

11月26日：第五章 多媒体信息分析与处理
12月3日：第六章 实时多媒体通信



多媒体技术基础

第三章 多媒体数据压缩

§ 3.4 视频数据的压缩标准



2024年11月18日



实验报告：12月8日前BB系统提交

电子工程与信息科学系



第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - **MPEG简介**
 - MPEG-1、MPEG-2视频压缩原理
 - MPEG-4视听对象编码简介
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介





MPEG是什么

- ◆ MPEG=Moving Picture Expert Group
- ◆ 动态图像专家组
 - 1988年5月国际标准化组织(ISO)和国际电工委员会(IEC)联合成立ISO/IEC JTC 1(**Joint Technical Committee**)
 - 正式名称： ISO/IEC JTC 1/SC 29/ WG 11
- ◆ 专家组的任务
 - 负责开发视频和声音的编解码和传输标准(MPEG标准)
- ◆ MPEG标准
 - 上世纪90年代，MPEG-1和MPEG-2标准获得的巨大成功
 - 随着因特网的迅速发展，2000年前后相继发布了MPEG-4、MPEG-7和MPEG-21标准
 - 其后又提议开发包括MPEG-H、MPEG-I在内的多种标准

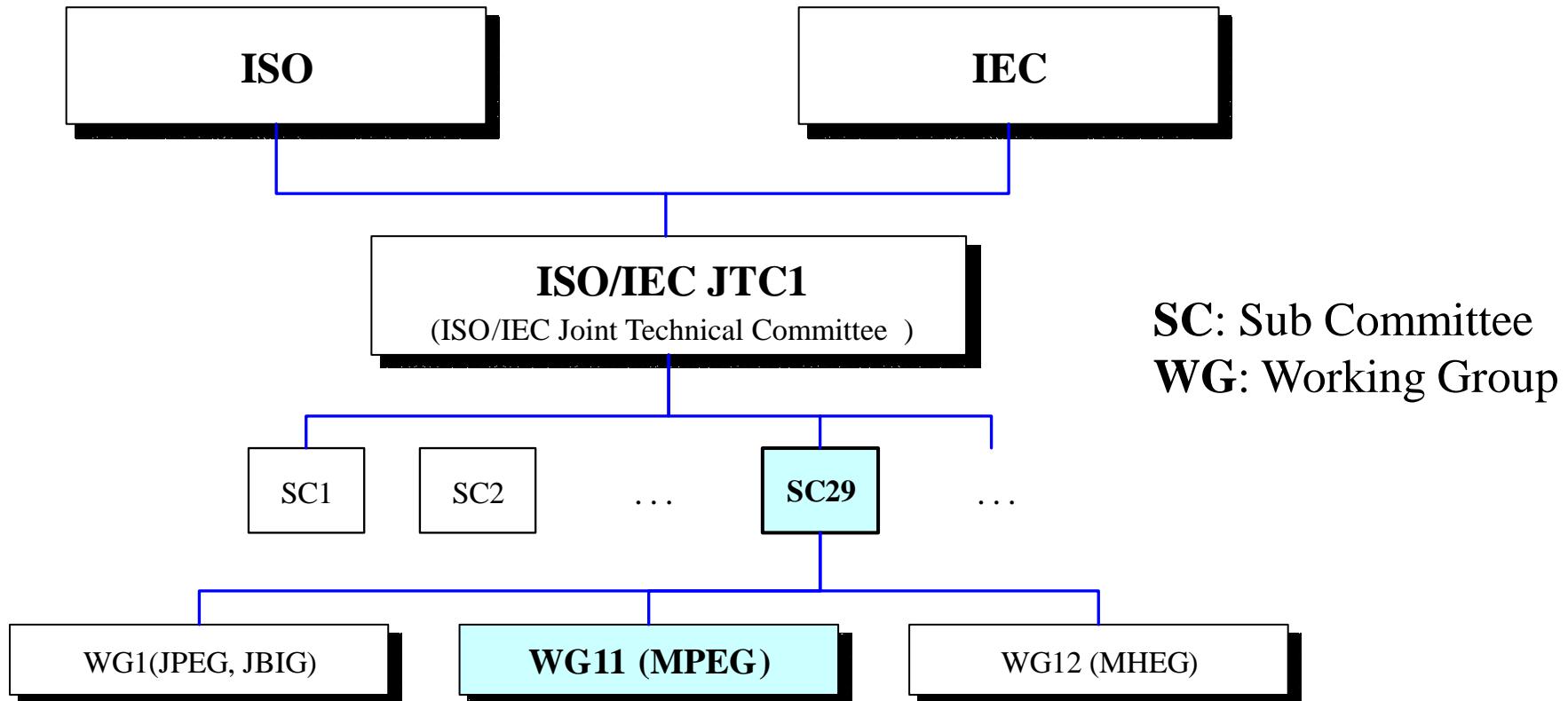




ISO/IEC联合成立的MPEG专家组

ISO: International Organization for Standardization

IEC: International Electrotechnical Commission



SC: Sub Committee

WG: Working Group





MPEG标准(2016年)

