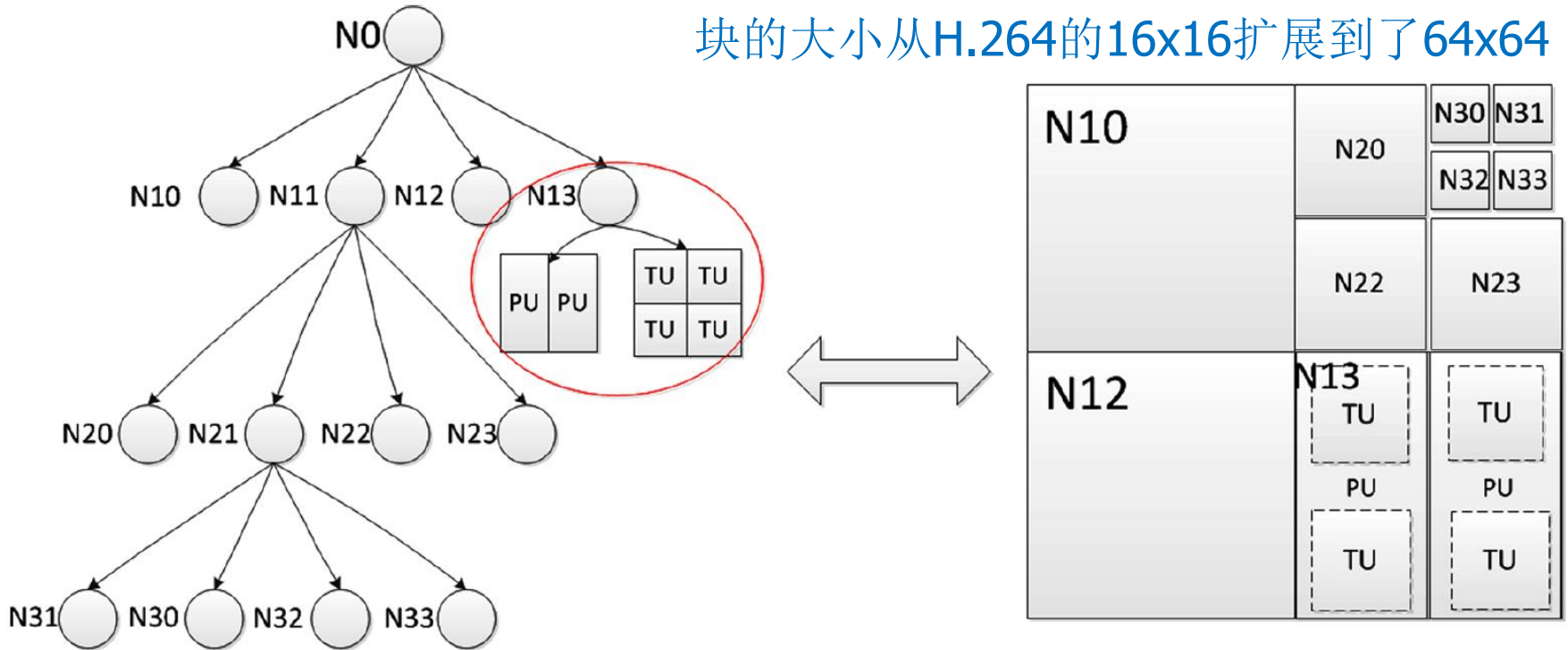
## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想：MPEG-4(1999)
- 视频编码可伸缩性 & H.264 SVC
- 多视角编码



# Quadtree-based coding structure

块的大小从H.264的16x16扩展到了64x64



**A treeblock can be further split into smaller blocks recursively, which leads to a quadtree based partition. As shown in Fig, N0 is a treeblock. It splits incrementally into N10 to N33 and generates a quadtree. Each leaf (e.g., N13, N20, N32) of the quadtree is referred to as a coding node, called a CU.**



# 灵活的编码结构

## CU、PU、TU的清晰分离

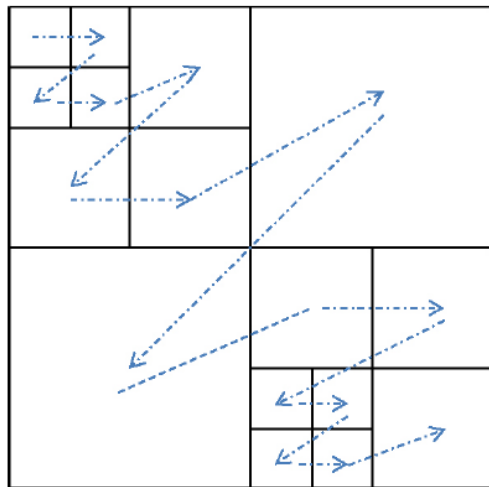
- ◆ In contrast to the fixed size  $16 \times 16$  **macroblock** structure of H.264/AVC, HEVC defines **three different units** according to their functionalities.
- ◆ 同时，采用了更加灵活的编码结构来提高编码效率，包括编码单元(CU, Coding Unit)、预测单元(PU, Predict Unit)和变换单元(TU, Transform Unit)。
- ◆ 其中编码单元类似于H.264/AVC中的宏块的概念，用于编码的过程，预测单元是进行预测的基本单元，变换单元是进行变换和量化的基本单元。这三个单元的分离，使得变换、预测和编码各个处理环节更加灵活，也有利于各环节的划分更加符合视频图像的纹理特征，有利于各个单元更优化的完成各自的功能。



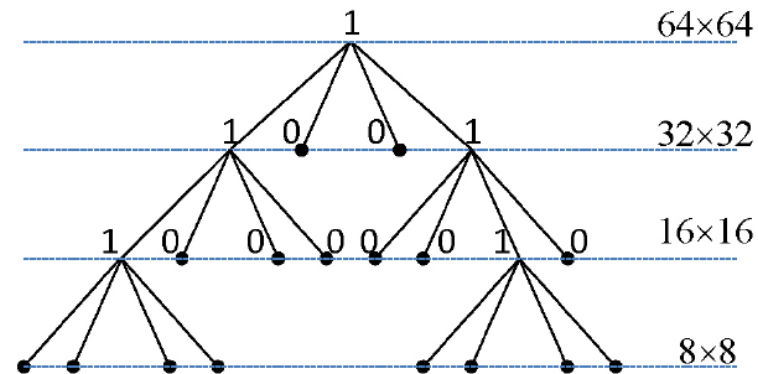
# Coding Tree Unit

Picture  $\rightarrow$  Slice  $\rightarrow$  CTU  $\rightarrow$  CU

**A slice contains an integer multiple of CTU**, which is an analogous term to the macroblock in H.264/AVC. Inside a slice, a raster scan method is used for processing the CTU.



(a)

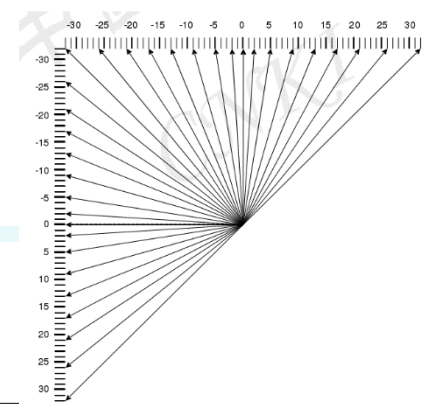


(b)

四叉树结构  
quadtree structure

Fig. 2. Example of CTU partitioning and processing order when size of CTU is equal to  $64 \times 64$  and minimum CU size is equal to  $8 \times 8$ . (a) CTU partitioning. (b) Corresponding coding tree structure.

# MC Prediction Unit



- 一个编码块CU可以分割1或2或4个预测块PU。
- 与H.264类似，提供三种分割模式
  - Skip mode
  - Intra mode
  - Inter mode
- 但支持非对称的PU分割

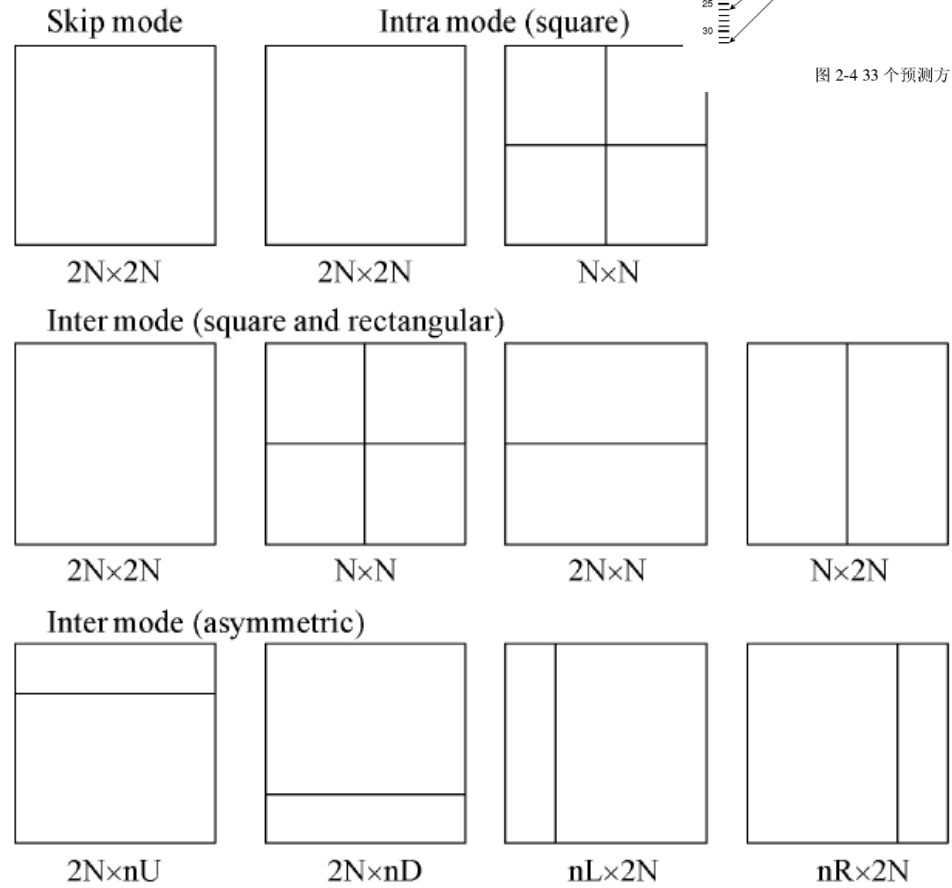


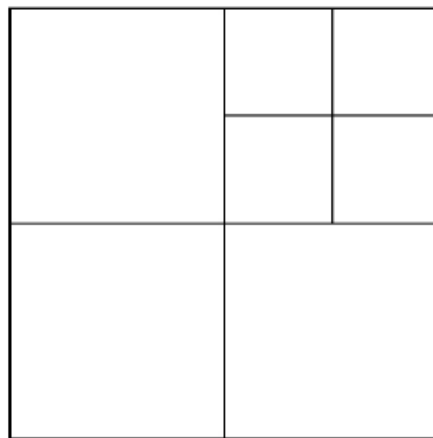
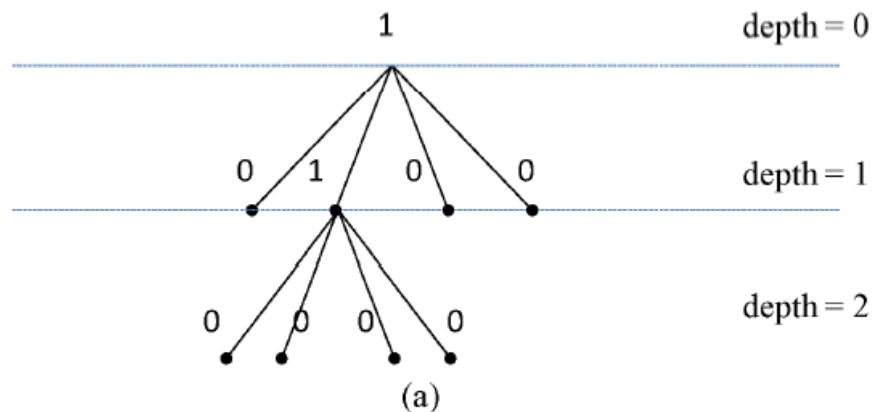
图 2-4 33 个预测方向

