

第10章 视频处理

1. 视频处理

- 在视频压缩前后，对视频图像质量增强的操作
- 视频编解码系统输出的图像主观质量不仅与压缩算法的性能有关，还受视频处理的影响
- 压缩之前对视频的处理称作预处理（Pre-processing）
- 压缩之后对视频的处理称作后处理（Post-processing）



2. 预处理

- 预处理的目的
 - 为了减少原图像受到的损害，保持原图像的重要特征，使原图像能被高效的压缩
 - 噪声污染
 - 光照差
 - 抖动
 - 为了进行视频格式转换
 - 去隔行
 - 空间缩放
 - 帧率转换
- 去噪声处理



受噪声污染的原图像



去噪声处理后的图像

- 图像噪声接近高斯分布
- 用高斯滤波器滤波能够有效去除噪声

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

- 其它简单滤波

1	3	1
3	16	3
1	3	1

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

$\frac{1}{273}$

二维高斯滤波整数近似 $\sigma=1.0$

- 光照



原图像



处理后图像

- 去隔行 (Deinterlace)
 - 隔行扫描的视频图像在逐行扫描播放设备上播放会有隔行效应，降低了主观质量



隔行图像



去隔行图像

3. 去隔行的方法

- 在单场内去隔行
 - 用奇数（偶数）行插值出偶数（奇数）行

X	X	X
?	?	?
X	X	X

x—偶数行像素值

? —奇数行像素值

- 两场间去隔行
 - 用两场的像素值（包括运动矢量信息）插值出一帧像素值

?	?	?
X	X	X
?	?	?

顶场

X	X	X
?	?	?
X	X	X

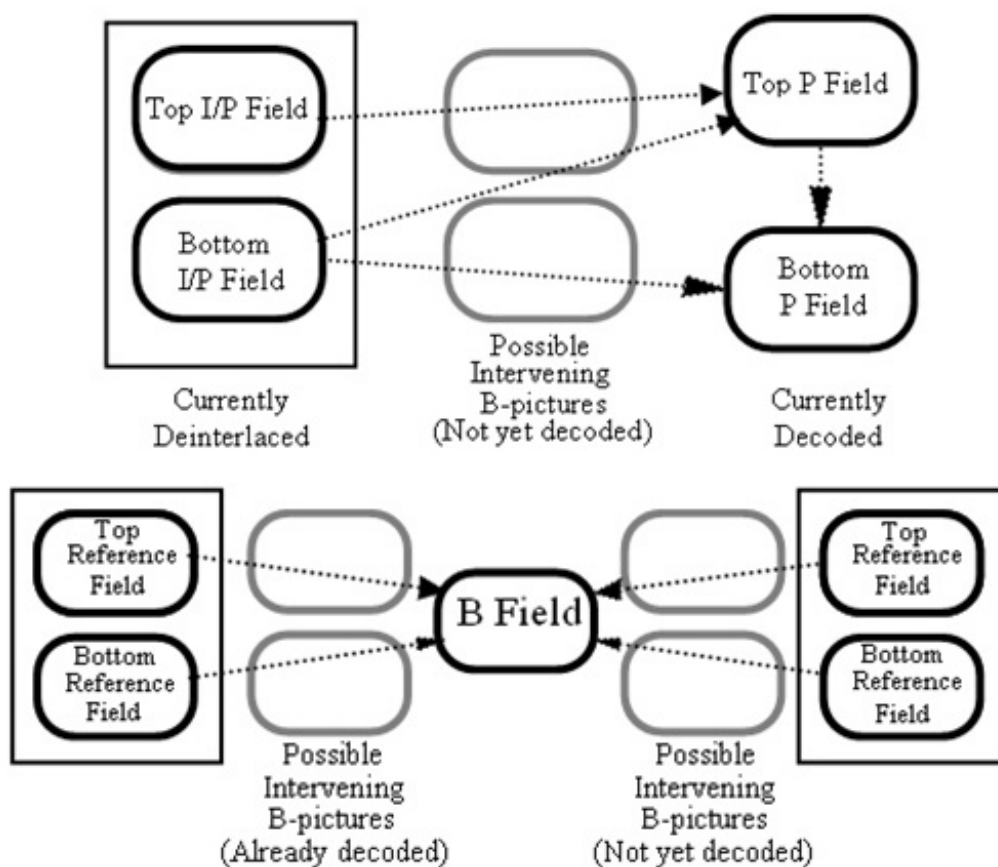
底场

- 运动补偿方法
 - 用前后向参考场的运动矢量信息，产生高质量的逐行帧

?	?	?	x	x	x	?	?	?
x	x	x	?	?	?	x	x	x
?	?	?	x	x	x	?	?	?
前向场			当前场			后向场		

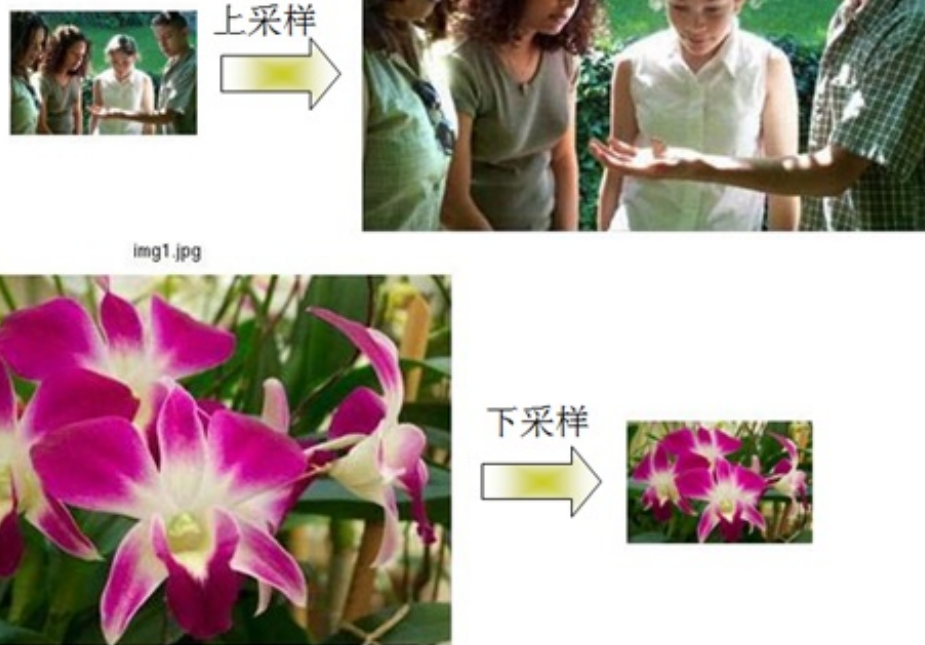
4. MPEG去隔行处理

- MPEG-2后处理去隔行



5. 空间缩放

- 空间缩放



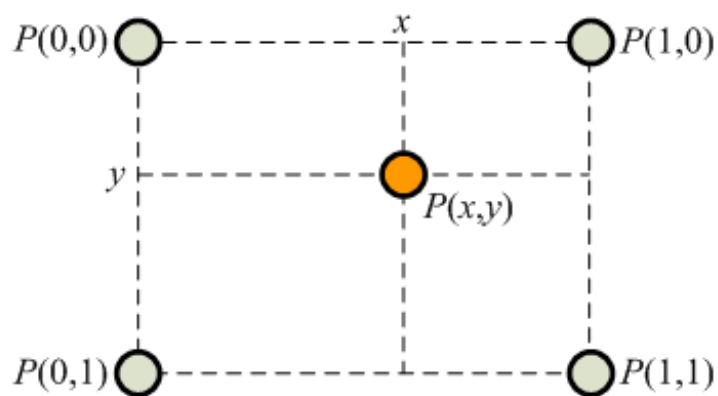
- 上采样 (Up sample) 方法

- 最近邻插值 (Nearest-neighbor interpolation)

- 用原图像上每一个像素点值作为上采样图像上对应四个相邻像素点的值
- 会产生阶梯形状的效应

- 线性，双线性插值

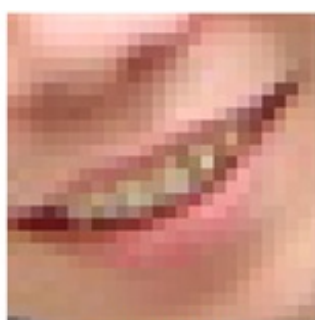
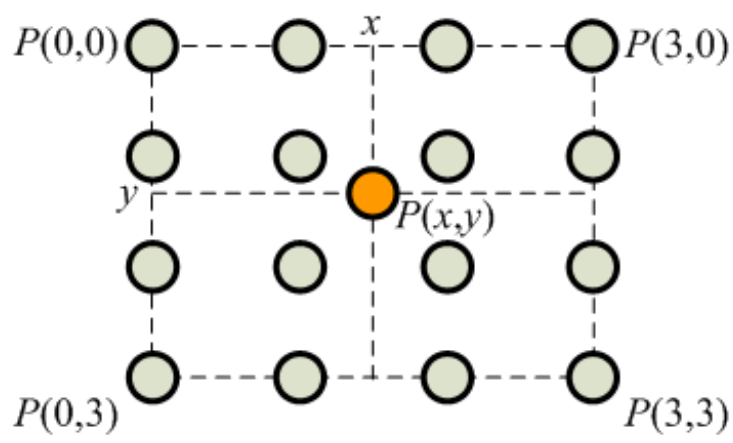
$$P(x,y)=P(0,0)xy+P(1,0)(1-x)y+P(0,1)x(1-y)+P(1,1)(1-x)(1-y)$$



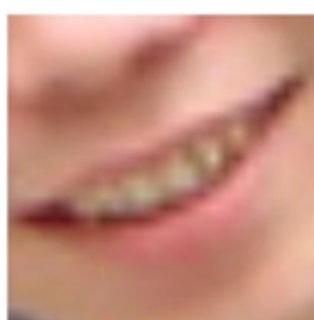
- 双三次插值 (Bicubic interpolation)

- 用邻近的16个像素值插值当前像素
- 有更平滑的插值效果

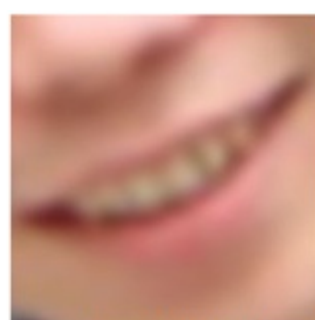
$$P(x,y) = \sum_{i=0}^3 \sum_{j=0}^3 P(i,j) \left(\frac{x}{4}\right)^i \left(\frac{y}{4}\right)^j$$