MPEG	标准编号	版本	标题
<u>MPEG-1</u>	ISO/IEC 11172	1992	Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mb/s (用于数据速率高达大约 1.5 Mb/s 的数字存储媒体的视像和伴音编码技术)
<u>MPEG-2</u>	ISO/IEC 13818	1994	Generic coding of moving pictures and associated audio information (动态图像及其伴音信息的通用编码)
<u>MPEG-3</u>	-	-	合并到MPEG-2
<u>MPEG-4</u>	ISO/IEC 14496	1999	Coding of audio-visual objects (视听对象编码)
<u>MPEG-7</u>	ISO/IEC 15938	2002	Multimedia content description interface (多媒体内容描述接口)
<u>MPEG-21</u>	ISO/IEC 21000	2001	Multimedia framework (MPEG-21) (多媒体框架)
MPEG-A	ISO/IEC 23000	2007	Multimedia application format (MPEG-A)(多媒体应用格式)
MPEG-B	ISO/IEC 23001	2006	MPEG systems technologies (多媒体系统技术)
MPEG-C	ISO/IEC 23002	2006	MPEG video technologies (MPEG视像技术)
<u>MPEG-D</u>	ISO/IEC 23003	2007	MPEG audio technologies (MPEG声音技术)
MPEG-E	ISO/IEC 23004	2007	Multimedia Middleware (多媒体中间件)
<u>MPEG-H</u>	ISO/IEC 23008	2013	High Efficiency Coding and Media Delivery in Heterogeneous Environments (高效视像编码和异构环境下的媒体传输)
MPEG-M	ISO/IEC 23006	2010	MPEG extensible middleware (MXM) (MPEG扩展中间件)
MPEG-U	ISO/IEC 23007	2010	Rich media user interfaces (富媒体用户接口)
MPEG-V	ISO/IEC 23005	2011	Media context and control (媒体前后关系和控制)
MPEG-DASH	ISO/IEC 23009	2012	Information technology – DASH (Dynamic Adaptive Streaming over HTTP) (在HTTP上的动态自适应数据流)





MPEG文档的创建过程

◆ 创建过程中的各阶段及其文档分别称为

- 工作草案(Working Draft, WD)
- 委员会草案(Committee Draft, CD)
- 最终草案(Final Committee Draft, FCD)
- 国际标准草案(Draft International Standard, DIS)
- 最终国际标准草案(FDIS)
- 国际标准(International Standard, IS)

◆ **WD → (CD → FCD) →(DIS → FDIS) → IS**





MPEG标准的重要性(待续)

- ◆ 对我们的生活和工作已经产生了巨大影响，并将继续产生深远影响，不论你是否意识到或感受到
 - VCD影视是MPEG-1标准的典型应用。在1990年代，世界上已销售无数的VCD播放机，在PC上几乎都安装了VCD播放器和CD-ROM驱动器
 - 销售的VCD影视光盘用“亿”做单位计算
 - MP3是MPEG-1标准另一个大规模应用的实例
 - MPEG-1, -2是数字电视设备的心脏，如VCD/DVD播放器和数字电视机顶盒等





MPEG标准的重要性

- ◆ MPEG-4标准推荐采用对象编码技术，以进一步减少视频数据和声音数据的数据量，这将对数字电视和多媒体通信产生深刻的影响
- ◆ MPEG-7和MPEG-21对多媒体的内容检索和多媒体数据库管理系统的发展产生深刻影响
- ◆ 21世纪，**H.264/MPEG-4 AVC**和**H.265/MPEG-H (HEVC)**的成功开发，进一步提高了数据压缩性能，在HDTV，UHDTV(4K和8K)电视广播中大显身手
- ◆ **H.266/MPEG-I (VVC)**在2020年发布...
- ◆ 标准是人类智慧的结晶，凝聚着人类认识世界和改造世界的历史经验