# MC Transform Unit

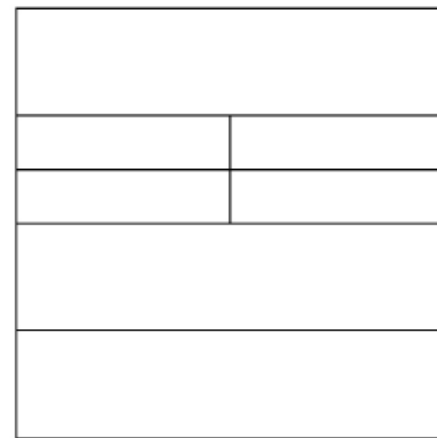
与PU类似一个CU可分割为1个或多个TU。

- H.264
  - 支持 $4 \times 4$ 和 $8 \times 8$ 的TU
  - 帧间预测时只要有块小于 $8 \times 8$ 就全部采用 $4 \times 4$
- H.265
  - 对于帧间编码，变换块的大小根据运动补偿块的大小进行自适应的调整
  - 对于帧内编码，变换块的大小根据帧内预测残差的特性进行自适应的调整

RQT(Residual Quad-tree Transform)



(b)



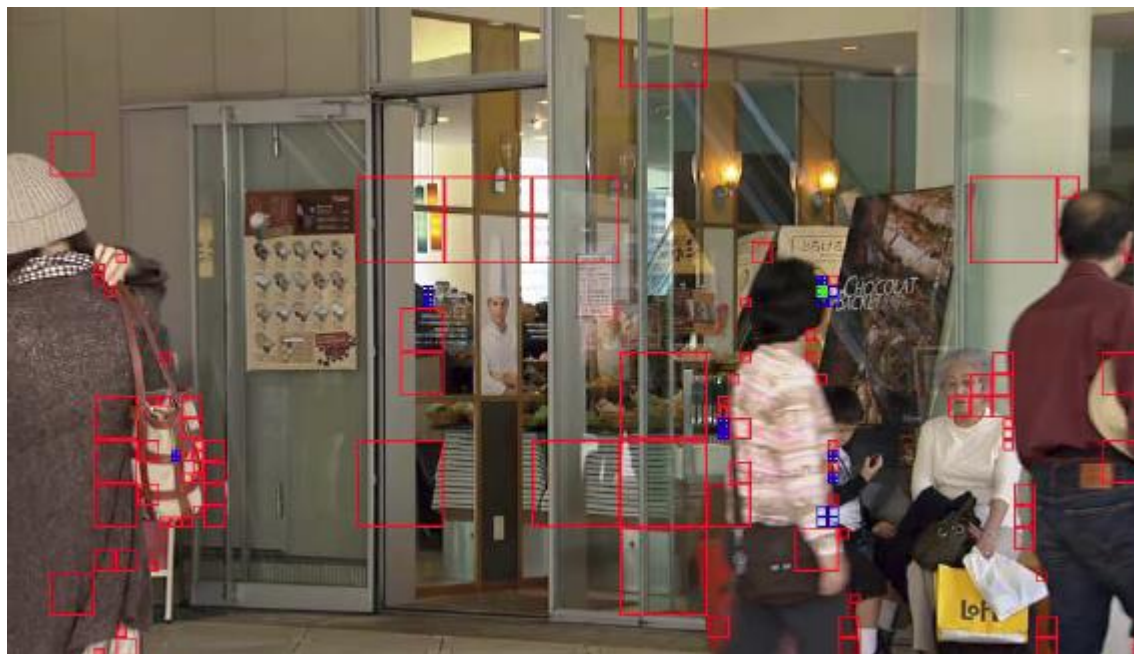
(c)

Fig. 6. Examples of transform tree and block partitioning. (a) Transform tree. (b) TU splitting for square-shaped PU. (c) TU splitting for rectangular or asymmetric shaped PU.

# MC CU和PU的关系

一个CU可能同时包含有高细节（高频）部分和低细节（低频）部分，因此CU直接用于变换处理不合适。H.265中允许将CU分割为TU，TU和PU并不需要对齐。

H.264中TU大小为 $4 \times 4$ 或 $8 \times 8$ ，TU是小于PU的



CUs in red have RQT transform tree of depth 1 (i.e. each TU spans more than 1 PU). CUs in blue and green have RQT transform tree of depths 2 and 3 respectively.

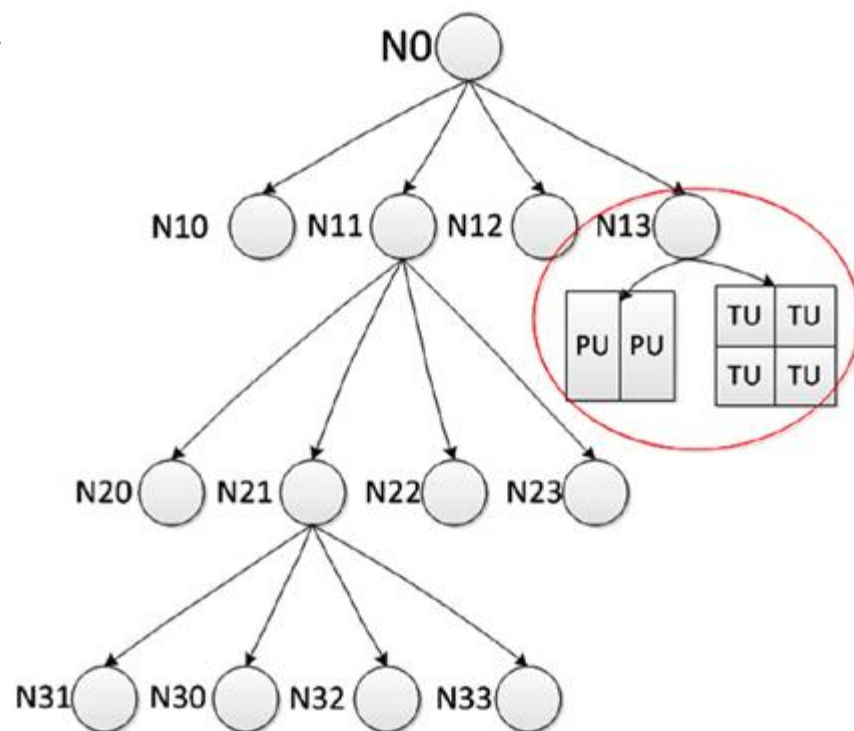
# MC 小结: Quadtree-based coding structure

## ◆ CU、PU、TU分离的思想

- CU, Coding Unit
- PU, Predict Unit
- TU, Transform Unit

## ◆ TU大小的自适应调整

- RQT
  - Residual Quad-tree Transform
- 非正方形区域的变换





# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
- HEVC(2013)视频标准
  - HEVC(H.265)产生的背景：4K时代来临
  - HEVC性能与编解码框架概述
  - Quadtree-based coding structure
  - **Deblocking & SAO Filter**
  - HEVC各模块运算量

## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想：MPEG-4(1999)
- 视频编码可伸缩性 & H.264 SVC
- 多视角编码

# MC Deblocking Filter

- ◆ The HEVC deblocking filter is **less complex** as compared to the H.264/AVC deblocking filter, while still having the capability to improve the subjective and objective quality.
- ◆ Another aspect that received significant attention in the HEVC deblocking filter design is its **suitability for parallel processing**.



# SAO: Sample Adaptive Offset

◆ 采样点自适应偏移(SAO)先**将像素分为不同的类型**,然后按照不同的类型为每个像素值**加上一个简单的偏移值**,达到减少失真的目的,从而提高压缩率,减少码流。

◆ 采用SAO后,平均可以减少2%~6%的码流,而编码器和解码器的性能消耗仅仅增加了约2%



## Two SAO types: EO, BO

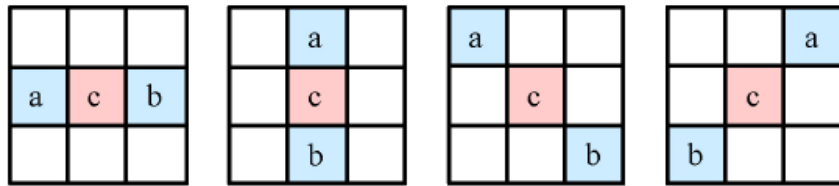
- ◆ SAO may use different offsets sample by sample in a region depending on the sample classification, and SAO parameters are adapted from region to region.
- ◆ Two SAO types that can satisfy the requirements of low complexity are adopted in HEVC: **edge offset (EO)** and **band offset (BO)**. For EO, the sample classification is based on comparison between **current samples and neighboring samples**. For BO, the sample classification is based on **sample values**.



# SAO: edge offset (EO)

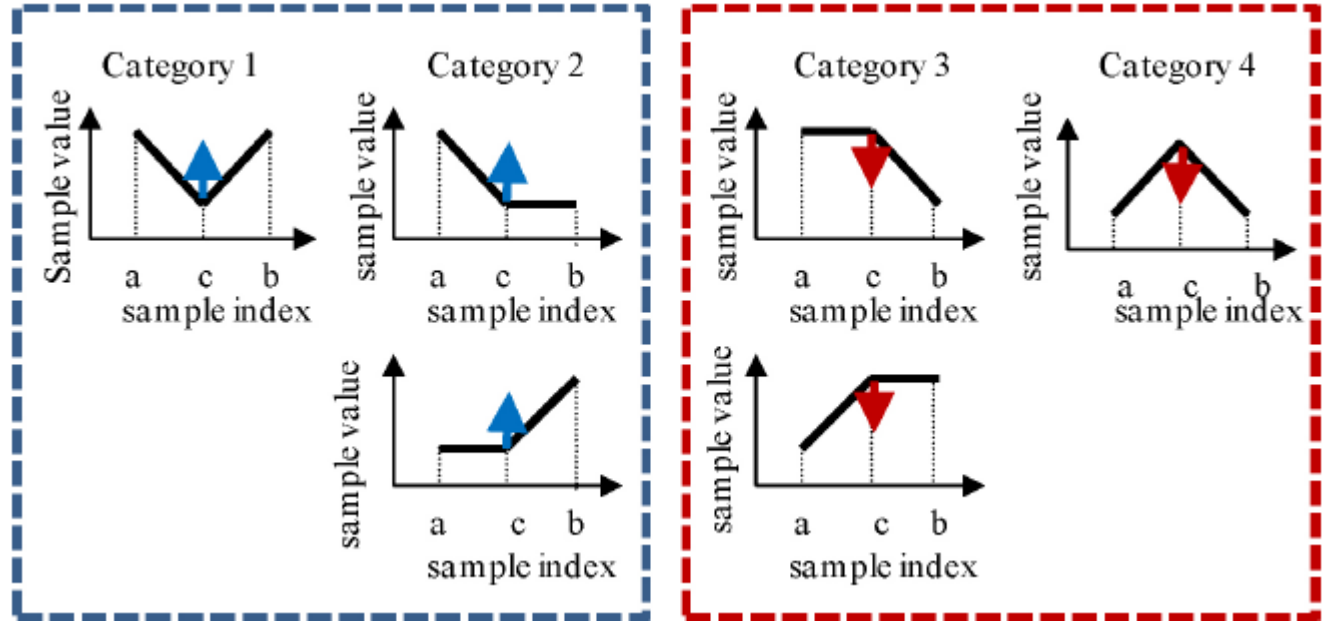
## Sample Classification Rules for Edge Offset

Four 1-D directional patterns for EO sample classification



Category	Condition
1	$c < a \ \&\& \ c < b$
2	$(c < a \ \&\& \ c == b) \    \ (c == a \ \&\& \ c < b)$
3	$(c > a \ \&\& \ c == b) \    \ (c == a \ \&\& \ c > b)$
4	$c > a \ \&\& \ c > b$
0	None of the above

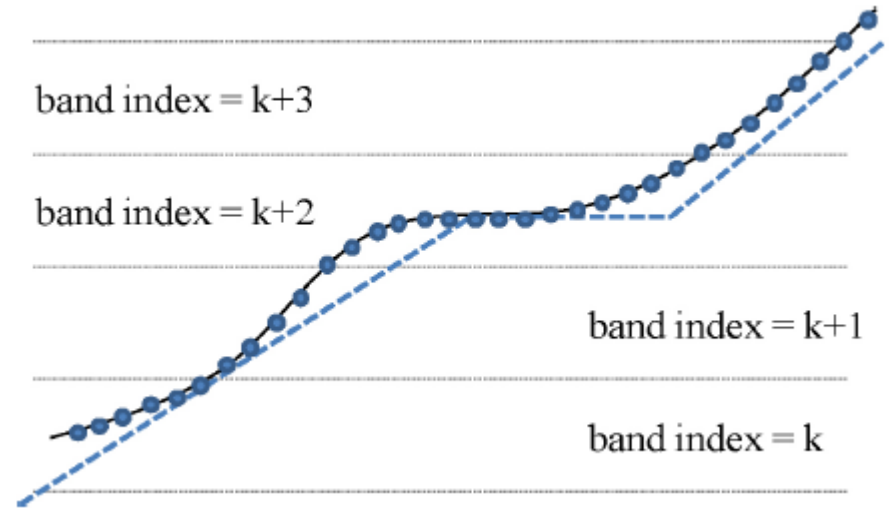
Positive offsets for EO categories 1 and 2 and negative offsets for EO categories 3 and 4 result in smoothing.