最近邻插值

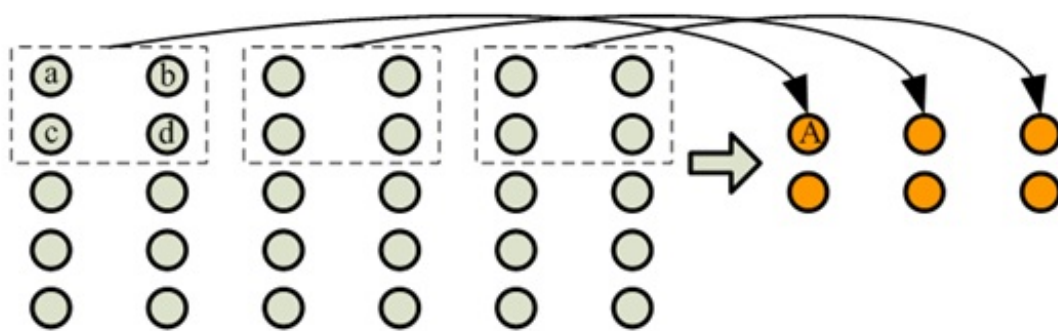


双线性插值



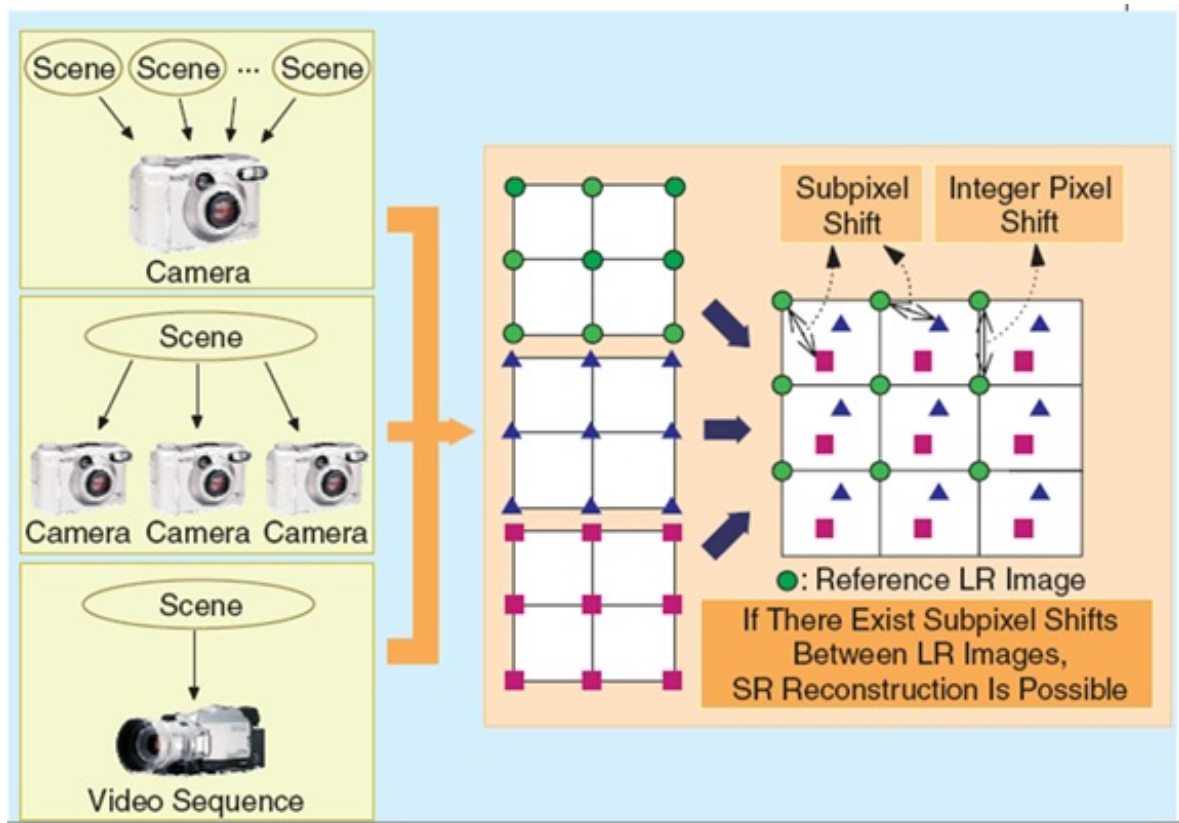
双三次插值

- 下采样 (Down sample) 方法
 - 缩小原始图像

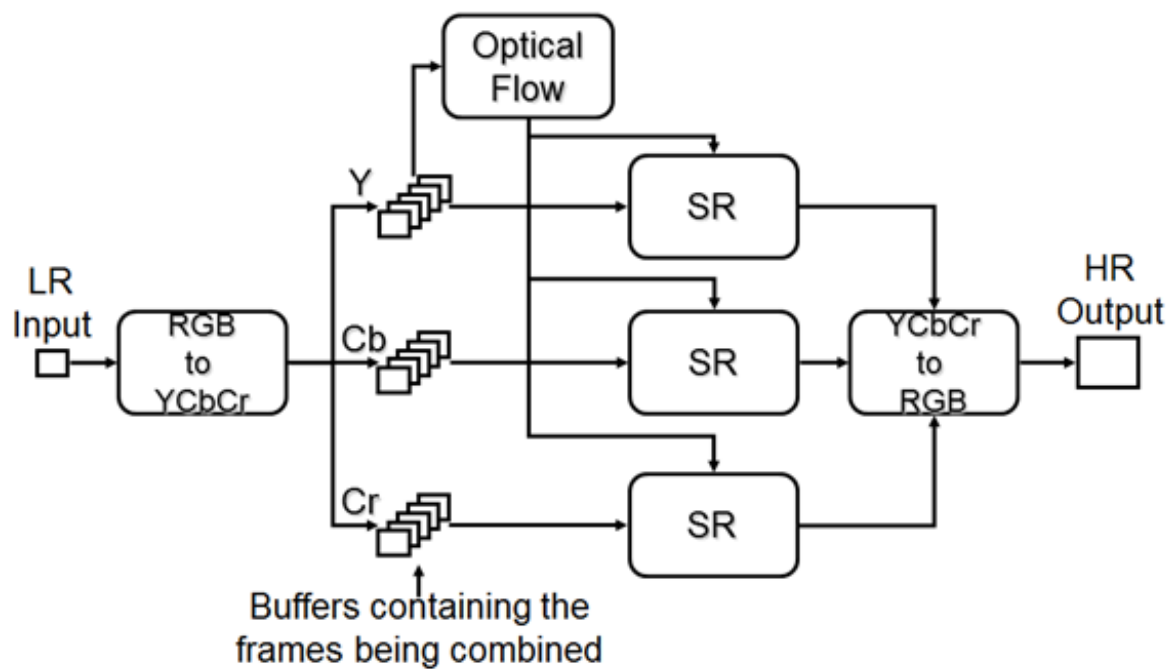


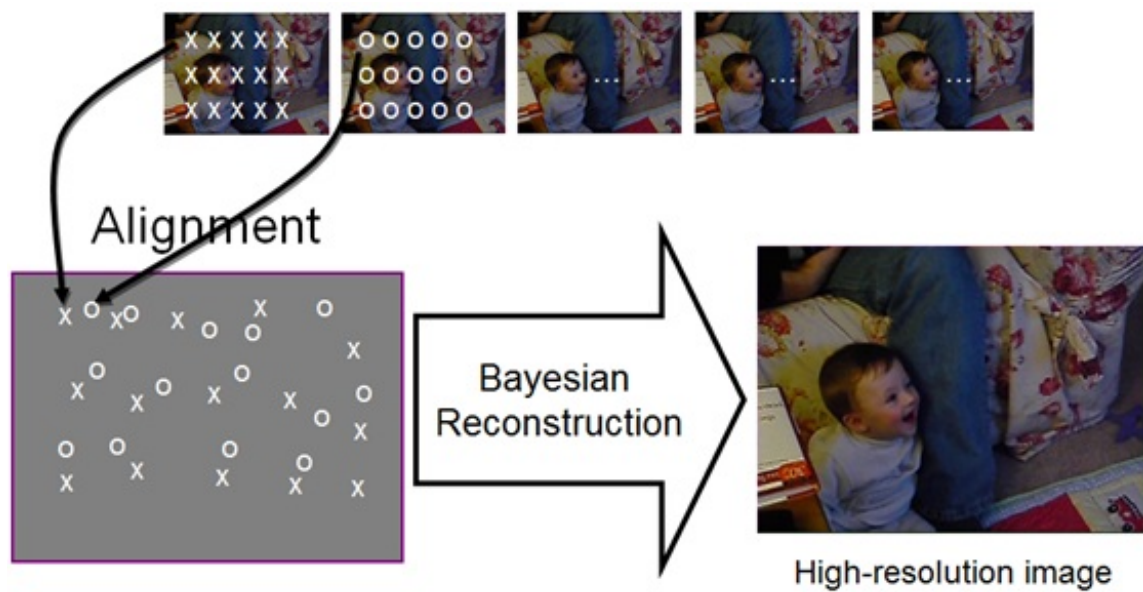
$$A = (a + b + c + d + 1) \gg 2$$

5. 超分辨率



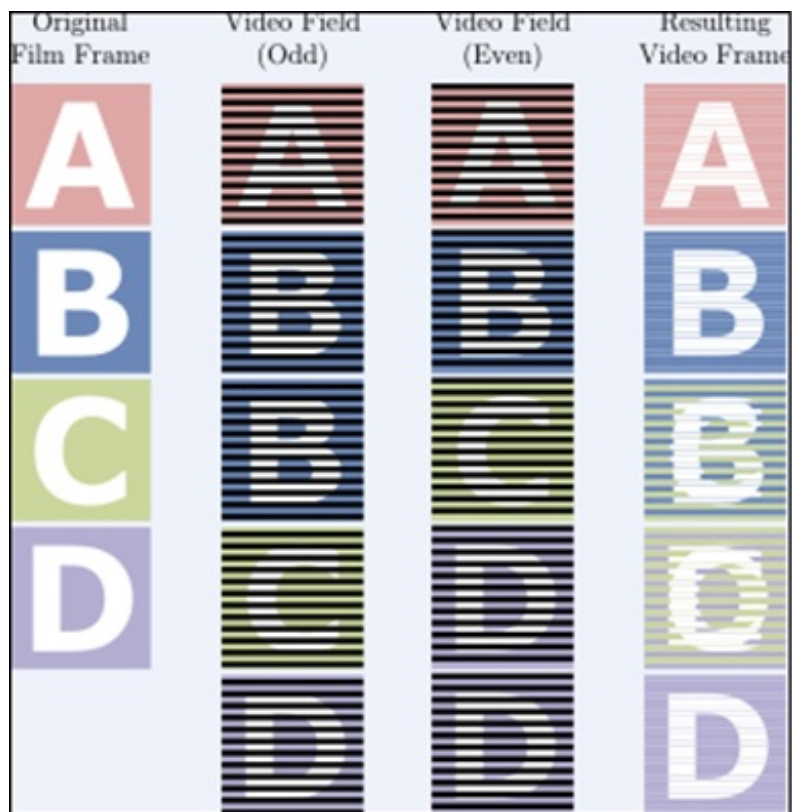
- 光流：像素的运动矢量



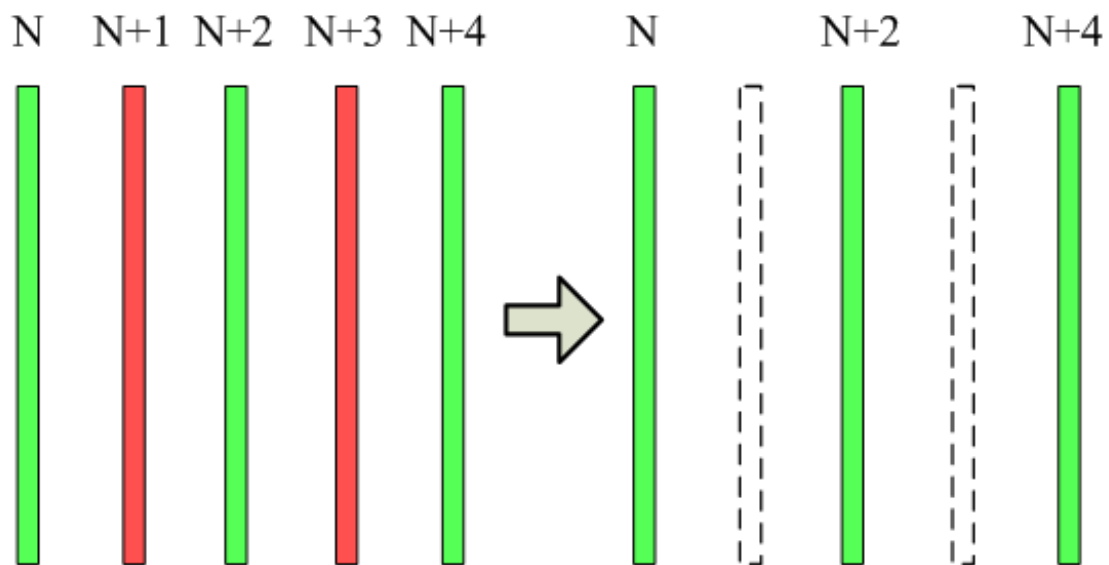


6. 帧率转换