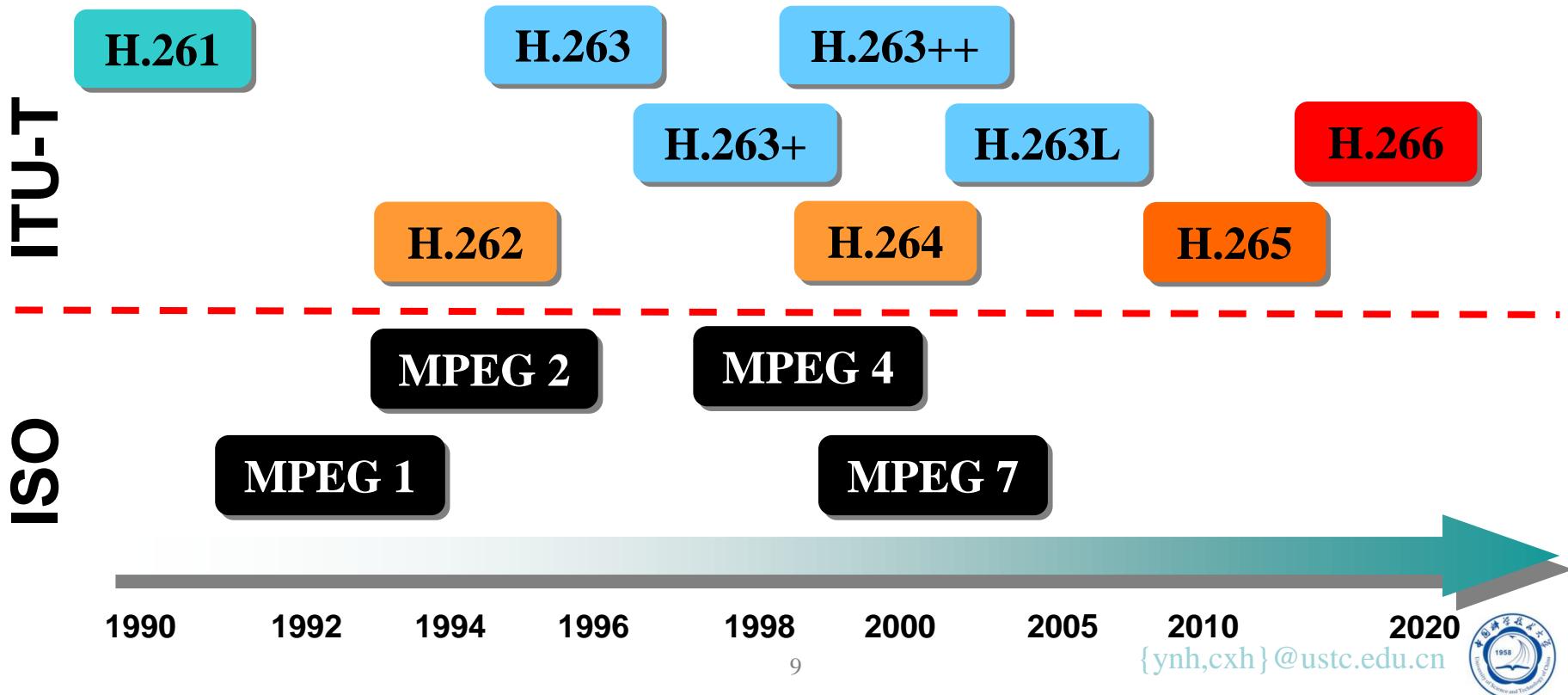




视频压缩标准的演进

◆ 两个著名的组织

- ITU-T VCEG(Video Coding Experts Group)专家组
- ISO/IEC MPEG(Moving Picture Experts Group)专家组





MPEG相关标准应用

- ◆ **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**: 2003年发布的视频压缩标准，主要是扩展MPEG-1、MPEG-2等标准以支持视频/音频对象（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-4 AVC即H.264
- ◆ **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。
- ◆ **H.264/MPEG-4 AVC**和面向高清视频，如HDTV
- ◆ **H.265/MPEG-H (HEVC)**预期在UHDTV(4K和8K)应用
- ◆ **H.266/MPEG-I (VVC)**在2020发布