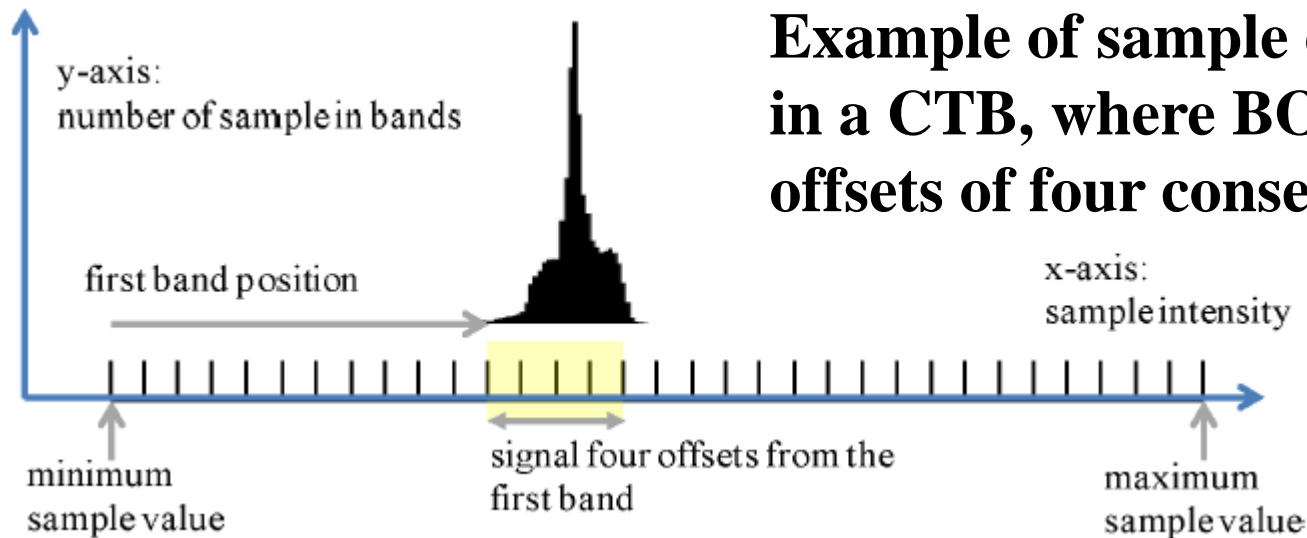


# SAO: band offset (BO)

→ Example of BO, where the dotted curve is the original samples and the solid curve is the reconstructed samples.



Example of sample distribution in a CTB, where BO send the offsets of four consecutive bands.





# 小结: Deblocking & SAO Filter

- ◆ deblocking filter
  - less complex
  - suitable for parallel processing
- ◆ SAO: Sample Adaptive Offset
  - edge offset (EO)
  - band offset (BO)



# 第2章 多媒体数据压缩国际标准

## ◆ 2.5 视频压缩的国际标准

- 视频编解码器的框架与标准简介
- MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
- MPEG-4 AVC/H.264(2003)视频标准
- HEVC(2013)视频标准
  - HEVC(H.265)产生的背景：4K时代来临
  - HEVC性能与编解码框架概述
  - Quadtree-based coding structure
  - Deblocking & SAO Filter
  - HEVC各模块运算量

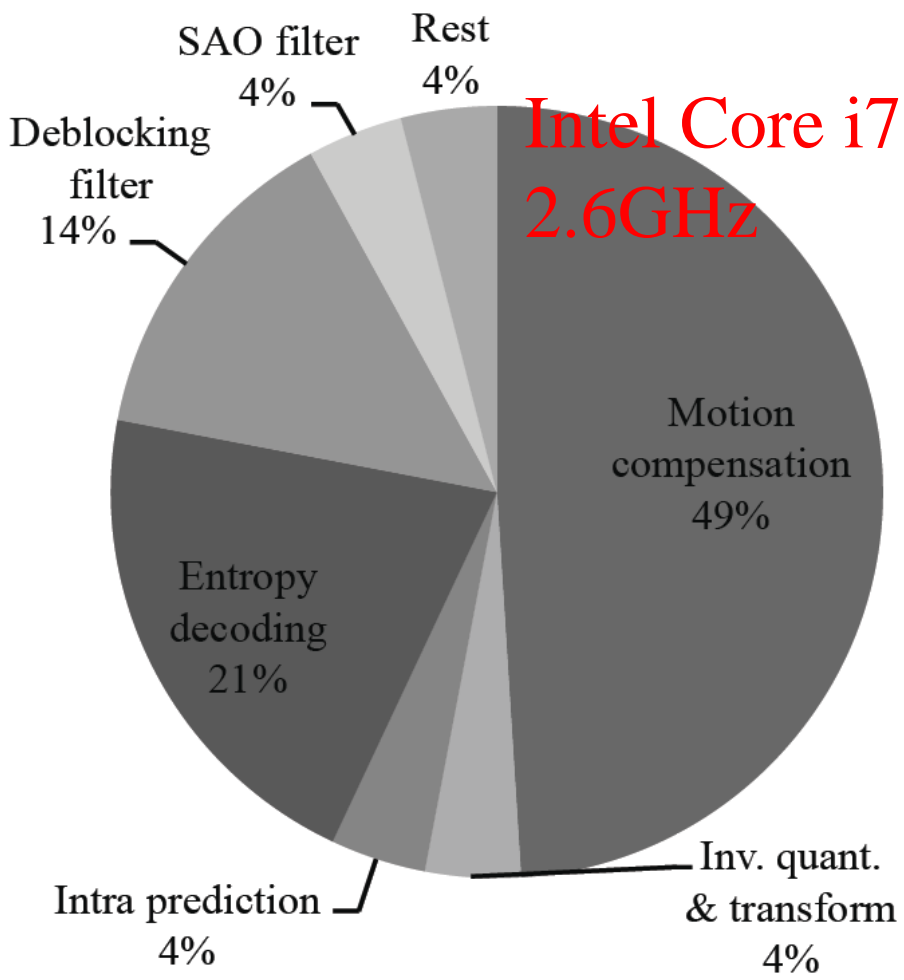
## ◆ 2.6 不同应用场景下的其他编码思想

- 基于模型/对象的编码思想：MPEG-4(1999)
- 视频编码可伸缩性 & H.264 SVC
- 多视角编码

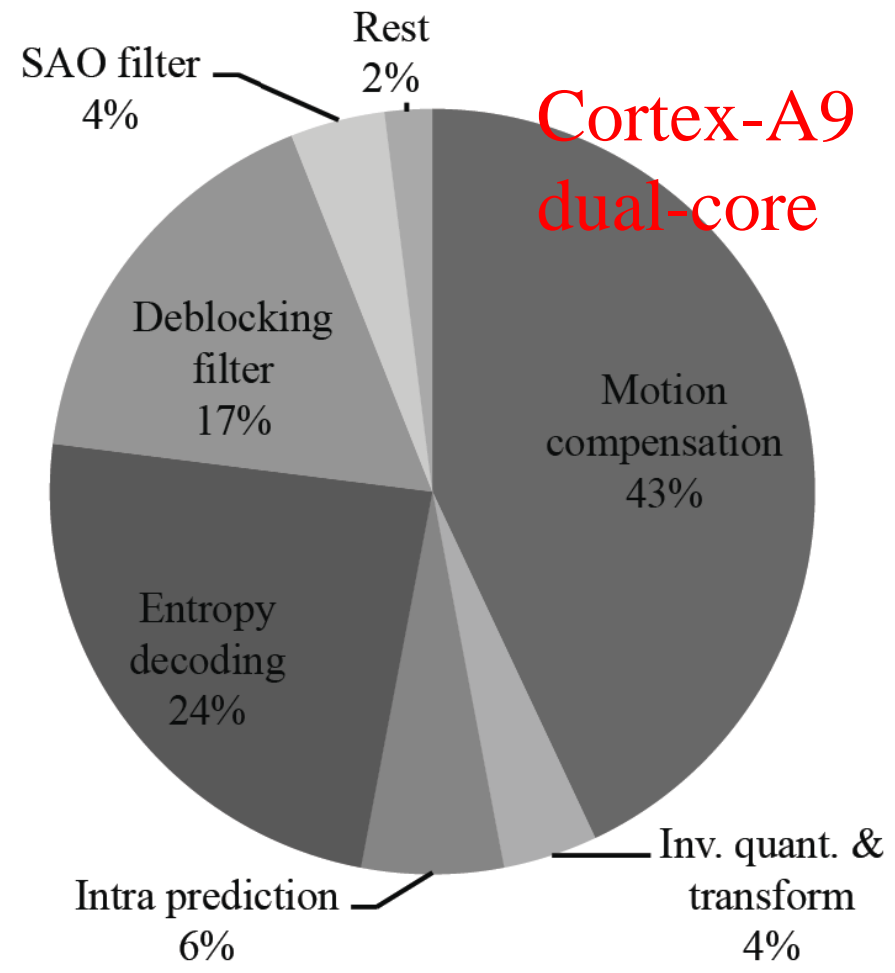


# 解码器不同模块的运算量 同一视频序列在不同处理器下对比

## Random Access (x86)



## Random Access (ARM)





# 解码器不同模块的运算量

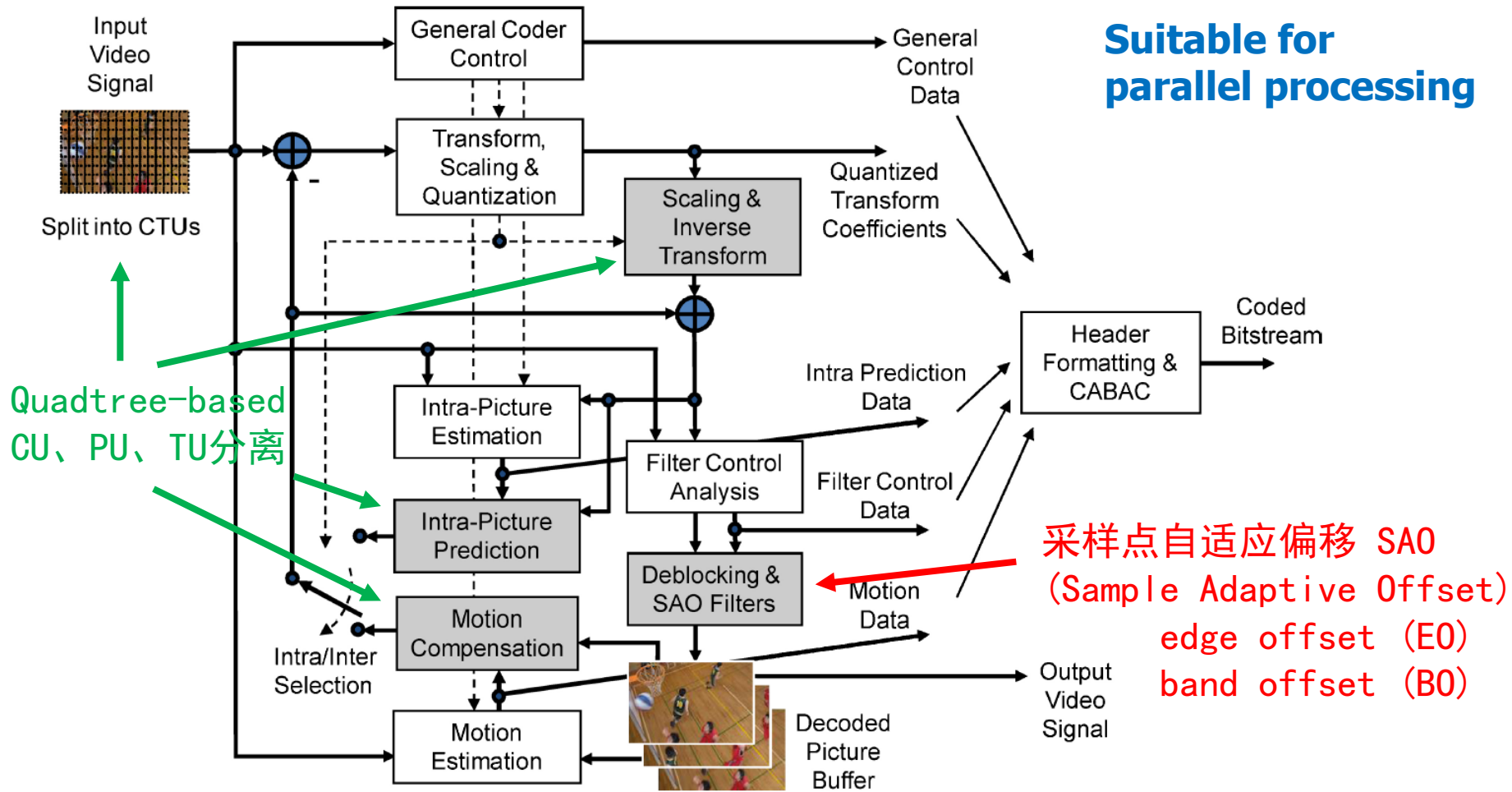
## Cortex-A9 dual-core不同视频序列对比

Sequence	MC [%]	QT [%]	PR [%]	ED [%]	DF [%]	SF [%]
BasketballDrill QP=27	36	5	7	26	19	5
BasketballDrill QP=32	44	4	6	19	20	3
BQMall QP=27	46	4	5	23	16	3
BQMall QP=32	53	3	4	16	17	2
PartyScene QP=27	40	5	5	31	13	3
PartyScene QP=32	49	3	5	22	15	3
RaceHorses QP=27	35	4	7	30	16	5
RaceHorses QP=32	43	4	6	22	19	4
Average	43	4	6	24	17	4

MC: motion compensation. QT: inverse quantization and transform. PR: intra prediction and picture construction process. ED: entropy decoding. DF: deblocking filter. SF: sample-adaptive offset filter.