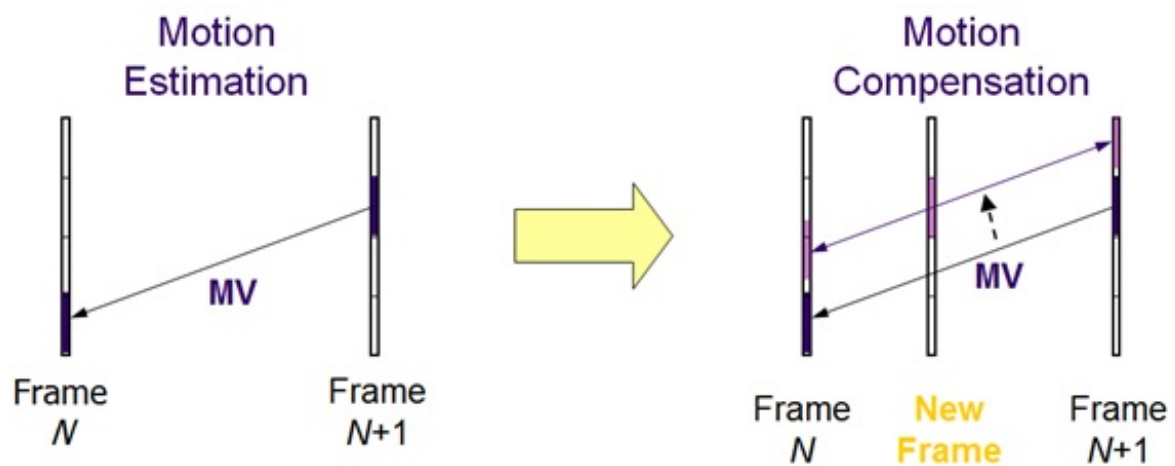
- 3 : 2下拉
 - 24帧/秒电影转换成60场/秒的电视视频
 - $24/30=4/5$



- 帧率下采样 (跳帧)



- 帧率上采样

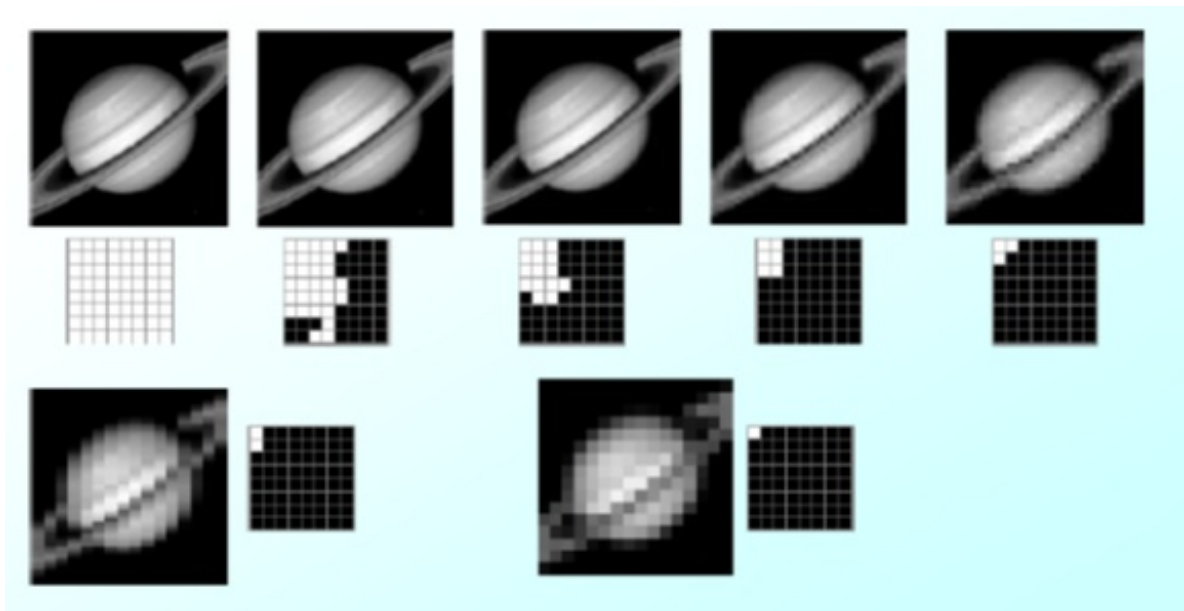


7. 后处理

- 后处理的目的是为了去除解码图像的各种效应，最大程度恢复损失的图像数据，提高解码图像的主观质量
 - 量化方法使变换系数信息损失导致解码图像产生失真和块边界的不连续性
 - 块效应 (Block Artifact) ，振铃效应 (Ringing Artifact)
 - 失真 (Distortion)
 - 视频码流传输及解码导致的解码图像质量问题
 - 跳帧 (Frame Skip)
 - 数据丢失 (Data Loss)