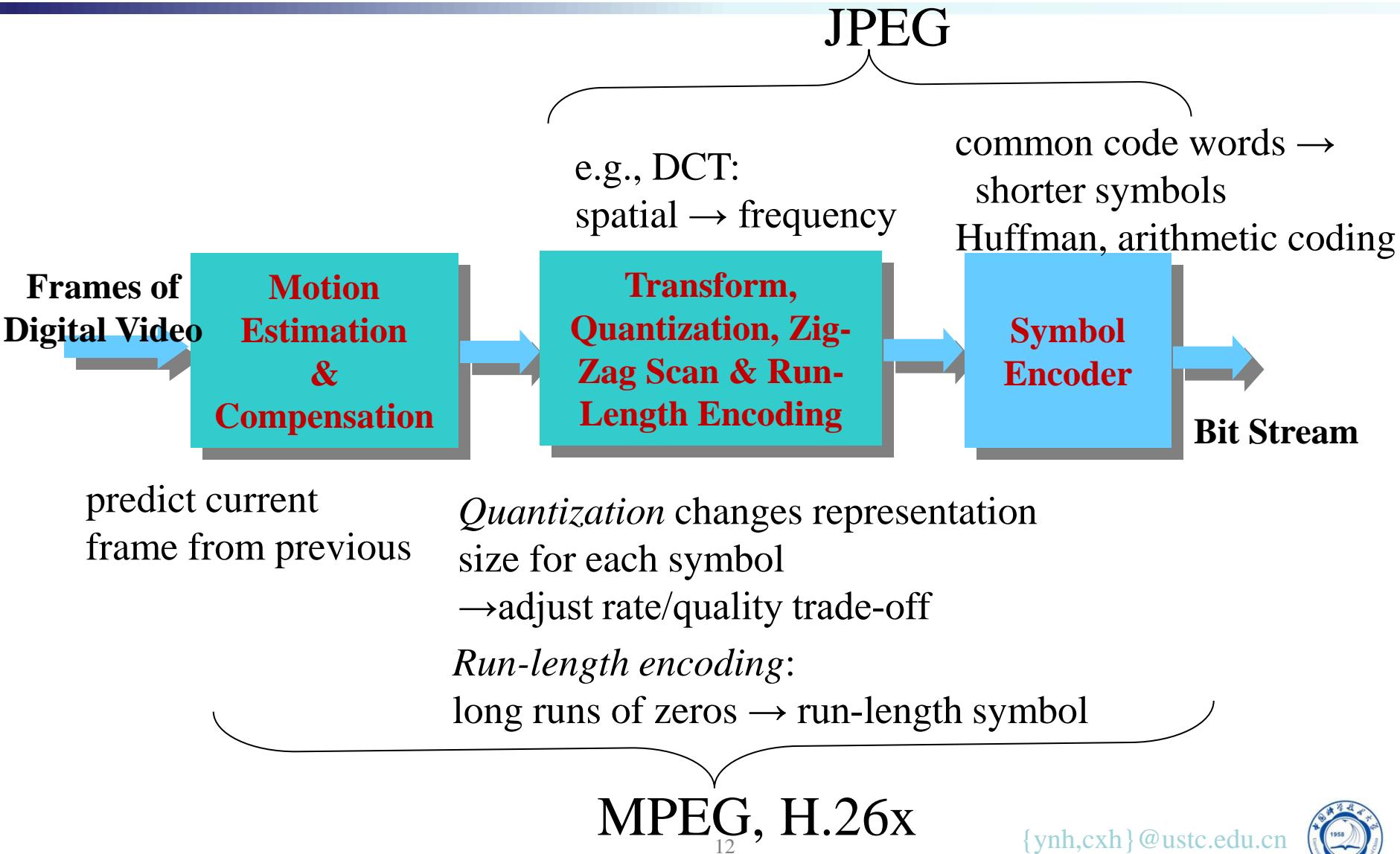
第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - MPEG简介
 - **MPEG-1、MPEG-2视频压缩原理**
 - MPEG-4视听对象编码简介
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介