# 小结: HEVC(H.265)创新点





# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想: MPEG-4(1999)
  - 视频编码可伸缩性 & H.264 SVC
  - 多视角编码



# MPEG1/MPEG2→MPEG4

- ◆ MPEG-1、MPEG-2、H.261、H.263
  - 基于图像块、变换、统计编码
- ◆ MPEG-4：提出基于模型/对象的压缩编码技术
  - 充分利用了人眼视觉特性，抓住了图像信息传输的**本质**，从**轮廓**、**纹理**思路出发，支持**基于视觉内容的交互功能**，这适应了多媒体信息的应用由播放型转向基于内容的访问、检索及制作的发展趋势。
- ◆ 不考虑内容→基于内容
- ◆ Block based → Object based

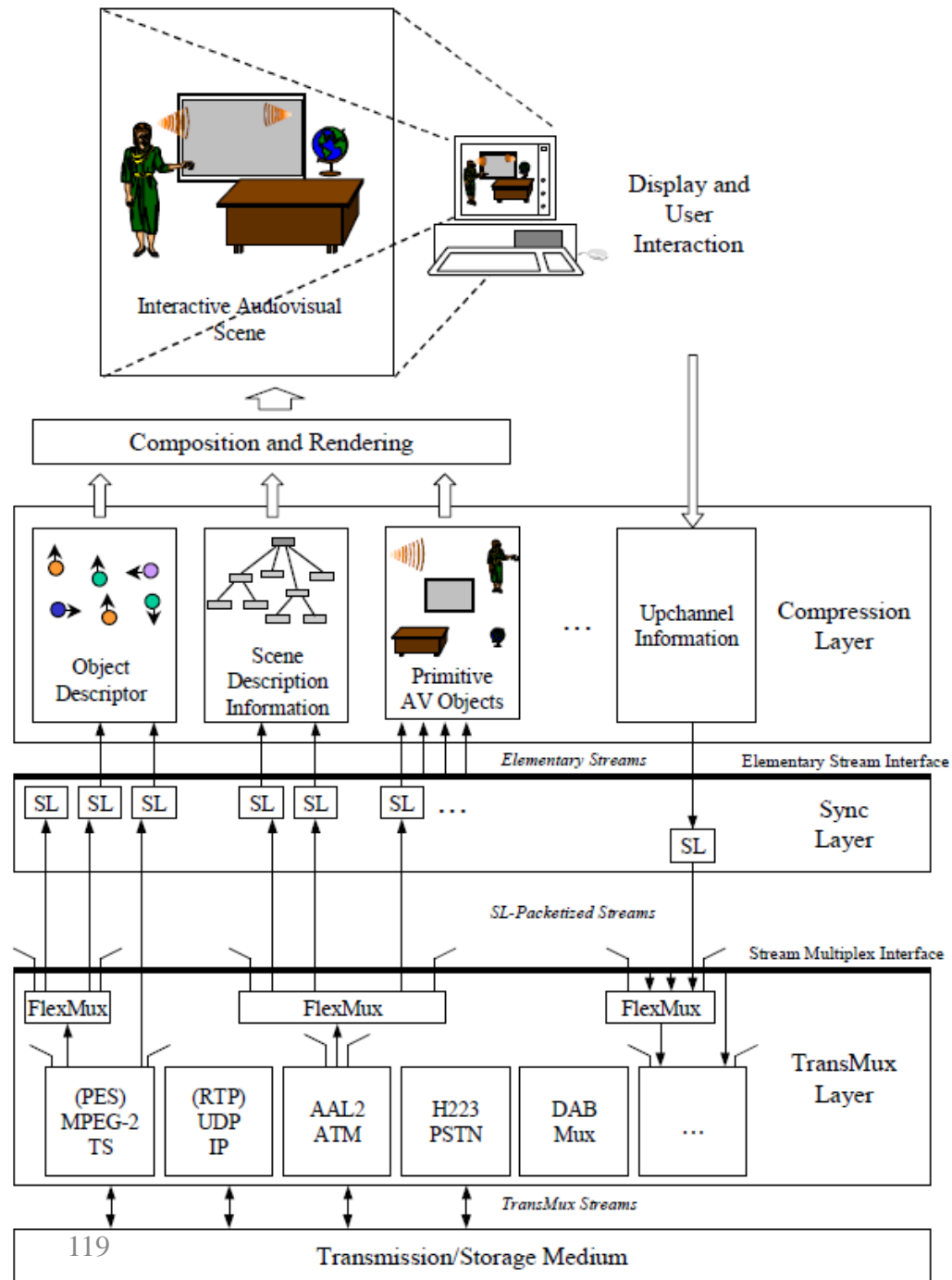
◆ 由于因特网的流行以及数字电视的即将兴起，ISO/IEC组织制定了MPEG-4国际标准，最初命名为“极低位率下的音视频编码”，到了1994年又更名为“音视频对象编码”，因此MPEG-4拥有极低位率及对象等特性的编码标准，其适用的范围主要为下列三大领域：数字电视，交互式图形应用程序，交互式多媒体。

◆ MPEG-4标准与先前的MPEG-1及MPEG-2标准有相当大的差异，过去是整个画面为基础的动画影像压缩技术，而MPEG-4则提供一套以**音视频对象**为基础的场景呈现方式，可使自然或合成的各种不同压缩编码方式的多媒体数据同时呈现在同1个场景中。通过音视频对象的方式，使用者对于音视频信息能够做到更加复杂的互动，不再只是单纯的播放、暂停、终止等简单的动作。



# MPEG-4 系统架构

- ◆ Compression Layer
- ◆ Synchronization Layer
- ◆ TransMux Layer





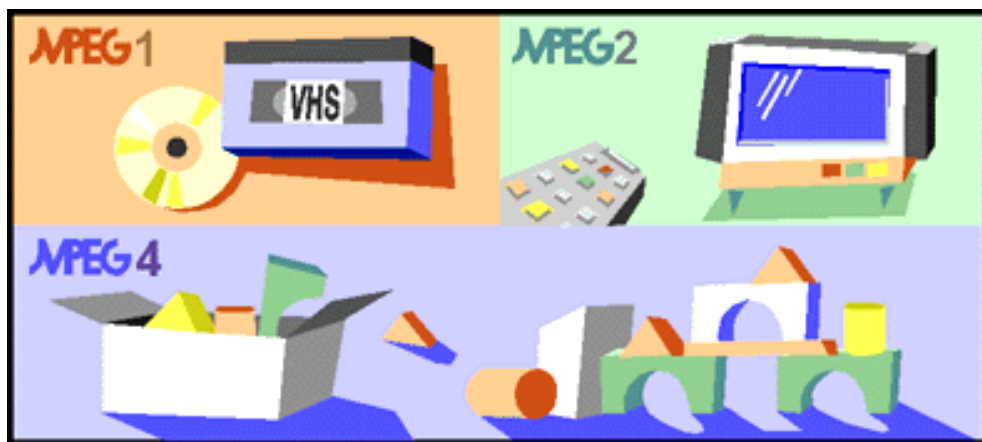
# MPEG-4由一系列的子标准组成

- ◆ part1 (ISO/IEC 14496-1): 系统: 描述视频和音频的同步以及混合方式(multiplexing)。
- ◆ **part2(ISO/IEC 14496-2): 视频: 定义了对视觉信息的编解码器。**
- ◆ part3(ISO/IEC 14496-3): 音频: 定义了一个对各种音频信号进行编码的编解码器的集合。
- ◆ part4(ISO/IEC 14496-4): 一致性: 定义了对本标准其他的部分进行一致性测试的程序。
- ◆ part5(ISO/IEC 14496-5): 参考软件。
- ◆ part6(ISO/IEC 14496-6): 多媒体传输集成框架。
- ◆ part7(ISO/IEC 14496-7): 优化的参考软件。
- ◆ part8(ISO/IEC 14496-8): 在IP网络上传输: 定义了在互联网上传输MPEG-4内容的方式。
- ◆ part9(ISO/IEC 14496-9): 参考硬件。
- ◆ **part10(ISO/IEC 14496-10): AVC, ITU-T H.264。**
- ◆ part12(ISO/IEC 14496-12): 基于ISO的媒体文件格式。
- ◆ **part13(ISO/IEC 14496-13): 知识产权管理和保护。**
- ◆ **part14(ISO/IEC 14496-14): MPEG-4文件格式。**
- ◆ part15(ISO/IEC 14496-15): AVC文件格式。
- ◆ part16(ISO/IEC 14496-16): 动画框架扩展。
- ◆ part17(ISO/IEC 14496-17): 同步文本字幕格式。
- ◆ part18(ISO/IEC 14496-18): 字体压缩和流式传输。
- ◆ part20(ISO/IEC 14496-20): 简单场景表示。
- ◆ part21(ISO/IEC 14496-21): 用于描绘(Rendering)的MPEG-J拓展。



# MPEG-4视频编码核心思想

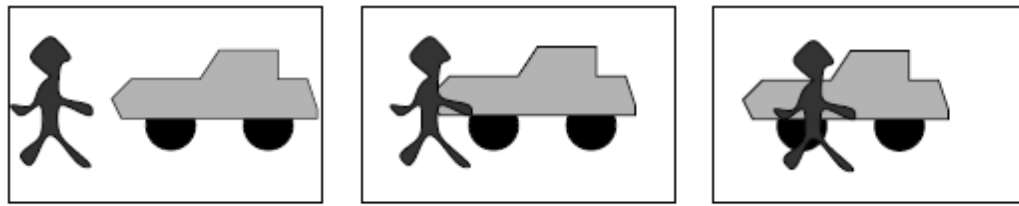
◆视听对象（AVO, Audio Visual Object）是MPEG-4为支持基于内容编码而提出的重要概念。对象是指在一个场景中能够访问和操纵的实体，对象的划分可根据其独特的纹理、运动、形状、模型和高层语义为依据。在MPEG-4中所见的视音频已不再是过去MPEG-1、MPEG-2中图像帧的概念，而是一个个视听场景，这些不同的视听场景由不同的视听对象组成。



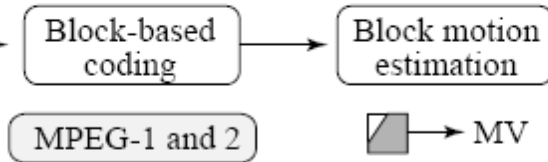
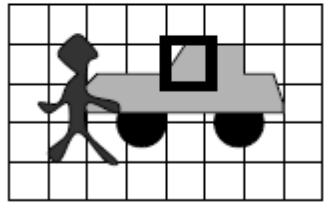
Previous frame