8. 块效应

- 块效应产生原因
 - 由于块被独立的处理

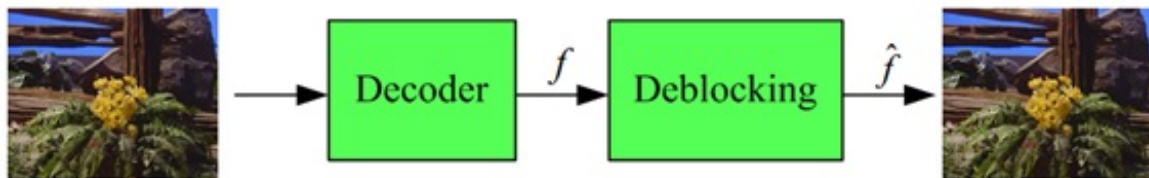


9. 去除块效应滤波

- 去除块效应滤波

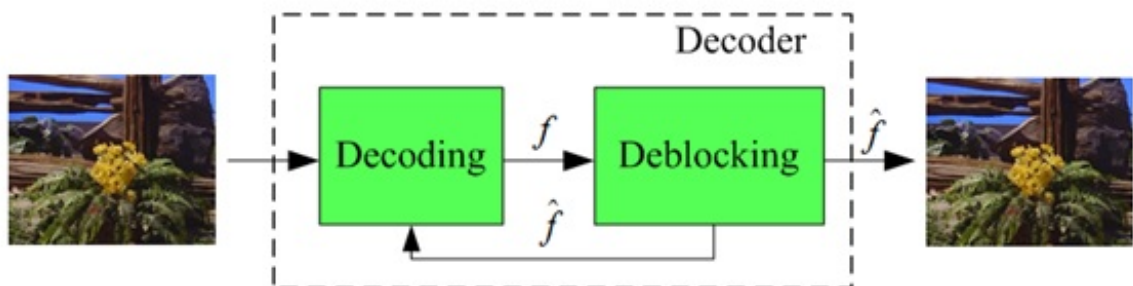
- 环外滤波（Out loop filter）：

- 滤波器在解码环路外，对显示缓冲区的图像进行滤波
 - 非标准部分，只有解码器需要环外滤波
 - MPEG-2，MPEG-4 SP



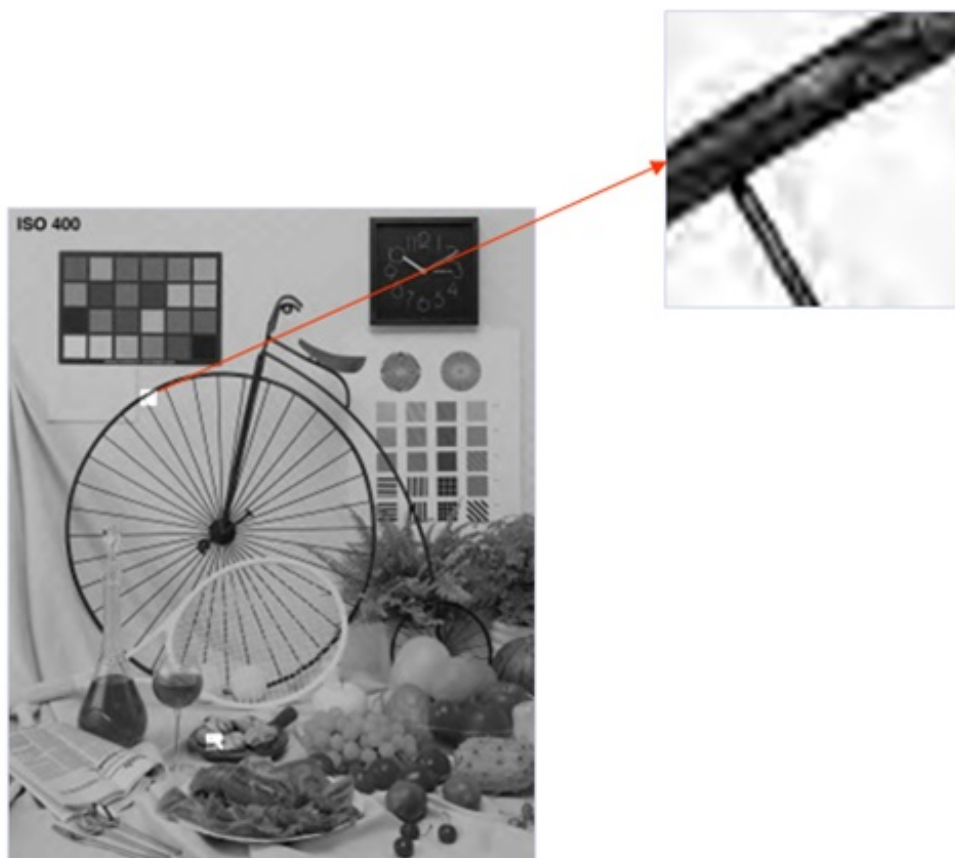
- 环路滤波（In loop filter）：

- 滤波器在解码环路内，滤波后的图像可以作为参考帧用于运动估计和补偿
 - 标准的一部分，编解码器都需要环路滤波
 - AVS，H.264/AVC



10. 振铃效应

- 振铃效应产生原因
 - 由于高频系数的量化严重



11. 数据丢失

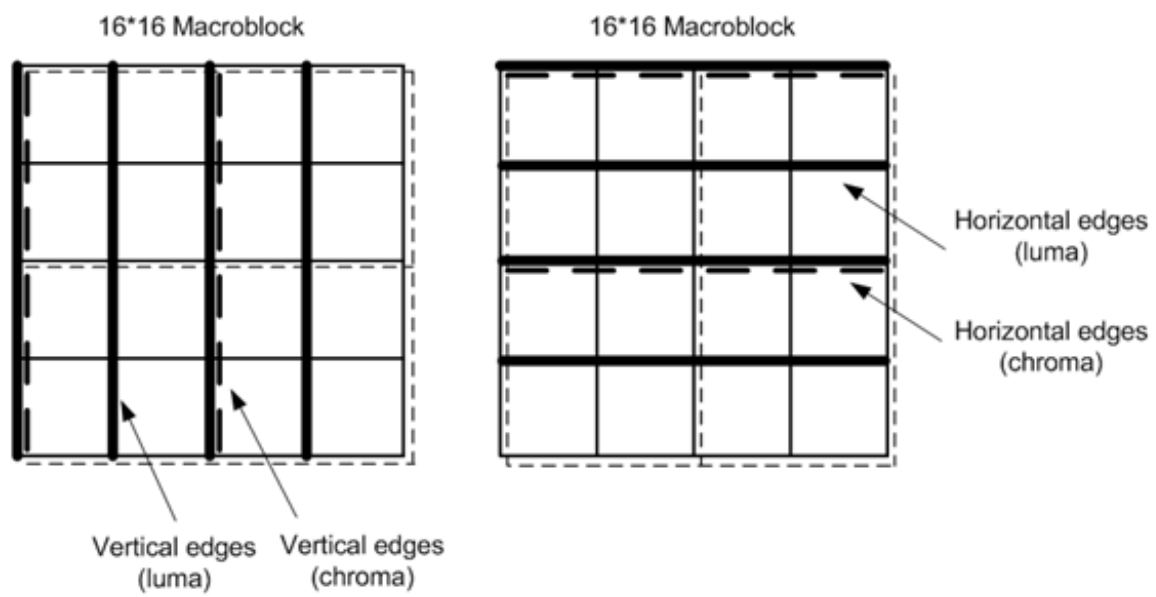
- 数据丢失



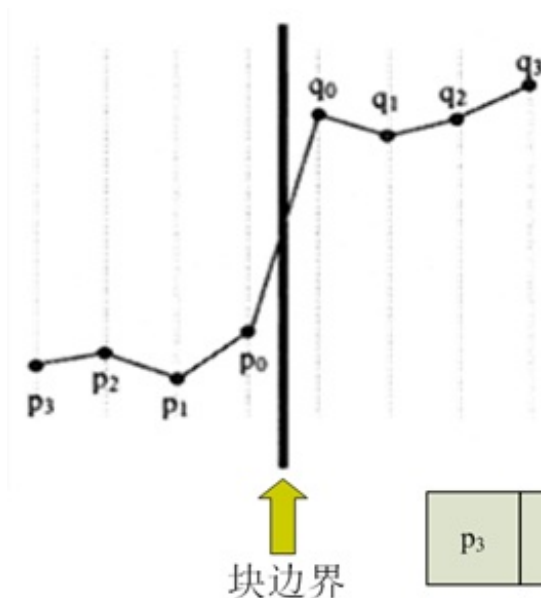
12. H.264的环路滤波



- 滤波边界



- 适应性滤波



●适应性滤波条件

●是否滤波

- 边界
- 块效应

●滤波程度

- 强滤波
- 弱滤波

p ₃	p ₂	p ₁	p ₀	q ₀	q ₁	q ₂	q ₃
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

●滤波强度

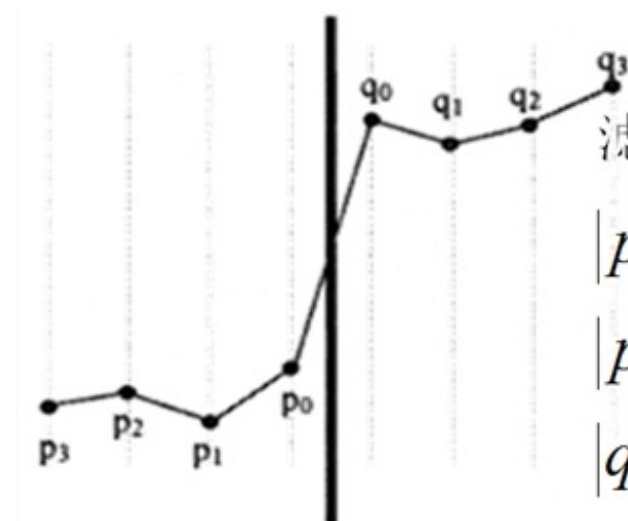
条件	Bs
P块或者Q块为帧内预测模式并且滤波边界是MB边界	4
P块或者Q块为帧内预测模式	3
P块或者Q块编码残差	2
P块和Q块的运动矢量差 ≥ 1	1
P块和Q块的参考帧不同	1
其它	0

●对相邻块P块和Q块的块边界进行滤波

●滤波强度Bs由P块和Q块的编码模式决定

●量化比运动补偿导致更多的块效应，因此用更强的滤波器滤波有变换量化系数的块边界

●滤波决策阈值



滤波条件

$$|p_0 - q_0| < \alpha(\text{Index}_A)$$

$$|p_1 - p_0| < \beta(\text{Index}_B)$$

$$|q_1 - q_0| < \beta(\text{Index}_B)$$

$$|p_0 - q_0| < \alpha(\text{Index}_A)$$

$$|p_1 - p_0| < \beta(\text{Index}_B)$$

$$|q_1 - q_0| < \beta(\text{Index}_B)$$

编码之前调整该参数，并传给解码器

$$QP = (QP_p + QP_q + 1) \ggg 1$$

$$\text{Index}_A = \min(\max(0, QP + \text{Offset}_A), 51)$$

$$\text{Index}_B = \min(\max(0, QP + \text{Offset}_B), 51)$$

$$\alpha(x) = 0.8(2^{x/6} - 1)$$

$$\beta(x) = 0.5x - 7$$

- 滤波强度

- 滤波强度和量化参数QP有关

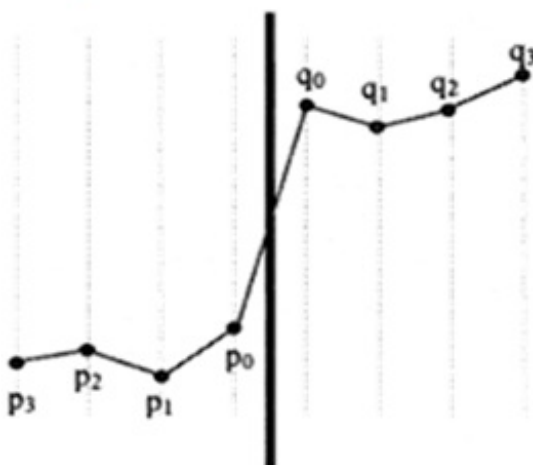
- QP越高，量化程度越大，会产生更多的块效应，需要更强的滤波
 - QP越低，量化程度越小，会产生更少的块效应，用弱一些的滤波来防止边界变得模糊

- 根据滤波强度 (Bs) 的不同有两种滤波器

- $1 \leq Bs \leq 3$ ，基本滤波操作
 - $Bs = 4$ ，强力滤波操作

- 边界滤波计算

- $1 \leq Bs \leq 3$



$$p'_0 = p_0 + \text{clip}(-c, c, \Delta_0)$$

$$q'_0 = q_0 - \text{clip}(-c, c, \Delta_0)$$

$$\Delta_0 = \left(\begin{array}{l} 4(q_0 - p_0) \\ + (p_1 - q_1) + 4 \end{array} \right) \ggg 3$$

- 截断操作参数C

	Index _A												
	1		31	32	33	34	35	36	37	38	39		51
Bs = 1	0	...	1	1	2	2	2	2	3	3	3	...	13
Bs = 2	0		2	2	2	2	3	3	3	4	4		17
Bs = 3	0		3	3	3	4	4	4	5	6	6		25

- C的值由表中的值得到
- 如果QP增加，则C更小，滤波更剧烈

- Bs=4
- 块效应更明显
- 需要更剧烈的滤波，去除块效应
- 当满足以下条件对边界进行强滤波
- 强滤波时修改 p_0, p_1, p_2 的像素值
- 弱滤波时只修改 p_0 的像素值

$$|p_0 - q_0| < (\alpha(\text{Index}_A) \gg 2) + 2$$

$$|p_2 - q_0| < \beta(\text{Index}_B)$$

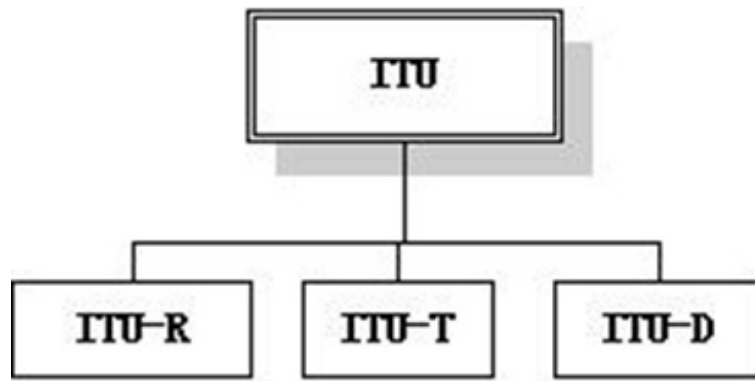
第11章 视频编码标准

1. 国际标准组织

- ITU (International Telecommunications Union) : 国际电信联盟
- ISO (International Standardization Organization) : 国际标准化组织
- IEC (International Electrotechnical Commission) : 国际电工委员会