视频编码器原理





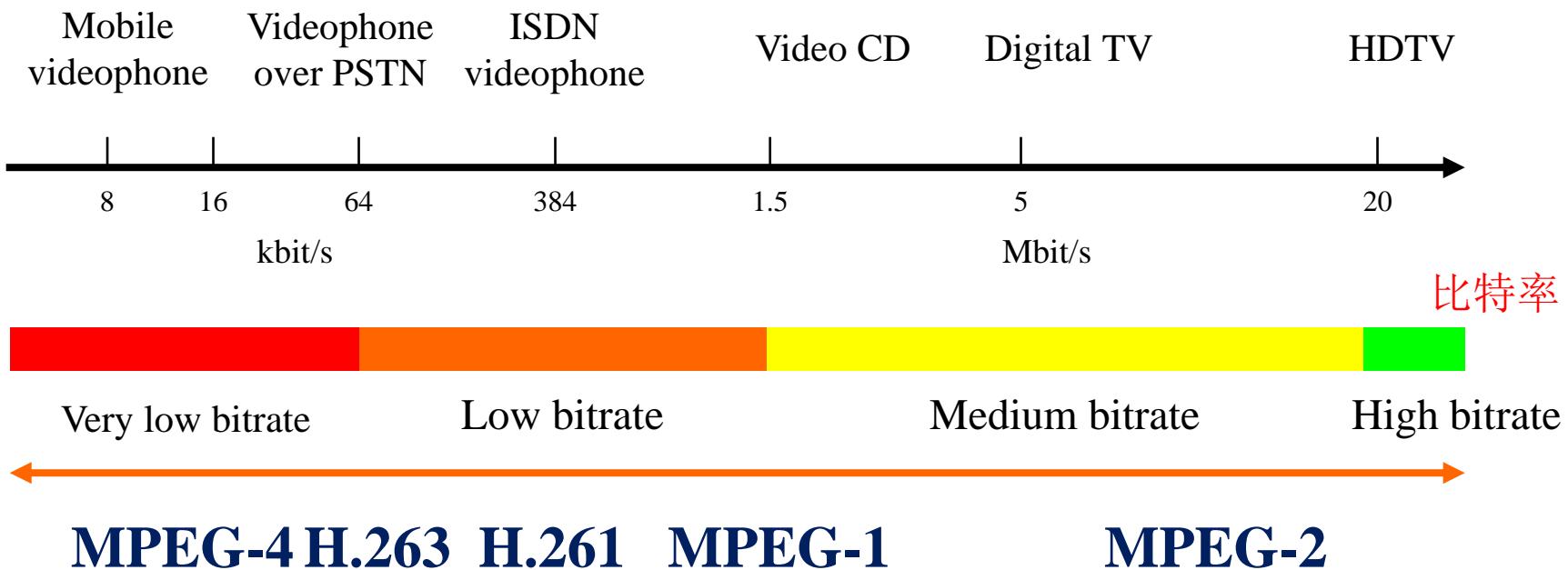
经典视频编解码器区别

- ◆ H.261 是第一个使用的数字视频压缩标准，之后标准中视频编解码器都是基于相同的设计框架。
- ◆ MPEG-1 part2 主要使用在 VCD 上。比起 H.261 标准，MPEG-1 增加了对半像素运动补偿和双向运动预测帧。
- ◆ MPEG-2 part2 等同于 H.262，比起 MPEG-1，MPEG-2 最大的改进在于增加了对隔行扫描视频的支持。
- ◆ H.263 主要用在视频会议、视频电话和网络视频上。性能上了较大的提升。尤其是在低码率端。
- ◆ MPEG-4 part2 标准比起 MPEG-2 和第一版的 H.263，它的压缩性能有所提高。引入 Object-oriented 的编码方法和一些提高压缩能力的技术，包括一些 H.263 的技术和 1/4 像素的运动补偿。
- ◆ MPEG-4 part10 和 ITU-T H.264 是相同的标准，也称作 AVC。