Current frame

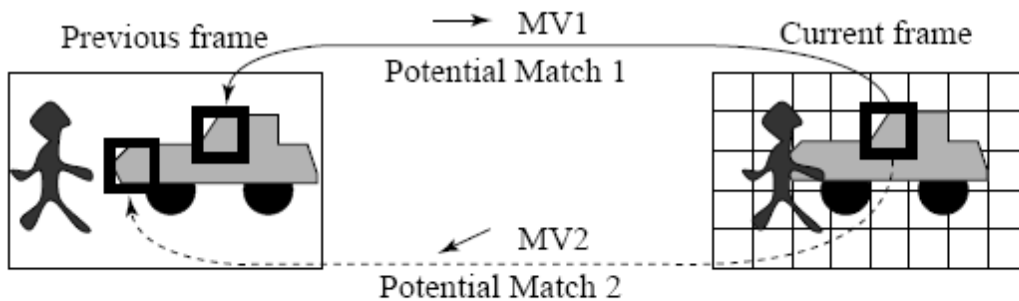
Next frame



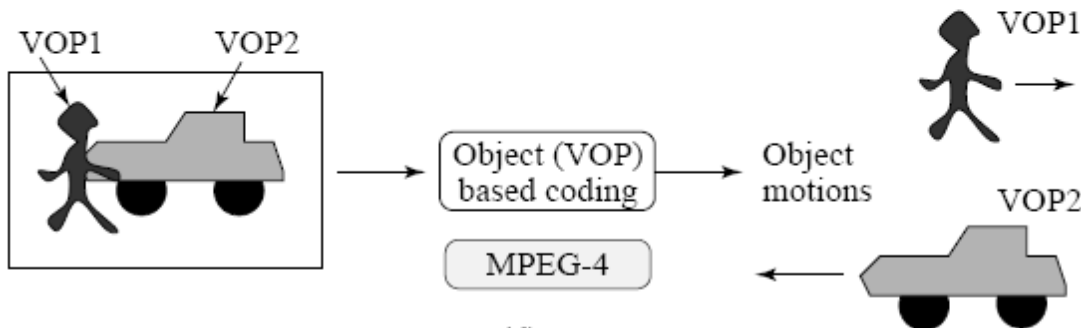
(a)



(b)



(c)



(d)

# Block based & Object based

◆ A

□ 原始图像

◆ B

□ 基于宏块的

◆ C

□ 引起宏块方式中预测错误

◆ D

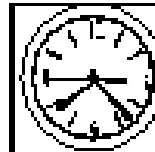
□ 基于对象的

# Audio Object

廣播聲

火車進  
站聲

Welcome to MPEG-4



Train to  
MPEG-4 07:30

Attention  
please!

背景音  
樂

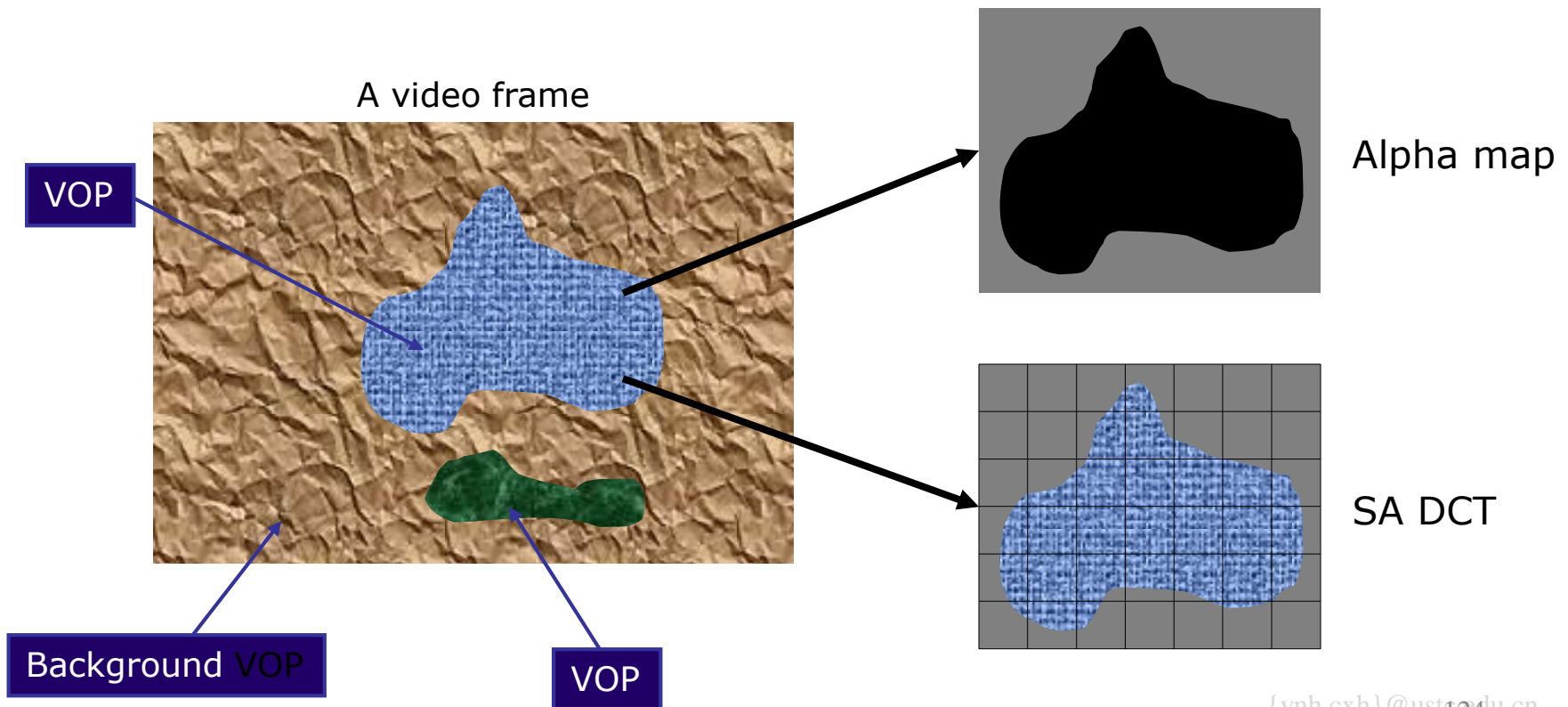
交談聲





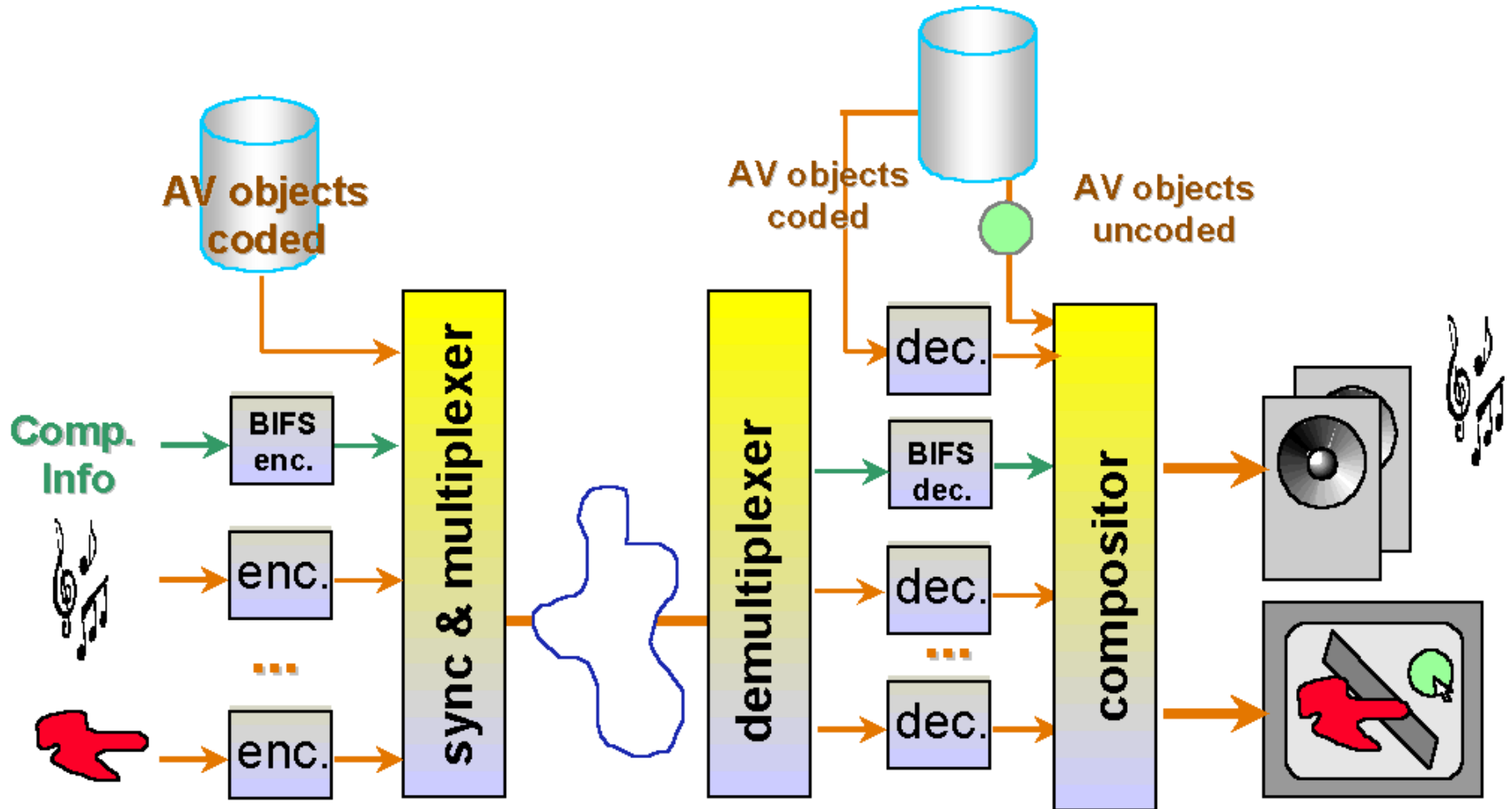
# MPEG-4 Video Object

- ◆ Instead of "frames": **Video Object Planes**
- ◆ **Shape Adaptive DCT**





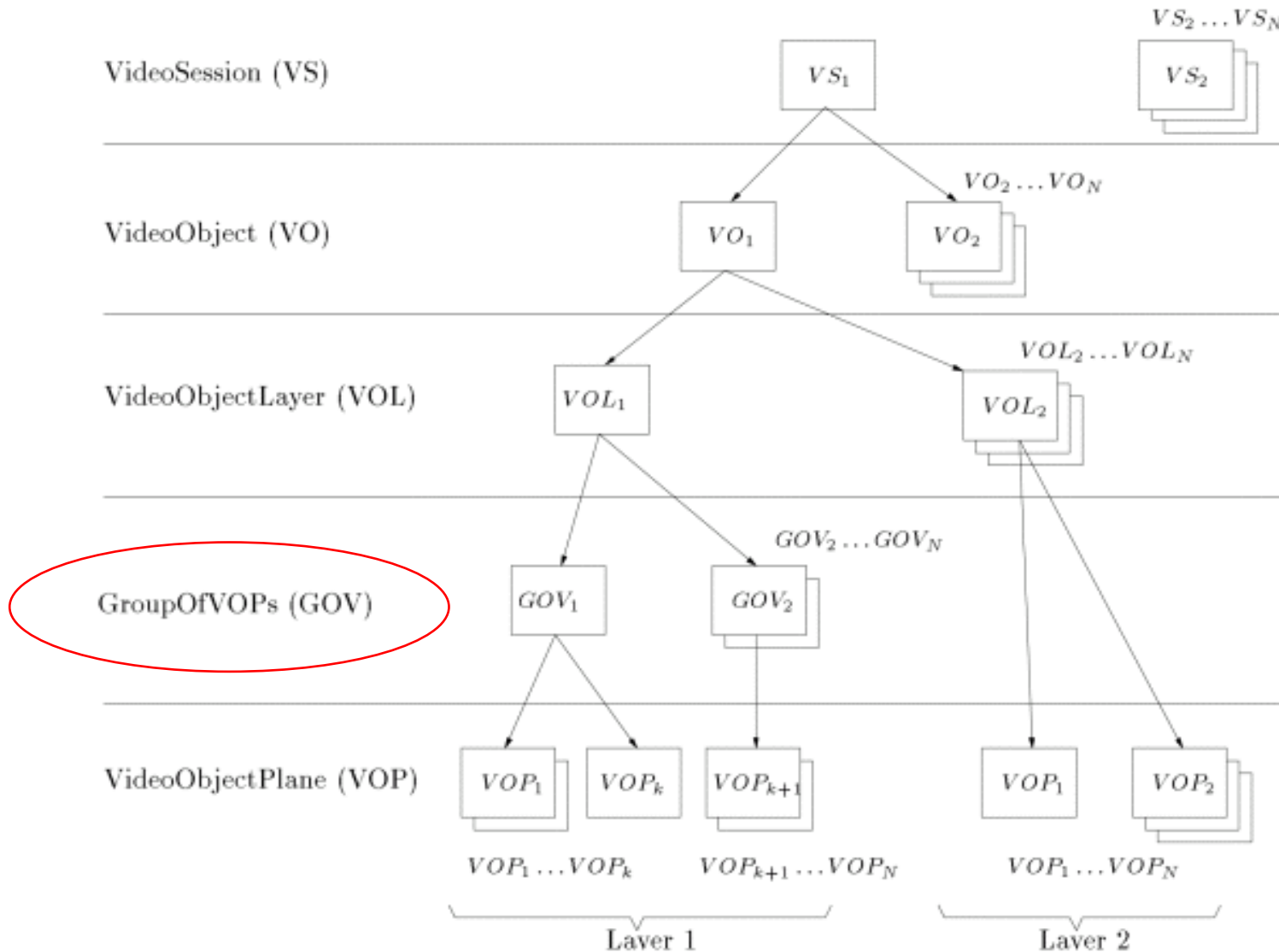
# MPEG-4 Structure





# GOP → GOV

## MPEG 4 – Structure and Syntax



# MC 小结：MPEG4

- ◆思路：不考虑内容→基于内容
- ◆处理单元：Block based → Object based
- ◆文件格式：GOP → GOV
  
- ◆part10(ISO/IEC 14496-10): AVC, ITU-T H.264
- ◆part13(ISO/IEC 14496-13): 知识产权管理和保护



# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想： MPEG-4(1999)
  - 视频编码可伸缩性 & H.264 SVC
  - 多视角编码



# MPEG-2的配置 profile

配置(Profile)	算法(Algorithms)
High (高档)	<ul style="list-style-type: none"><li>● 支持由空间分辨率可变配置(Spatial Scalable Profile)提供的所有功能和其他规定功能</li><li>● 子采样各种: YUV 4:2:2 用于进一步提高图像质量</li></ul>
Spatial scalable (空间分辨率可变)	<ul style="list-style-type: none"><li>● 支持信噪比可变配置(SNR Scalable Profile)提供的所有功能和空间分辨率可变(Spatial scalable coding)算法(2层)</li><li>● 子采样格式: YUV: 4:2:0</li></ul>
SNR scalable (信噪比可变)	<ul style="list-style-type: none"><li>● 支持基本配置(Main Profile)提供的所有功能和信噪比可变编码(SNR scalable coding)算法(2层)</li><li>● 子采样格式: YUV: 4:2:0</li></ul>
Main (基本)	<ul style="list-style-type: none"><li>● 非可变速率编码算法支持随机存取, B图像预测方式</li><li>● 子采样格式: YUV: 4:2:0</li></ul>
Simple (简化)	<ul style="list-style-type: none"><li>● 除不支持基本配置(Main Profile)提供的B图像预测功能外, 基本配置的其他所有功能都支持</li><li>● 子采样格式: YUV: 4:2:0</li></ul>



# MPEG-2等级 level

## ◆ MP@ML

- Main Profile, Main Level
- 720 × 576 × 30
- 4:2:0
- 15 Mb/s

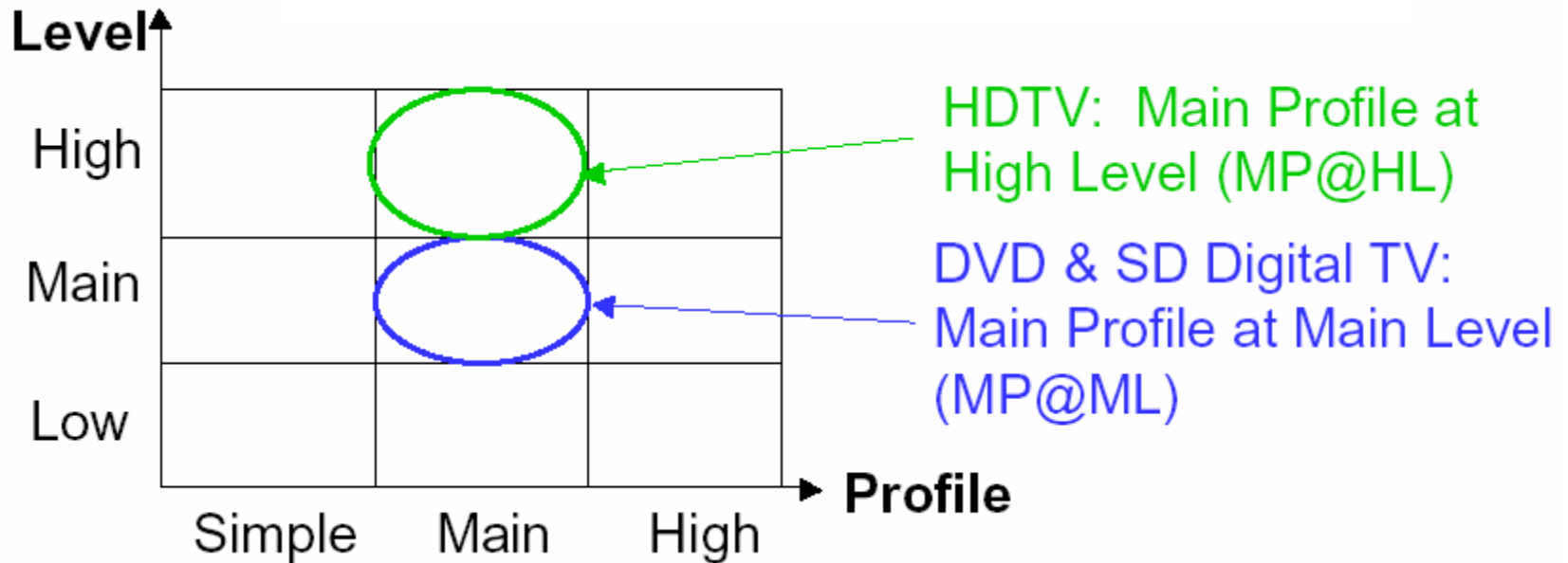
## ◆ MP@HL

- Main Profile, High Level
- 1920 × 1152 × 60
- 4:2:0
- 80 Mb/s

等级 (Level)	参数 (Parameters)	说明
	1920 samples/line	1920 样本/行
HIGH	1152 lines/frame	1152 行/帧
(高级)	60 frames/s	60 帧/秒
	80 Mb/s	80 兆比特/秒
	1440 samples/line	1440 样本/行
HIGH 1440	1152 lines/frame	1152 行/帧
(高级1440)	60 frames/s	60 帧/秒
	60 Mb/s	60 兆比特/秒
	720 samples/line	720 样本/行
MAIN	576 lines/frame	576 行/帧
(基本级)	30 frames/s	30 帧/秒
	15 Mb/s	15 兆比特/秒
	352 samples/line	352 样本/行
LOW	288 lines/frame	288 行/帧
(低级)	30 frames/s	30 帧/秒
	4 Mb/s	4 兆比特/秒



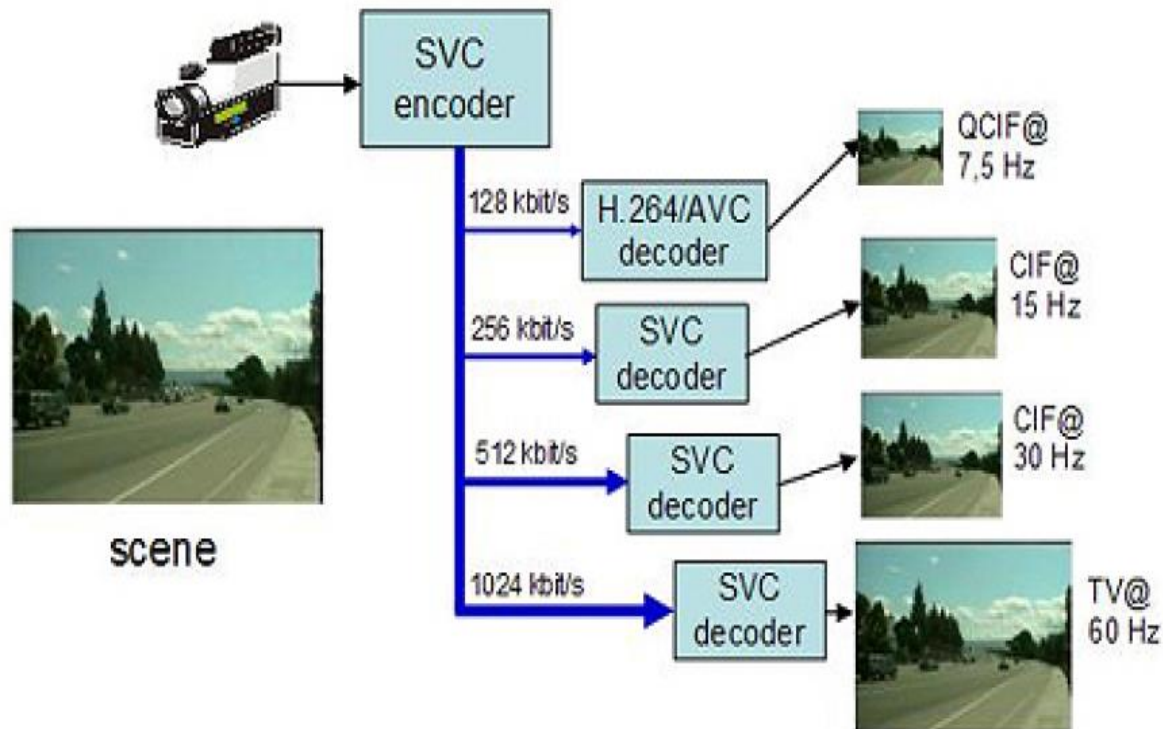
# MPEG2中常用Profile@Level





# Scalable Video Coding (SVC) extension of the H.264/AVC standard

- ◆可伸缩性编码只需对视频源编码一次，在解码时就可以获得不同分辨率、码率、帧率或质量的视频





# H.264: SVC Basics

- ◆ In **Spatial scalability** and **Temporal Scalability** the subset of the bit-stream represent the source content with reduced picture size (Spatial Resolution) or frame rate (Temporal Resolution).
- ◆ In case of **quality scalability**, also known as fidelity or SNR scalability, the subset of the bit-stream provides lower quality. (Lower SNR).
- ◆ In rare cases, “**region-of-interest**” and **object based scalability** is also required, wherein the subsets of the bit-stream represent spatially contiguous regions of original picture area.
- ◆ Multiple scalability features can be combined to support various spatio-temporal resolutions and bit rates within single bit-stream.