2. ITU

- ITU-T (Telecommunication) : 电信标准化部门
- ITU-R (Radiocommunication) : 无线电通信部门
- ITU-D (Development) : 电信发展部门
- VCEG (Video Coding Experts Group) : 视频编码专家组，属于ITU-T的Study Group 16/Question 6 (ITU-T SG16/Q.6)



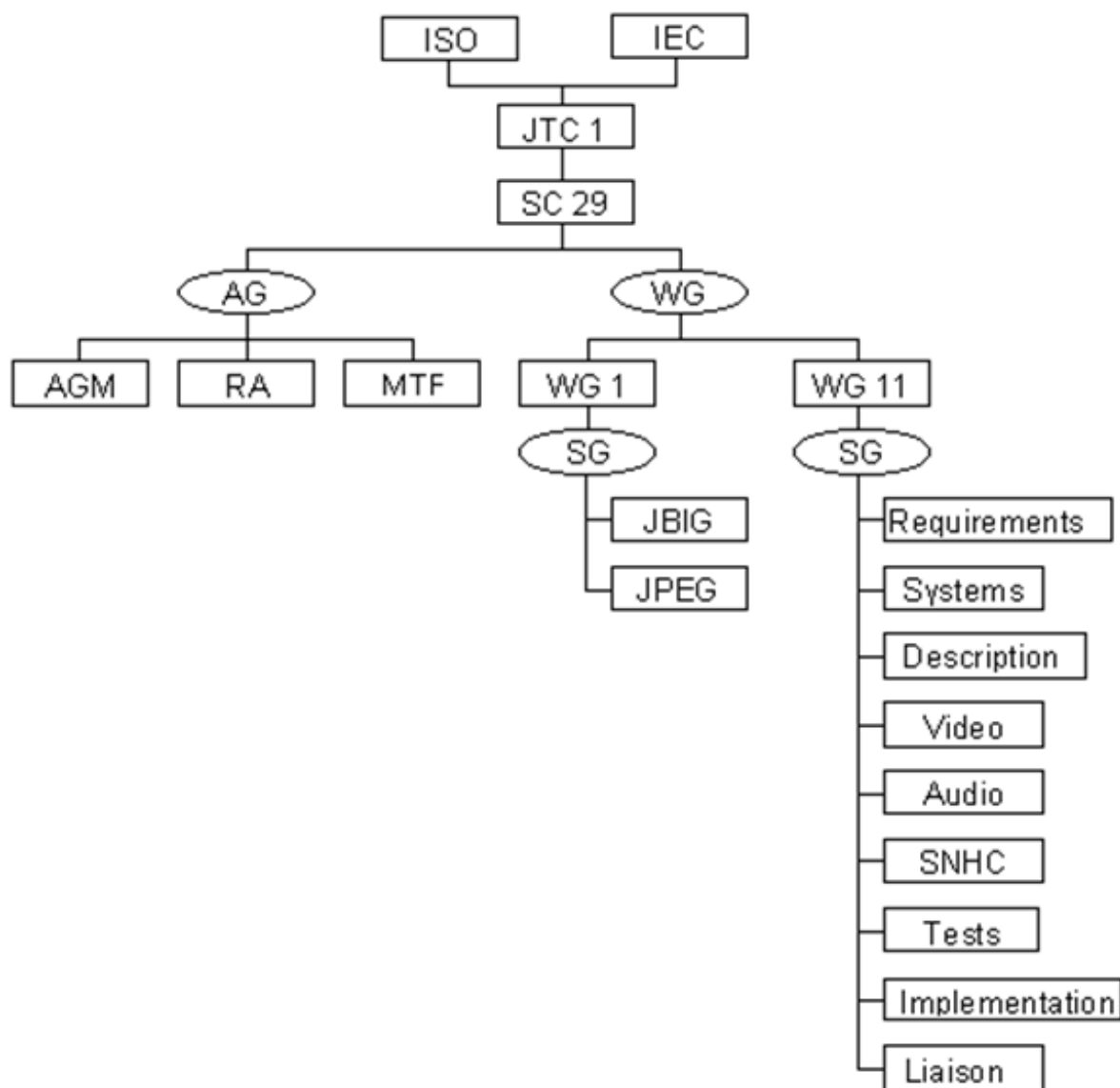
3. VCEG

- VCEG制订标准：
 - H.261：第一个视频编码标准，用于ISDN（px64Kbps）传输的视频会议，主要采用了基于整像素的运动补偿
 - H.262 (MPE-2)：与MPEG合作制定的标准
 - H.263 (H.263+， H.263++)：用于低码率的视频编码，主要用于PSDN低于54Kbps的视频会议和视频电话，采用了1/2像素运动补偿，考虑了数据损失和错误鲁棒性的需要
 - H.26L：H.264的前身标准
 - H.264 (MPEG-4 AVC)：目前编码效率最高的标准

4. ISO/IEC

- JPEG：ISO/IEC JTC1/SC29/WG1

- MPEG : ISO/IEC JTC1/SC29/WG11

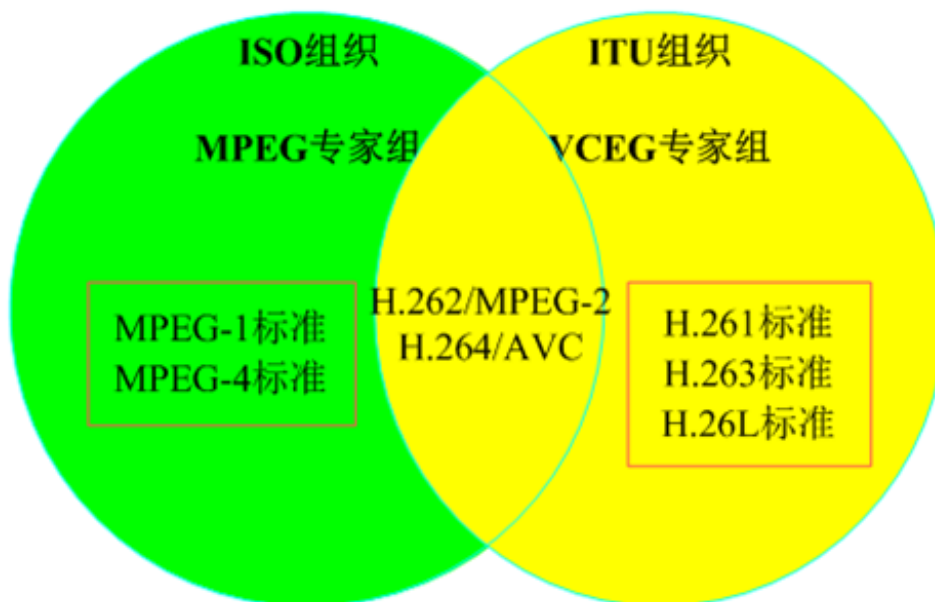


5. MPEG

- MPEG制订标准：
 - MPEG-1 : H.261+半像素运动补偿+双向预测+条带结构编码，用于VCD，LAN视频（最大1.5Mbps）
 - MPE-2 : MPEG-1+隔行视频编码+缩放功能，用于HDTV（18Mbps）、SDTV（2-5Mbps）数字信号传输和DVD（6-8Mbps）存储，兼容MPEG-1标准
 - MPEG-4 ASP (P-2) : 从H.263标准发展而来，基于视频对象平面编码，能在解码端控制视频对象，合成场景，并有音视频交互功能
 - MPEG-4 AVC (P-10) (H.264/AVC) : 与VECG合作制定的标准

6. 视频编码标准

- 国际标准组织及其制定的编码标准



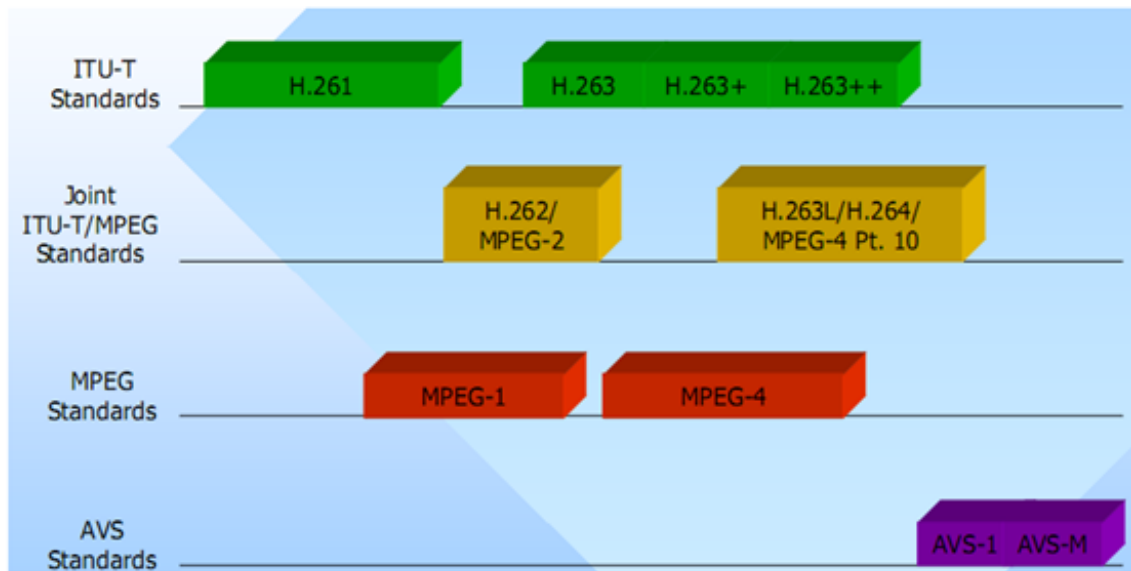
7. 标准的应用

Standard	Main Applications	Year
JPEG, JPEG2000	Image	1992-1999, 2000
JBIG	Fax	1995-2000
H.261	Video Conferencing	1990
H.262, H.262+	DTV, SDTV	1995, 2000
H.263, H.263++	Videophone	1998, 2000
MPEG-1	Video CD	1992
MPEG-2	DTV, SDTV, HDTV, DVD	1995
MPEG-4	Interactive video	2000
MPEG-7	Multimedia Content description Interface	2001
MPEG-21	Multimedia Framework	2002
H.264/MPEG-4 part 10	Advanced Video Coding	2003
	Fidelity Range Extensions (High profile), Studio editing, Post processing, Digital cinema	2004 August

8. 其他标准组织

- AVS (China Audio and Video Standard)
 - AVS1.0 : 面向高清数字电视广播
 - AVS-M : 面向移动视频
 - AVS-S : 面向监控视频
- SMPTE (The Society of Motion Picture and Television Engineers)
 - VC-1

9. 视频编码标准发展



10. MPEG-1

- 标准文档：ISO/IEC 11172-2
- 编码技术：基于H.261
 - B帧编码：B帧不能被其它帧参考
 - 16x16 ½精度运动补偿
 - 码率控制：宏块级的码率控制，调整宏块的量化参数控制码率
 - 支持视频随机访问，回放，码流编辑，音视频同步
- 参考软件：Simulation Model 3 (SM3)