经典编解码器的速率





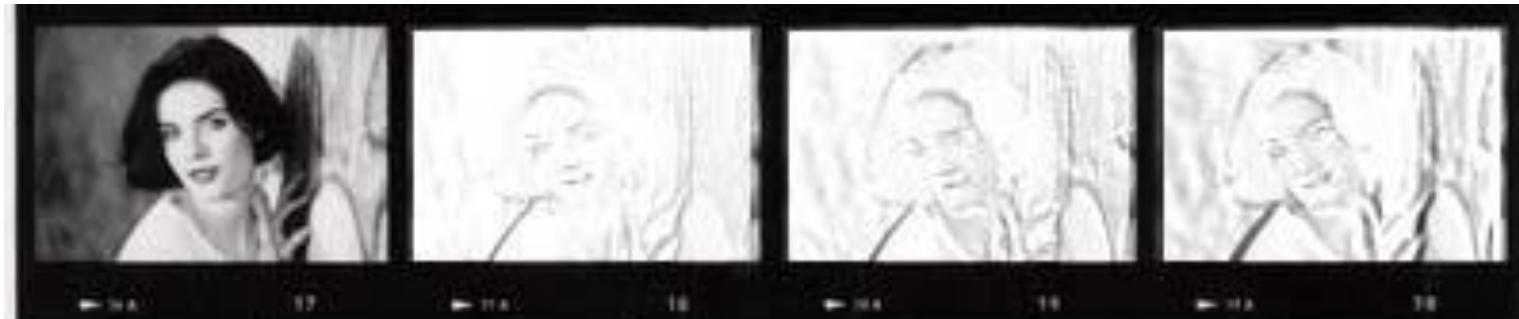
MPEG视频压缩利用各种冗余

- ◆ 空间方向用JPEG算法来去掉冗余
- ◆ 时间方向用运动补偿(motion compensation)算法来去掉冗余信息

种类	内容	目前用的主要方法
统计	空间冗余	变换编码, 预测编码
特性	时间冗余	帧间预测, 移动补偿
图像构造冗余	图像本身的构造	轮廓编码, 区域分割
知识冗余	收发两端对人物的共有认识	基于知识的编码
视觉冗余	人的视觉特性	非线性量化, 位分配
其他	不确定性因素	



帧间冗余是最大的冗余



Frame 1

Frame 2

Frame 3

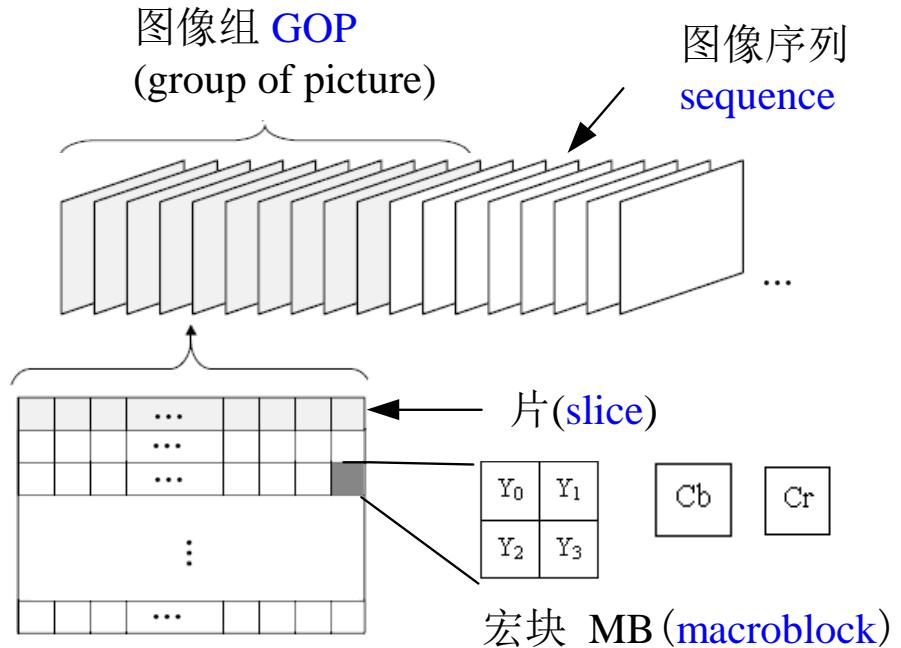
Frame 4





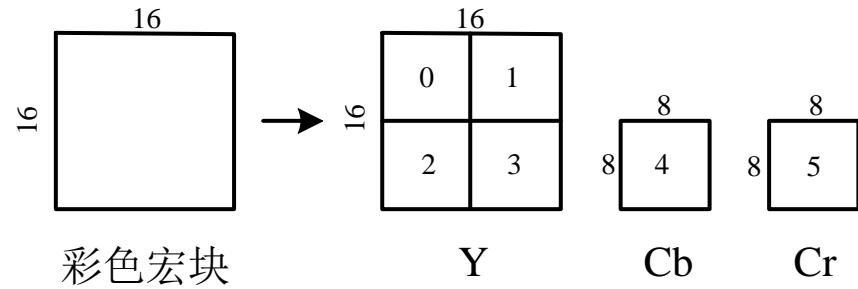
视频数据结构

GOP→picture→slice→MB



图像(picture)