# 小结：视频编码的可伸缩性

## ◆ MPEG1/2

- Spatial scalability
- Quality scalability (SNR scalability)

## ◆ MPEG4 AVC/H.264

- Temporal Scalability
- Spatial scalability
- Quality scalability (SNR scalability)

## ◆ 未来？

- region-of-interest
- object based scalability

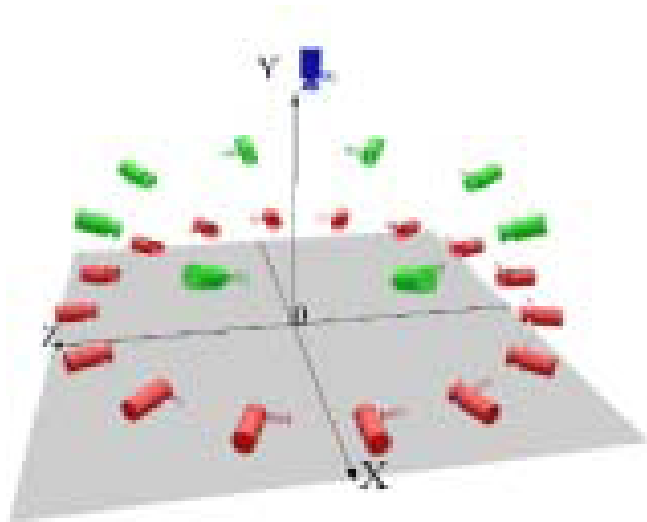


# 第2章 多媒体数据压缩国际标准

- ◆ 2.1 多媒体数据压缩编码的重要性和分类
- ◆ 2.2 常见数据压缩方法分类与基本原理
- ◆ 2.3 音频压缩标准
- ◆ 2.4 静态图像压缩编码的国际标准
- ◆ 2.5 视频压缩的国际标准
  - 视频编解码器的框架与标准简介
  - MPEG-1(1993) & MPEG-2/H.262(1996)视频标准
  - MPEG-4 AVC/H.264(2003)视频标准
  - HEVC(2013)视频标准
- ◆ 2.6 不同应用场景下的其他编码思想
  - 基于模型/对象的编码思想： MPEG-4(1999)
  - 视频编码可伸缩性 & H.264 SVC
  - 多视角编码： FVV、MVC



# 任意视角视频 (FVV) Free Viewpoint Video



**FVV** 采集设备图例



同一时刻不同视角的视频

# MC 任意视角视频 (FVV)

## 问题:

总数据量 = 单视角的数据量  $\times$  视角数

随着视角数的增加，要传输的数据量很快变得无法承受，必须想办法减少数据量。

减少数据量  $\rightarrow$  消除冗余

## 问题:

原有的视频编码框架已经较好的解决了帧间冗余（时域冗余）、帧内冗余（空间冗余）、统计冗余问题，那么，在FVV中又可以利用什么冗余信息？



# 任意视角视频 (FVV) 视角间冗余



视角1



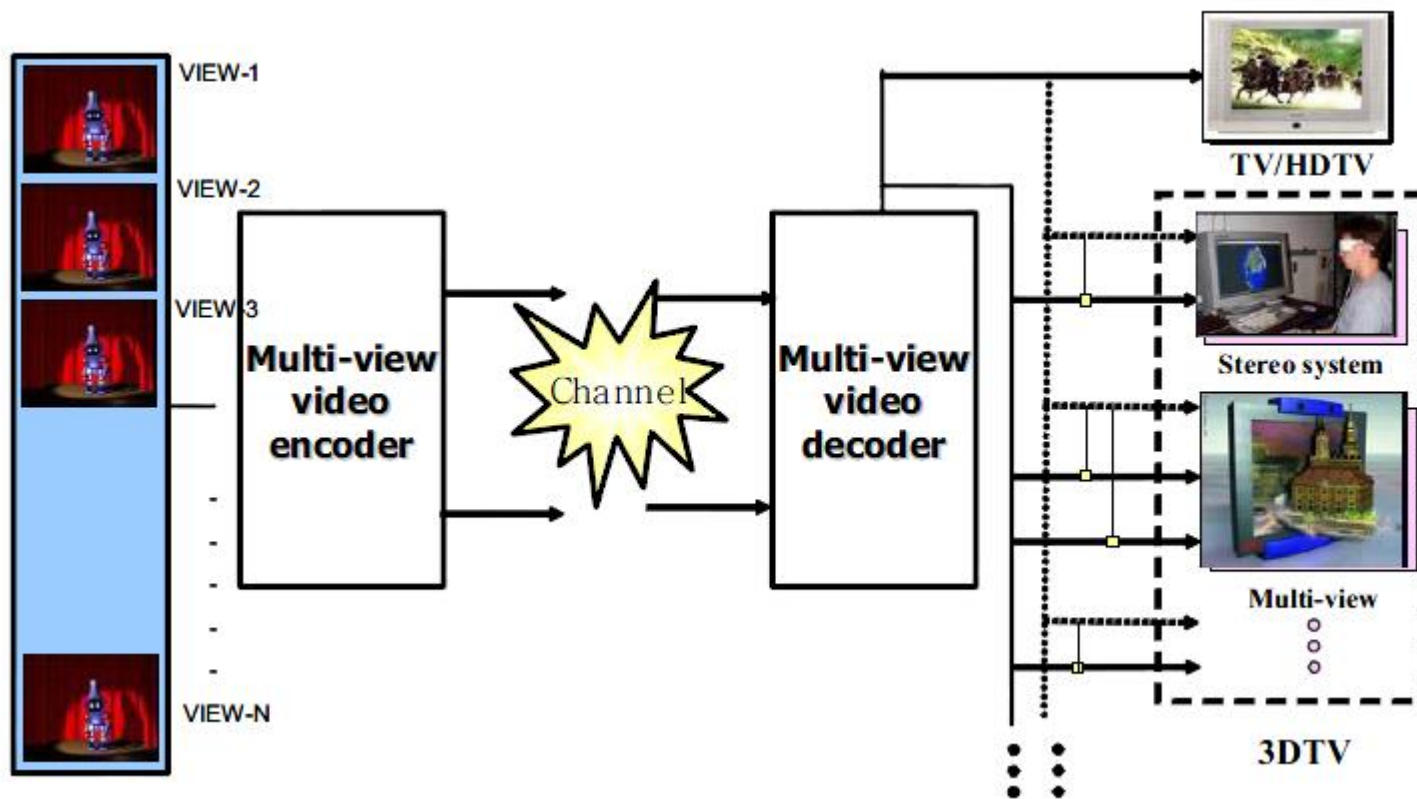
视角2

## 视角间冗余

为满足FVV的需要，人们提出多视角编码 (Multi-view Video Coding, MVC)，以充分利用视角间冗余，减少传输或存储的数据量。

# MC 多视角编码 (MVC)

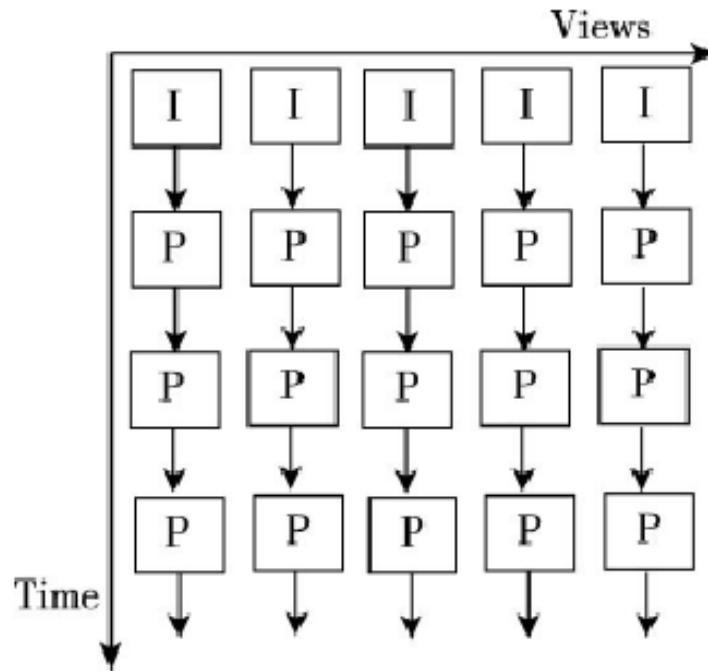
多视角编码 (**Multi-view Video Coding**)的系统框架:



# MC MVC的预测模型

## (1) Simulcast结构:

时间上只采用了前向预测；未考虑视角间的相关性。  
编码效率低

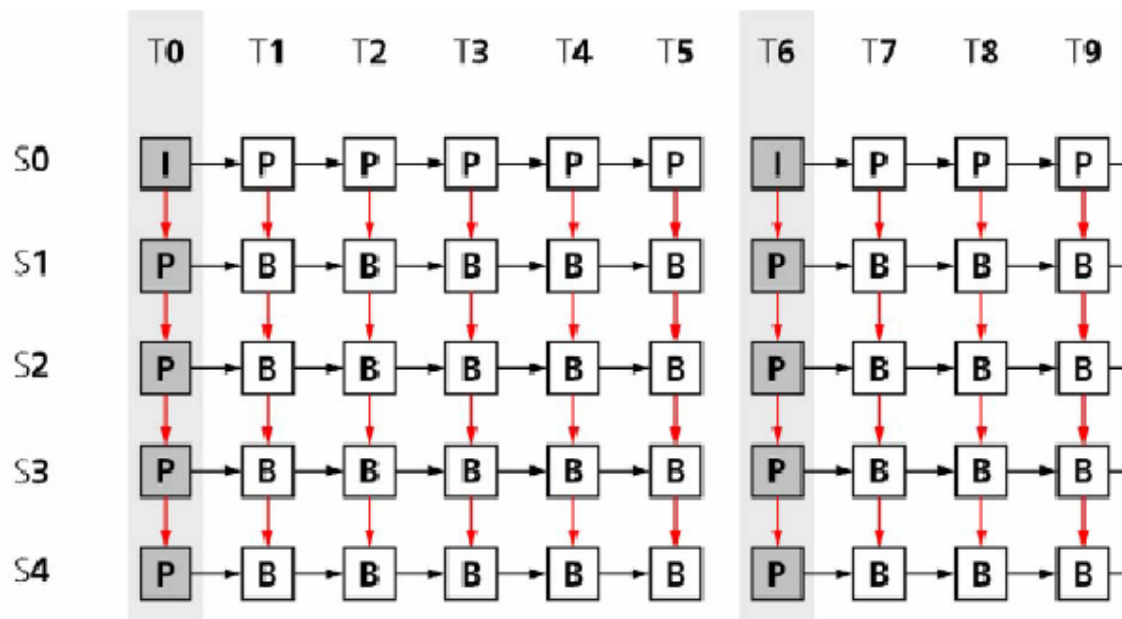


# MC MVC的预测模型

## (2) 视间和时间双向预测结构:

第一个视角的第一帧为关键帧，采用帧内预测编码(I帧)，其余帧采用时间上的帧间预测编码;