11. MPEG-2

- 标准文档：ISO/IEC 13818-2
- 编码技术：基于MPEG-1
 - 帧和场运动补偿
 - 三种类型的缩放
 - 空间缩放
 - 时间缩放
 - SNR缩放
- 参考软件：Test Model 5 (TM5)
- Profile：工具集
- Level：参数限制
- 不同的应用使用不同的Profile和Level的组合

Level	Profile				
	Simple 4:2:0	Main 4:2:0	SNR Scalable 4:2:0	Spatially Scalable 4:2:0	High 4:2:0 or 4:2:2
High 1920x1152 (60 frames/s)		80 Mbit/s			100 Mbit/s for 3 layers
High-1440 1440x1152 (60 frames/s)		60 Mbit/s		60 Mbit/s for 3 layers	80 Mbit/s for 3 layers
Main 720x576 (30 frames/s)	15 Mbit/s	15 Mbit/s	15 Mbit/s for 2 layers		20 Mbit/s for 3 layers
Low 352x288 (30 frames/s)		4 Mbit/s	4 Mbit/s for 2 layers		

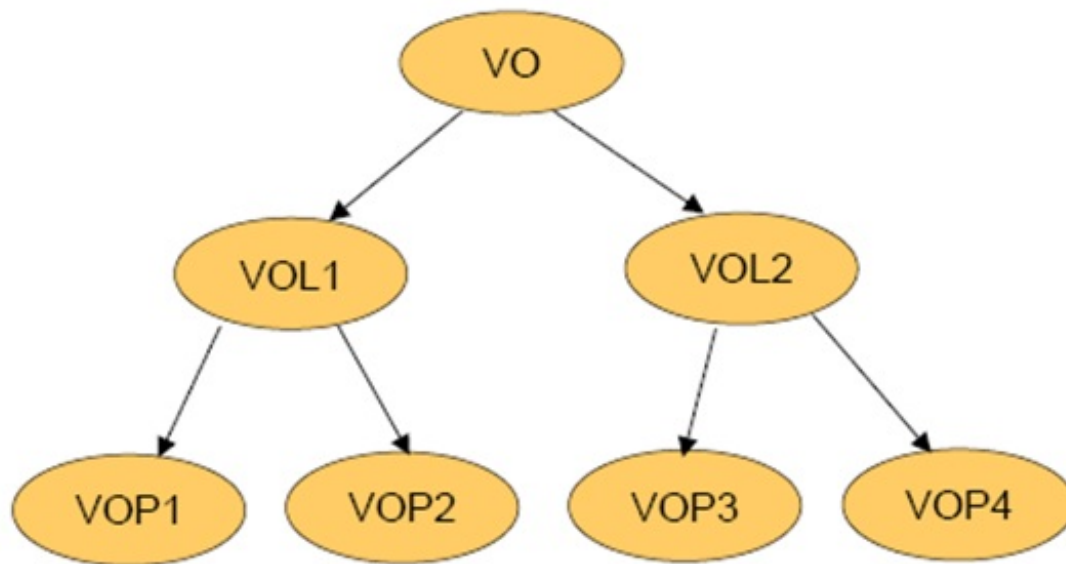
12. H.263系列

- H.263编码技术：基于H.261
 - 半像素精度运动补偿
 - 可选的数据流层次结构
 - 四种高效编码模式
 - D选项—无限制的运动矢量模式
 - E选项—基于语法的算术编码
 - F选项—高级预测模式，OBMC
 - G选项—PB帧模式
 - 支持更多的视频图像分辨率
 - 参考软件TMN2.0（Test Model Near-term 2.0）

- H.263+编码技术：基于H.263
 - 允许自定义视频图像格式
 - 增加了12个选项提高压缩效率
 - I选项—高级的帧内预测模式
 - J选项—去块效应滤波
 - K选项—条带结构模式
 - L选项—补充增强信息，用于解码器的显示相关的信息
 - M选项—增强的PB帧模式
 - N选项—参考帧选择
 - O选项—分级，时间分级，空间分级，SNR分级
 - P选项—参考帧重采样
 - Q选项—减少的分辨率更新
 - R选项—独立的片段解码
 - S选项—可替代的inter-VLC
 - T选项—改进的量化
- H.263++编码技术：基于H.263+
 - 增加了三个选项，增强抗误码能力
 - U选项—增强型参考帧选择
 - V选项—数据分区条带，变换量化系数和运动矢量数据分离
 - W选项—增加的补充增强信息，包括定点IDCT的选项，图像信息，信息类型，重复的图像头，交替场指示，稀疏参考帧识别

13. MPEG-4

- 标准文档：ISO/IEC 14496-2
- 编码技术：基于H.263
 - 增加了基于对象的编码和操作
 - 形状分割编码
 - 缩放的静态纹理小波变换
 - 网格动画编码
 - 人脸动画编码
 - 内容合成，半合成编码
 - 10比特，12比特深度采样
 - 参考软件VM7（Verification Model 7）
- 基于对象编码
 - 场景被分成多个对象
 - 每个对象用形状，运动和纹理特征描述

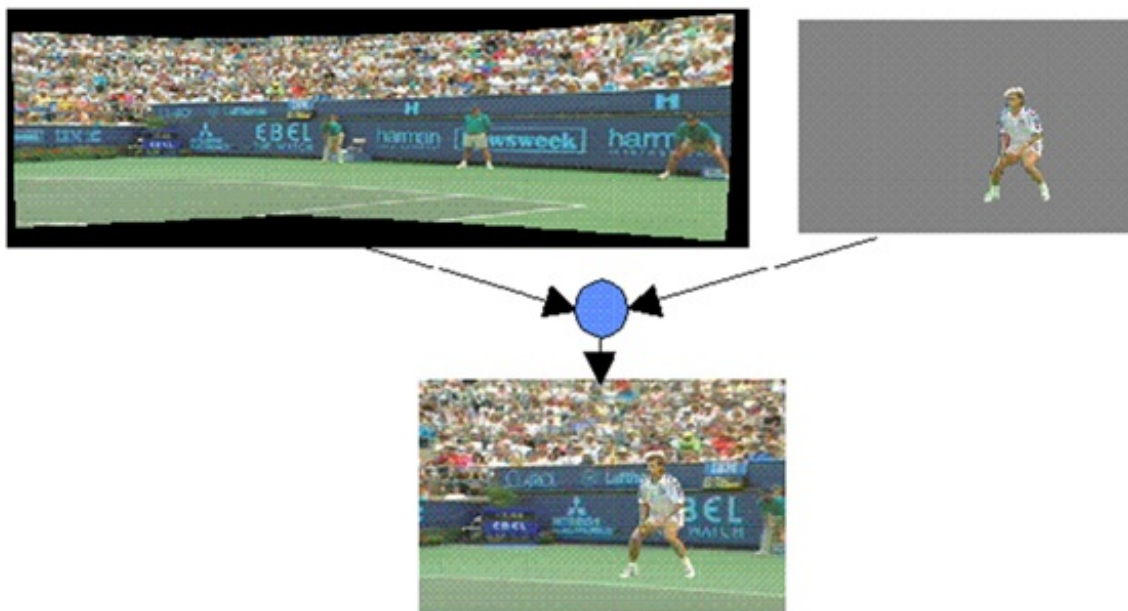


- Video Object Plane (VOP)