视频数据的组织



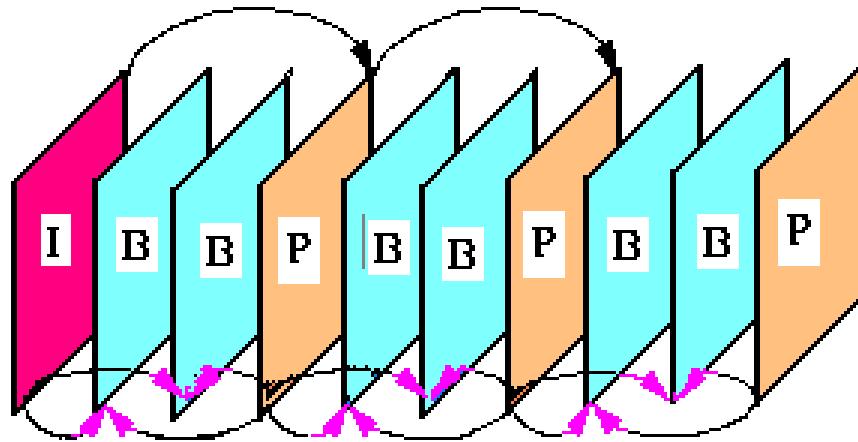
宏块的结构(4:2:0)





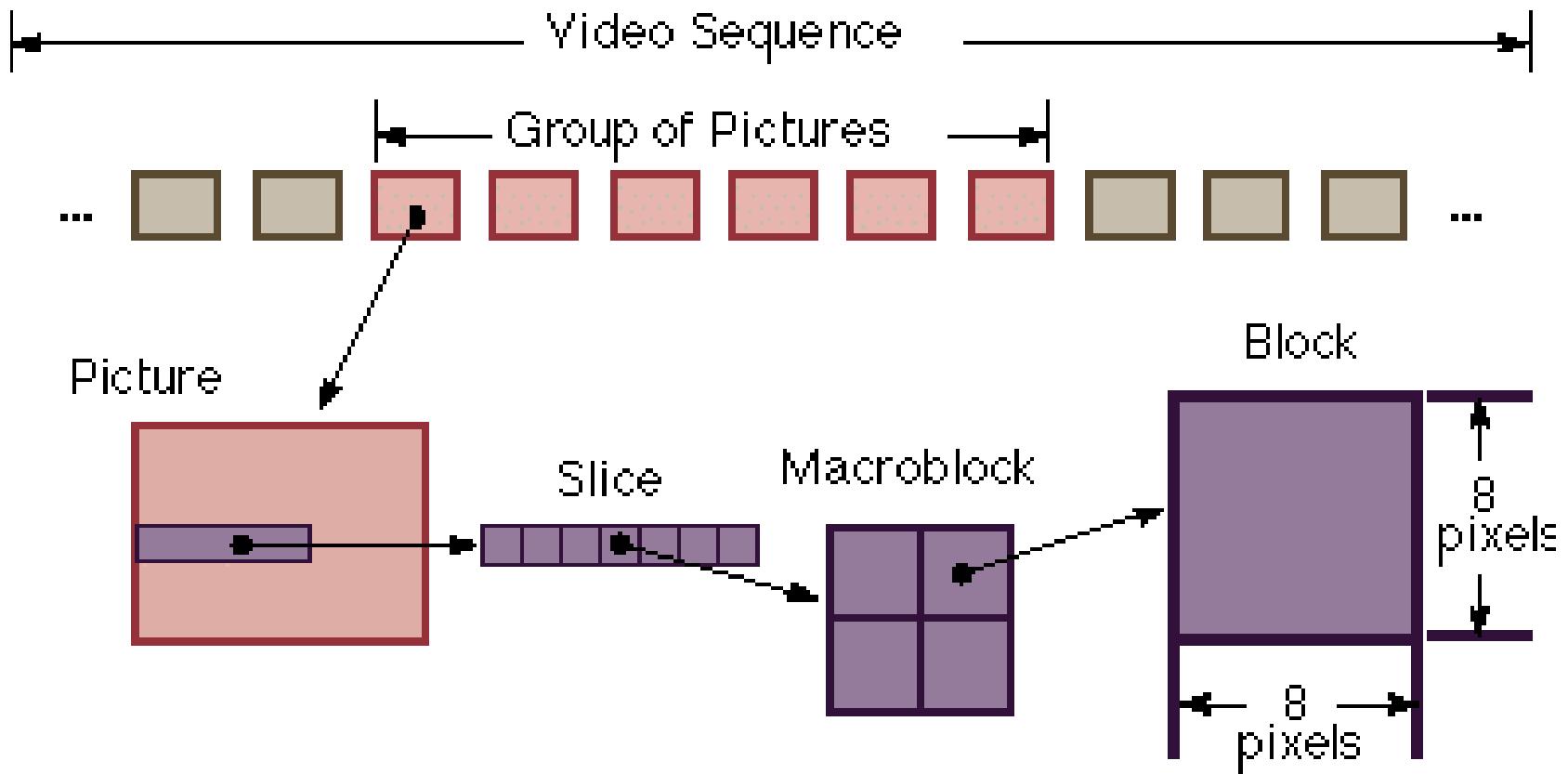
MPEG1、2中的三种图象帧

- ◆ 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)