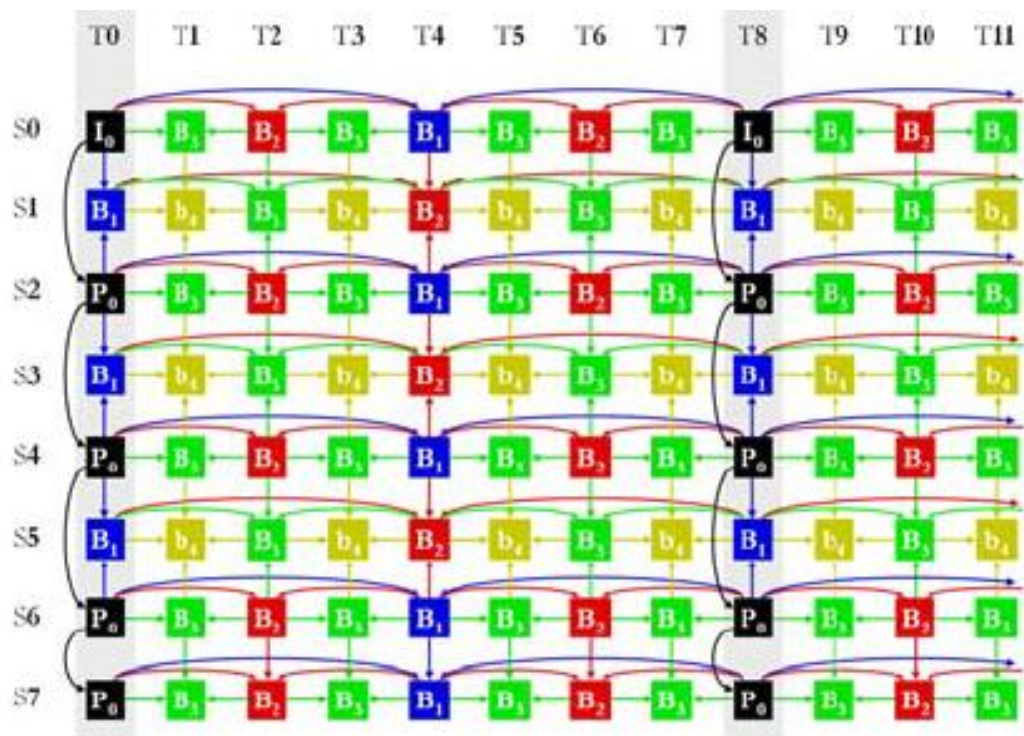
其余视角的第一帧参考关键帧采用视间预测编码，其余帧采用帧间和视间双向预测编码。



# MC MVC的预测模型

## (3) 基于分层B帧的预测结构:

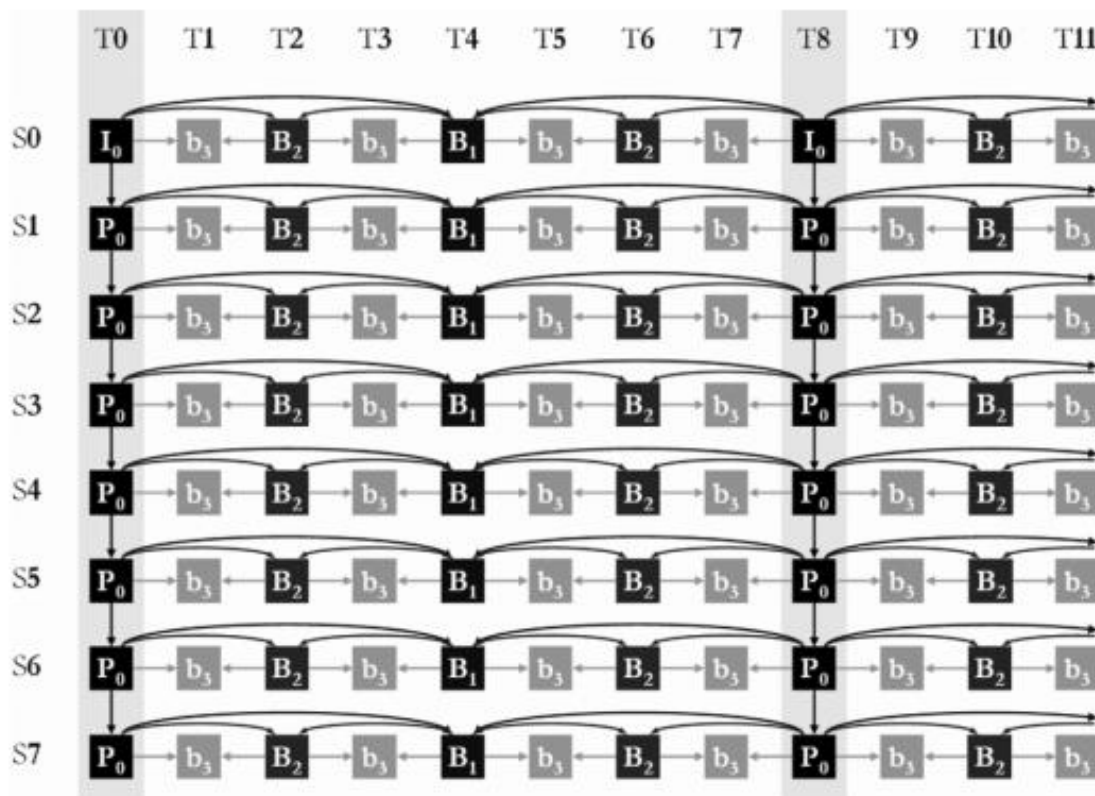
视间预测发生在奇数视角，如：**S1**、**S3**和**S5**；偶数视角位置的**B**帧不进行视间预测，只进行时间上的预测。



# MC MVC的预测模型

## (4) 仅关键帧作视间预测:

只对各非基本视角 (S1-S7) 的第一帧进行视间预测, 其余帧均采用时域帧间预测

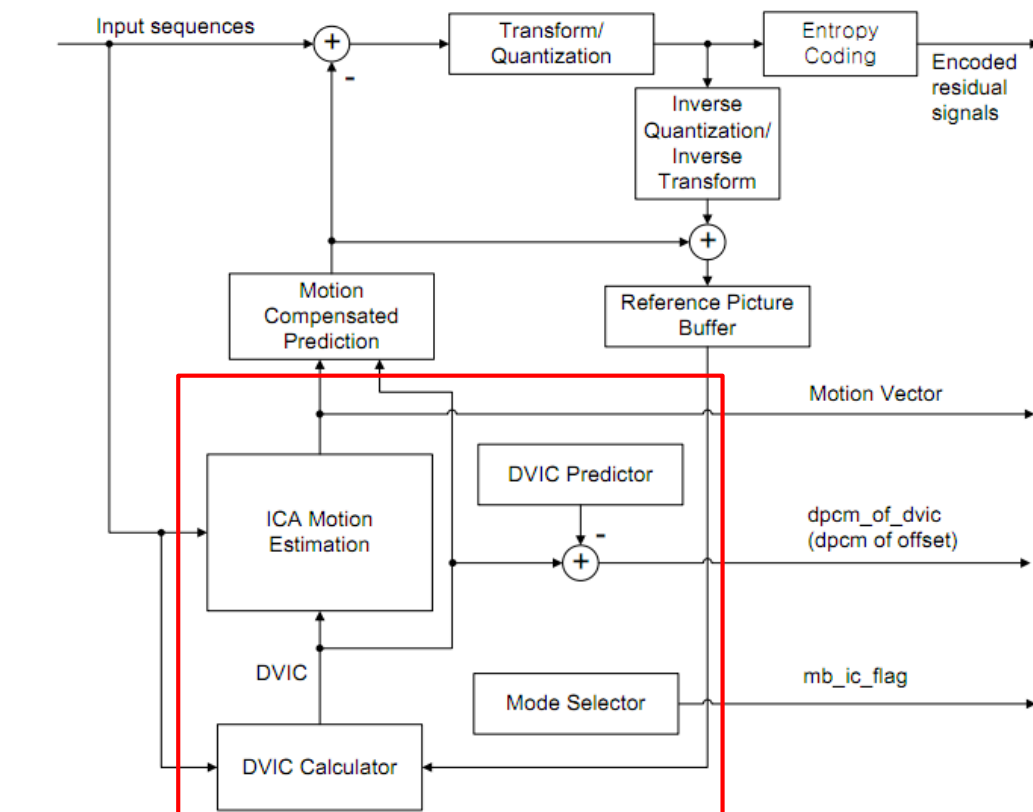


# MC MVC的编码器框架

与单视角编码相比，MVC编码器的框架中包含了视间预测的部分，如红色方框所示。

DVIC Calculator模块：计算编码块和参考块亮度均值的差；

ICA Motion Estimation模块：在亮度补偿后的参考视角帧中搜索最匹配块。



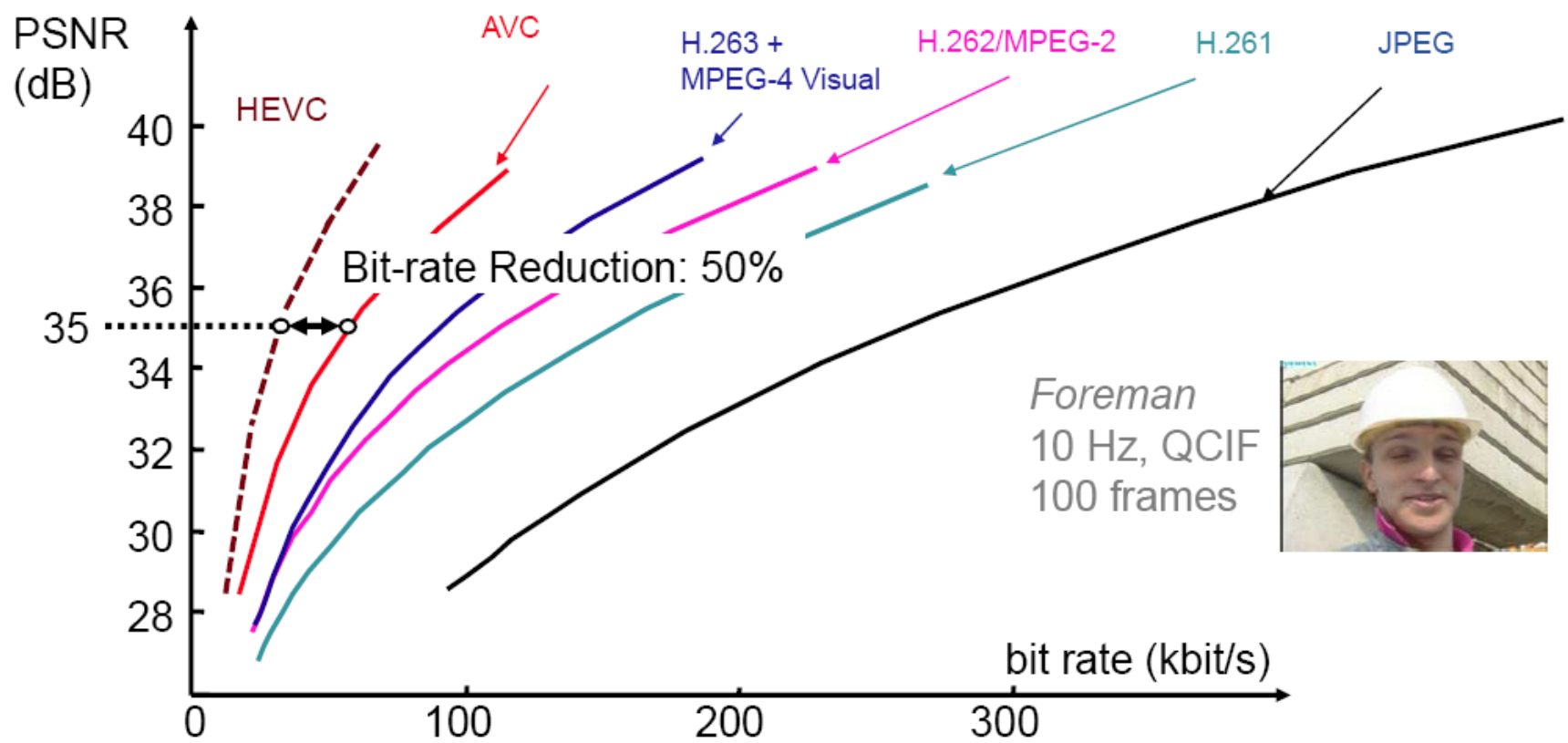
# MC 小结：多视角视频编码

- ◆ Temporal
  - ◆ Spatial
  - ◆ Entropy
  - ◆ **Inter-View**
- 
- ◆ 不同冗余的联合处理？
    - Simulcast结构
    - 视间和时间双向预测结构
    - 基于分层B帧的预测结构
    - 仅关键帧作视间预测



# 未来? 是否必要开发新的视频压缩标准?

## Performance history of standard generations





# 下一代的视频编码标准

## VVC(Versatile Video Coding)

In October 2015, MPEG and VCEG formed Joint Video Exploration Team (JVET) to evaluate available compression technologies and study the requirements for a next-generation video compression standard. The new algorithm should have **30-50% better compression rate** for the same perceptual quality, with support for lossless and subjectively lossless compression. It should also support YCbCr 4:4:4, 4:2:2 and 4:2:0 with 10 to 16 bits per component, BT.2100 wide color gamut and **high dynamic range (HDR)** of more than 16 stops (with peak brightness of 1000, 4000 and 10000 nits), auxiliary channels (for depth, transparency, etc.), variable and fractional frame rates from 0 to 120 Hz, scalable video coding for temporal (frame rate), spatial (resolution), SNR, color gamut and dynamic range differences, **stereo/multiview** coding, **panoramic** formats, and still picture coding. Encoding complexity of **10 times** that of HEVC is expected. JVET issued a final "Call for Proposals" in October 2017, with the first working draft of the Versatile Video Coding standard released in April 2018; the final standard is to be approved before the end of 2020.



# 下一代的视频编码标准 VVC的增强

- **Block structure**

- Larger Coding Tree Unit (up to 256x256) and transforms (up to 64x64)

- Quadtree plus binary tree (QTBT) block structure replaced quadtree structure of HEVC

- **Intra prediction tools**

- 65 intra prediction directions

- 4-tap interpolation filter for intra prediction

- Boundary filter applied to other directions in addition to horizontal and vertical ones

- Cross-component linear model (CCLM) prediction

- Position dependent intra prediction combination (PDPC)

- **Transform**

- Explicit multiple core transform

- Mode dependent non-separable secondary transforms in intra coding

- **Inter prediction**

- Subblock level motion vector prediction

- Locally adaptive motion vector resolution (AMVR)

- 1/16 pel motion vector storage accuracy

- Overlapped block motion compensation (OBMC)

- Local illumination compensation (LIC)

- Affine motion prediction

- Pattern matched motion vector derivation at decoder

- Bi-directional optical flow (BIO) for improved motion compensation at decoder

- **In-loop filters**

- Adaptive loop filter (ALF)

- Bilateral filter

- **CABAC design**

- Context model selection for transform coefficient levels

- Multi-hypothesis probability estimation, improved initialization for context models

# MC 总结：视频编码

## ◆ 基于块的预测/变换混合编码框架

### □ MPEG-1(1993) & MPEG-2/H.262(1996)

- GOP → Picture → Slice → Macroblock → Block
- I帧、P帧、B帧

### □ MPEG4 AVC/H.264 (2003)

- 运动估计、帧内预测、去块效应滤波、整数变换、熵编码

### □ HEVC (2013)

- 灵活的编码结构
- CU、PU、TU分离
- 采样点自适应偏移 (SAO)

## ◆ 其他编码思想

### □ 基于模型/对象的编码思想： MPEG-4(1999)

### □ 视频编码可伸缩性 & H.264 SVC

### □ 多视角编码

## ◆ VVC(Versatile Video Coding)