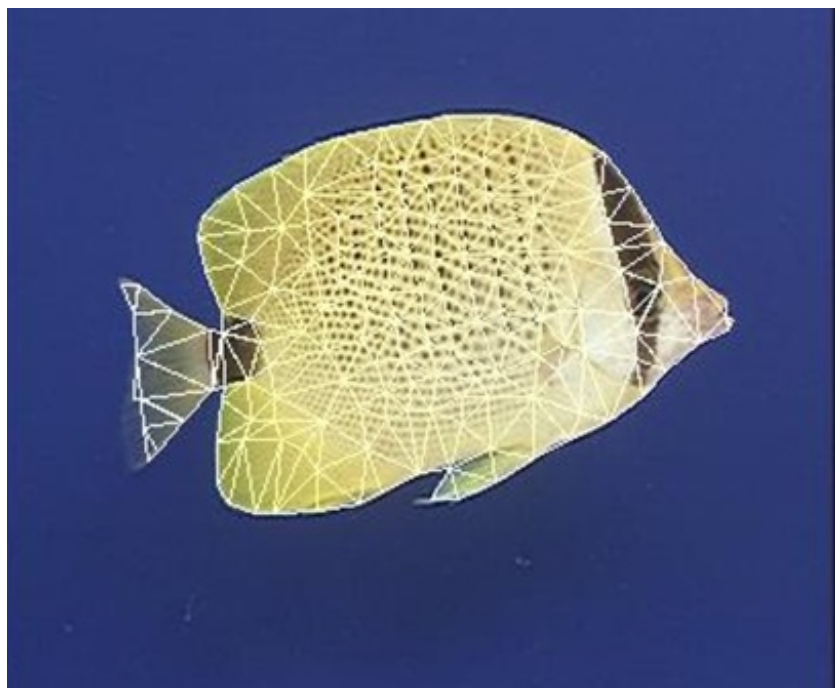
- Sprite编码

- 在序列开始编码整个视频背景和仿射映射，通过仿射映射将部分背景内容放到解码图像上显示

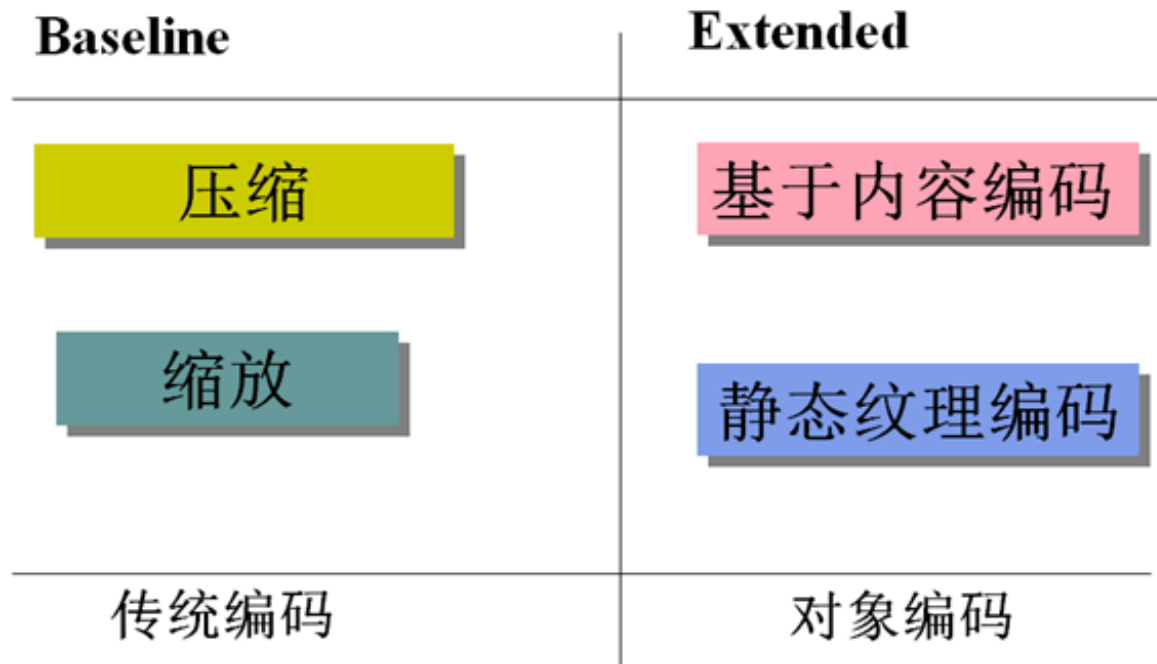


- 网格（Mesh）动画编码

- 对象可由初始的网格和节点的运动矢量描述



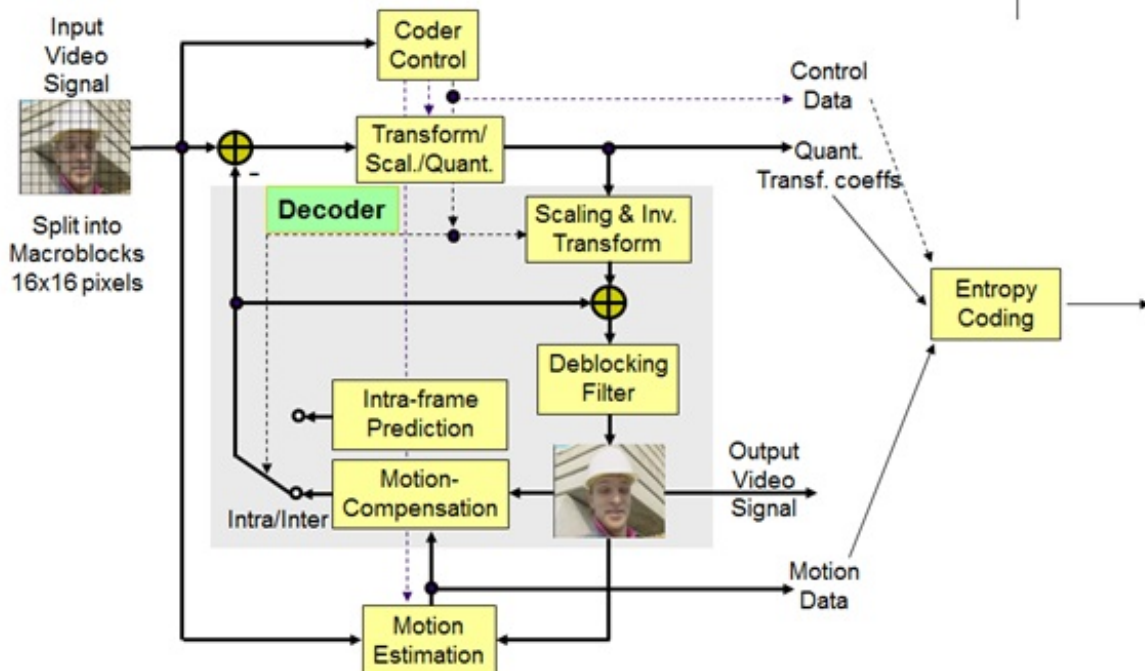
- Profile



14. H.264/AVC

- 标准文档：ISO/IEC 14496-10 (ITU-T H.264)
- 前身为：H.26L(Long Range)，基于MPEG-2/4
- 参考软件JM (Joint Model)

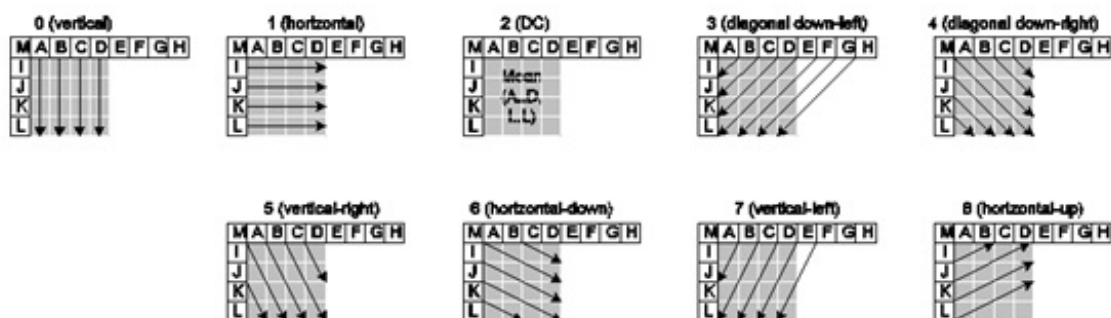
15. H.264编码结构



16. H.264-Intra编码

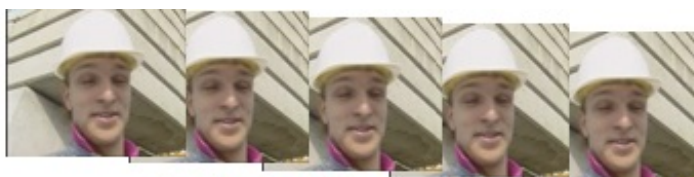
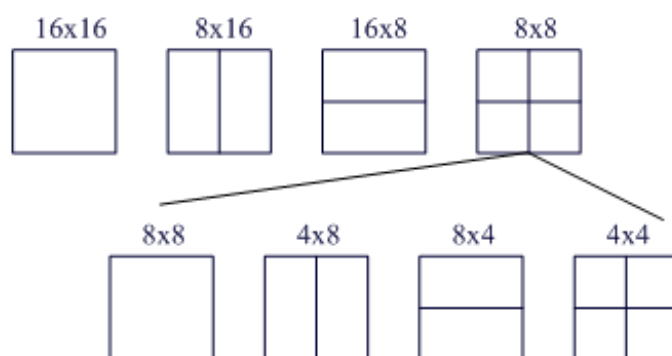
- 4个亮度16x16的模式
- 9个亮度4x4的模式
- 4个色度8x8的模式
- 编码复杂度高

- 不利于错误恢复



17. H.264-运动补偿

- 可变块运动补偿
- 1/4像素精度亮度运动补偿
 - 1/2像素滤波器 (1, -5, 20, 20, -5, 1)
- 1/8像素精度色度运动补偿
 - 色度采用双线性滤波
- 图像外运动矢量
- 多参考帧
 - 每个宏块的四个8x8子块都可以使用不同的参考帧



多参考帧运动补偿

18. H.264多参考帧

- P帧
 - 只参考一个list的参考帧，List0
- B帧
 - 参考两个list的参考帧，List0和List1

- 每个List最大可包含16个参考帧
 - List0：以前的参考帧缓冲
 - List1：将来的参考帧缓冲

19. H.274-加权预测

- P帧

$$\hat{P}(x, y) = \alpha P_{l_0}(x, y) + \beta$$

- B帧

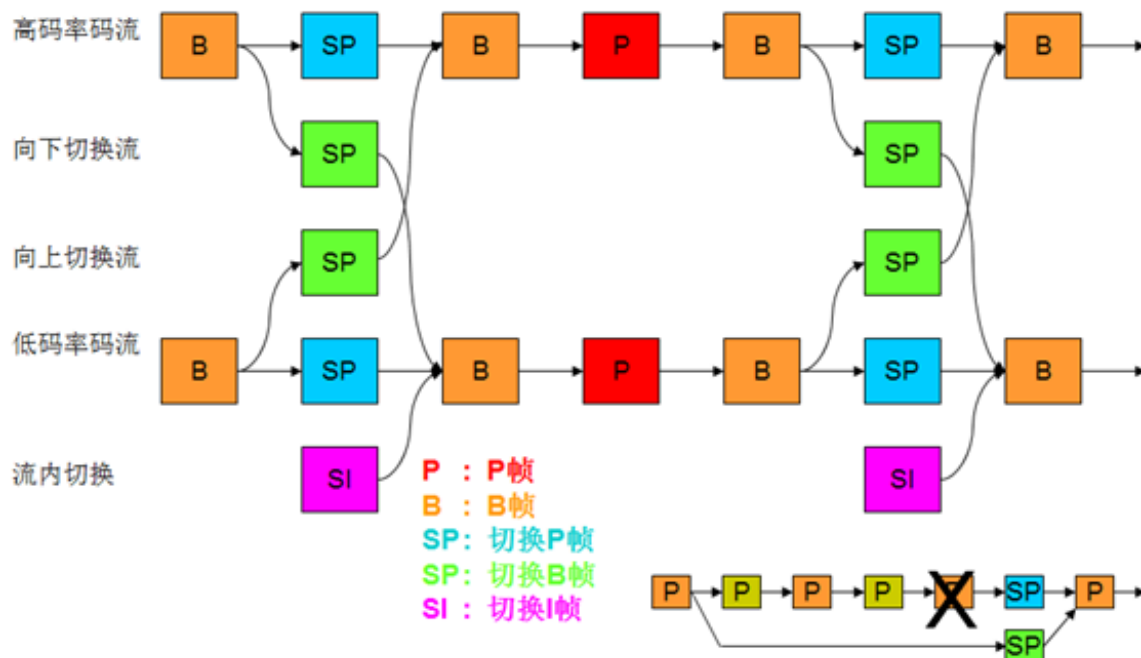
$$\hat{P}(x, y) = \alpha_0 P_{l_0}(x, y) + \alpha_1 P_{l_1}(x, y) + \beta$$

- 对消逝和过度的视频场景编码效果好



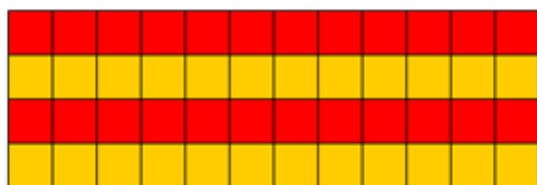
20. H.264-切换图像

- SP/SI帧实现
 - 码流切换
 - 随机切入
 - 错误恢复
 - 也可以实现快进，快退
- 特点
 - 对于不同的参考图像可以得到完全相同的重构图像
 - SP帧的编码效率低于P帧但高于I帧
 - SI帧的编码效率低于I帧



21. H.264-FMO

- 灵活宏块顺序 (Flexible MB Order, FMO)
 - 宏块有不同的组合构成条带
 - 七种不同的宏块组合方式
- FMO增强了解码的错误恢复能力



两个隔行的条带划分



四个散布的条带划分

22. H.264-数据分区

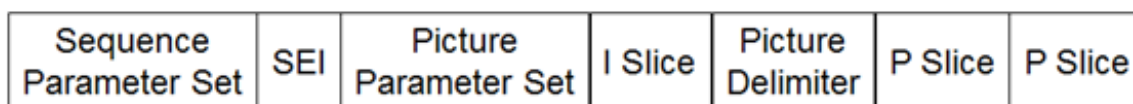
- 上层的语法元素的丢失会严重影响压缩码流的解码
- 将码流按照重要程度分成三部分
 - A : 头信息 : 包括宏块类型, 运动矢量等
 - B : Intra系数信息
 - C : Inter系数信息
- 用不同的优先级传输不同重要程度的分区来实现错误恢复

23. H.264-ASO, RP

- 任意条带顺序（Arbitrary Slice Order，ASO）
 - 条带之间独立解码
 - 条带可以按任意顺序传输
 - 丢失的条带可以被解码的条带恢复出来
- 冗余图像（Redundant Picture）
 - 在码流里放入冗余的图像，如果图像数据丢失了，可以用冗余图像代替

24. H.264-视频编码层

- 网络独立的编码层
- VCL层包含：
 - 序列参数集：Sequence Parameter Set（SPS），存放序列头信息
 - 图像参数集：Picture Parameter Set（PPS），存放图像头信息
 - 条带以下数据
 - 额外的增强信息（Supplemental Enhancement Information，SEI）



25. H.264-网络提取层

- 提取的VCL数据
- 适合于传输层的数据传输和媒介的存储
- NAL单元为数据包和码流系统定义了一种通用的传输格式



- NAL单元连接在一起构成了NAL流

- 面向码流传输的NAL单元



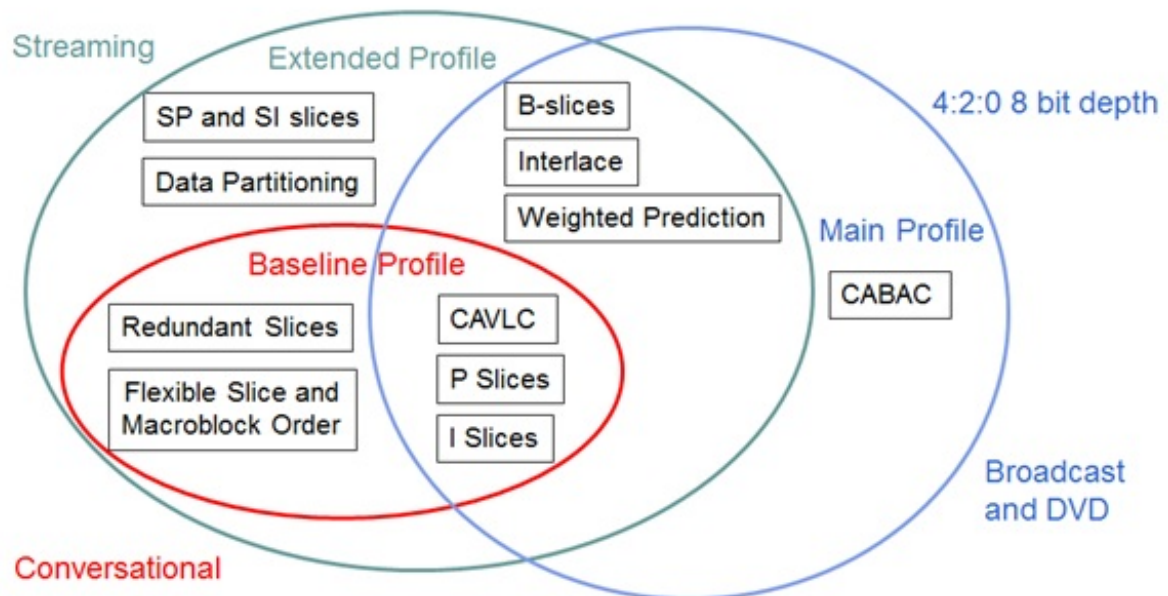
- 面向数据包传输的NAL单元



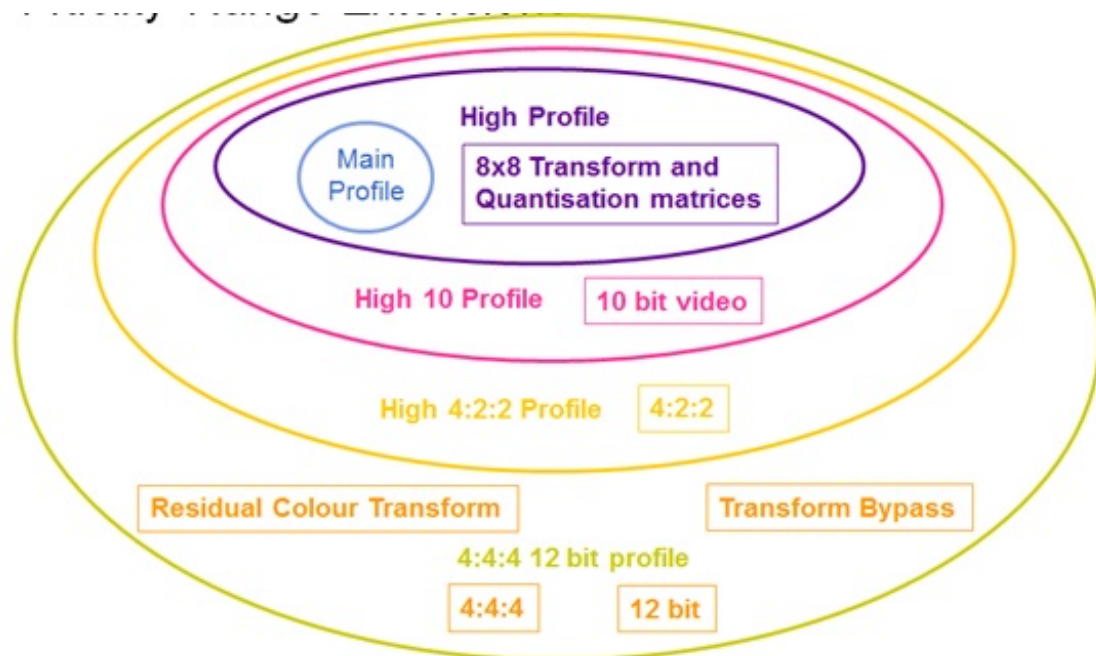
26. H.264-Profiles

- Profile : 码流语法子集

Profile	Applications
Baseline	Video Conferencing Videophone
Main	Digital Storage Media Television Broadcasting
Extended	Streaming Video
High	Content contribution Content distribution Studio editing Post processing



- Fidelity Range Extensions



27. H.264-Levels

- Level : 参数限制

Level 1	15Hz QCIF at 64Kbit/s
Level 1b	15Hz QCIF at 128Kbit/s
Level 1.1	30Hz QCIF at 192Kbit/s
Level 1.2	15Hz CIF at 384Kbit/s
Level 1.3	30Hz QCIF at 768Kbit/s
Level 2	30Hz QCIF at 2Mbit/s
Level 2.1	25Hz 625HHR at 4Mbit/s
Level 2.2	12.55Hz 625SD at 4Mbit/s

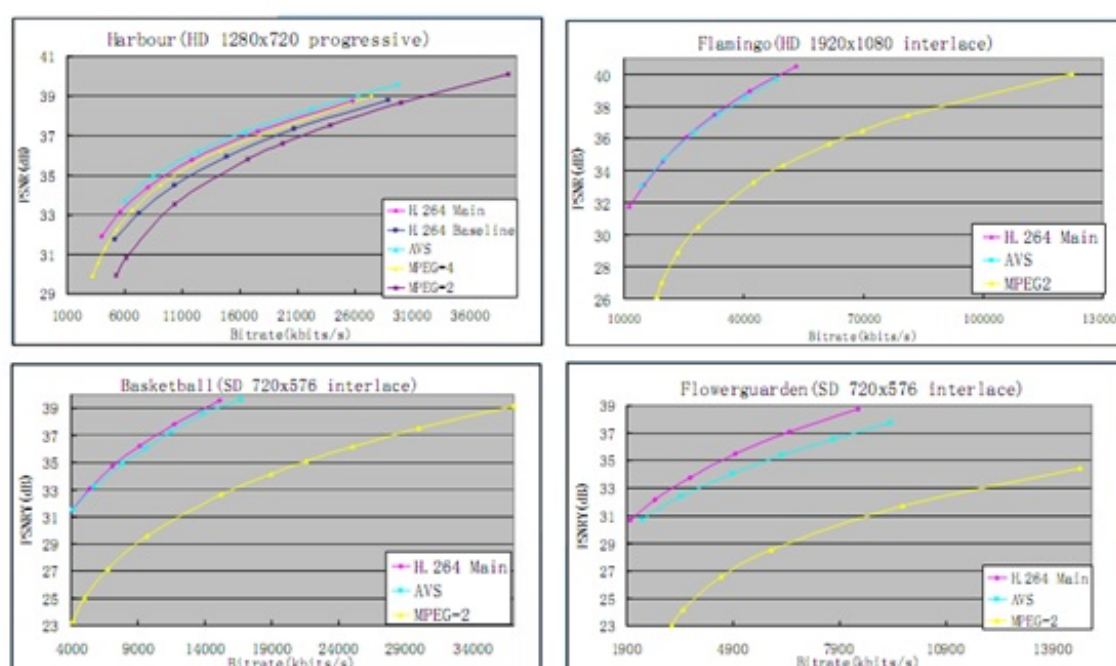
Level 3	25Hz 625SD at 10Mbit/s
Level 3.1	30Hz 720p at 14Mbit/s
Level 3.2	60Hz 720p at 20Mbit/s
Level 4	30Hz 1080 at 20Mbit/s
Level 4.1	30Hz 1080 at 50Mbit/s
Level 4.2	60Hz 1080 at 50Mbit/s
Level 5	30Hz 16VGA at 135Mbit/s
Level 5.1	30Hz 4Kx2K at 240Mbit/s

28. AVS vs. H.264

Comparison	AVS Part-2	H.264/AVC
Interpolation	4-tap filters for 1/2- and 1/4-pixel	6-tap filter for 1/2-pixel 2-tap filter for 1/4-pixel
MC minimum block size	8x8	4x4
Transform	8x8 Pre-scaled ICT	4x4 and 8x8 ICT
Quantization	64 QPs nearly 8-order exponential periodic	52 QPs 6-order exponential periodic
Perceptual-based quant.	No	Yes
Adaptive scan	Yes	No
Entropy coding	CA-2D-VLC and CABAC	CAVLC and CABAC
Deblocking filter	8x8-based, 3 BS levels	4x4-based, 4 BS levels
Error robustness	Weak	Strong
Supporting lossless coding	No	Yes

Comparison	AVS Part-2	H.264/AVC
Number of profiles	2	4
Picture type	I, P and B	I, P, B, SI and SP
Interlace handling	PAFF and MBAFF	PAFF and MBAFF
Color sampling format	4:2:0 and 4:2:2	4:2:0, 4:2:2 and 4:4:4
Bits per sample	8	up to 12
Block size for intra prediction	8x8 for luma and chroma	4x4, 8x8 and 16x16 for luma 8x8 for chroma
Intra prediction mode	5 modes for luma 4 modes for chroma	13 modes for luma 4 modes for chroma
Number of reference frame	at most 2	depends on the level, can be much more than 2
Bi-directional prediction	at most 1 MV transmitted	at most 2 MVs transmitted
MV prediction	Geometrical median prediction	Median prediction

29. 性能比较



30. AVS会员