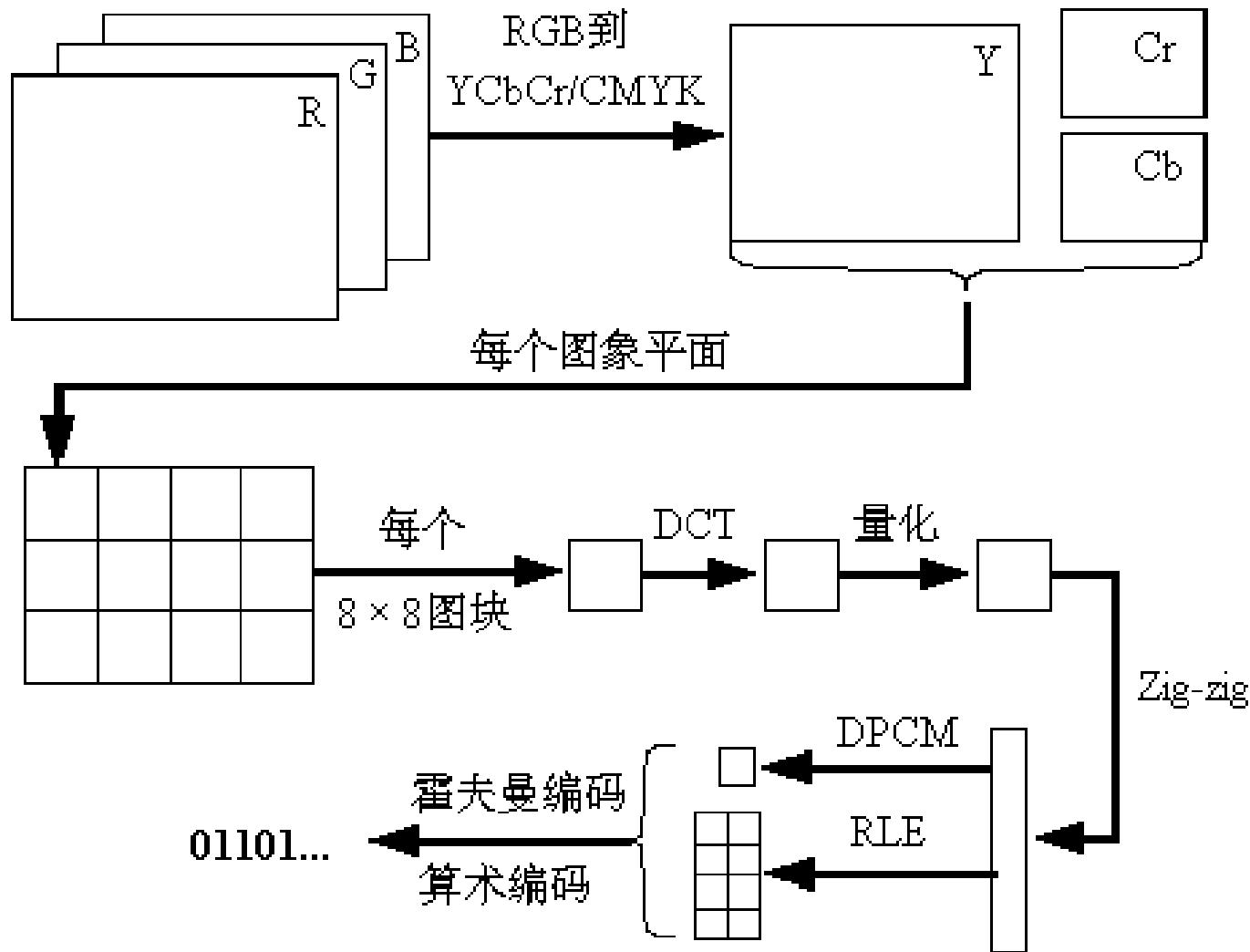


视频序列的组成



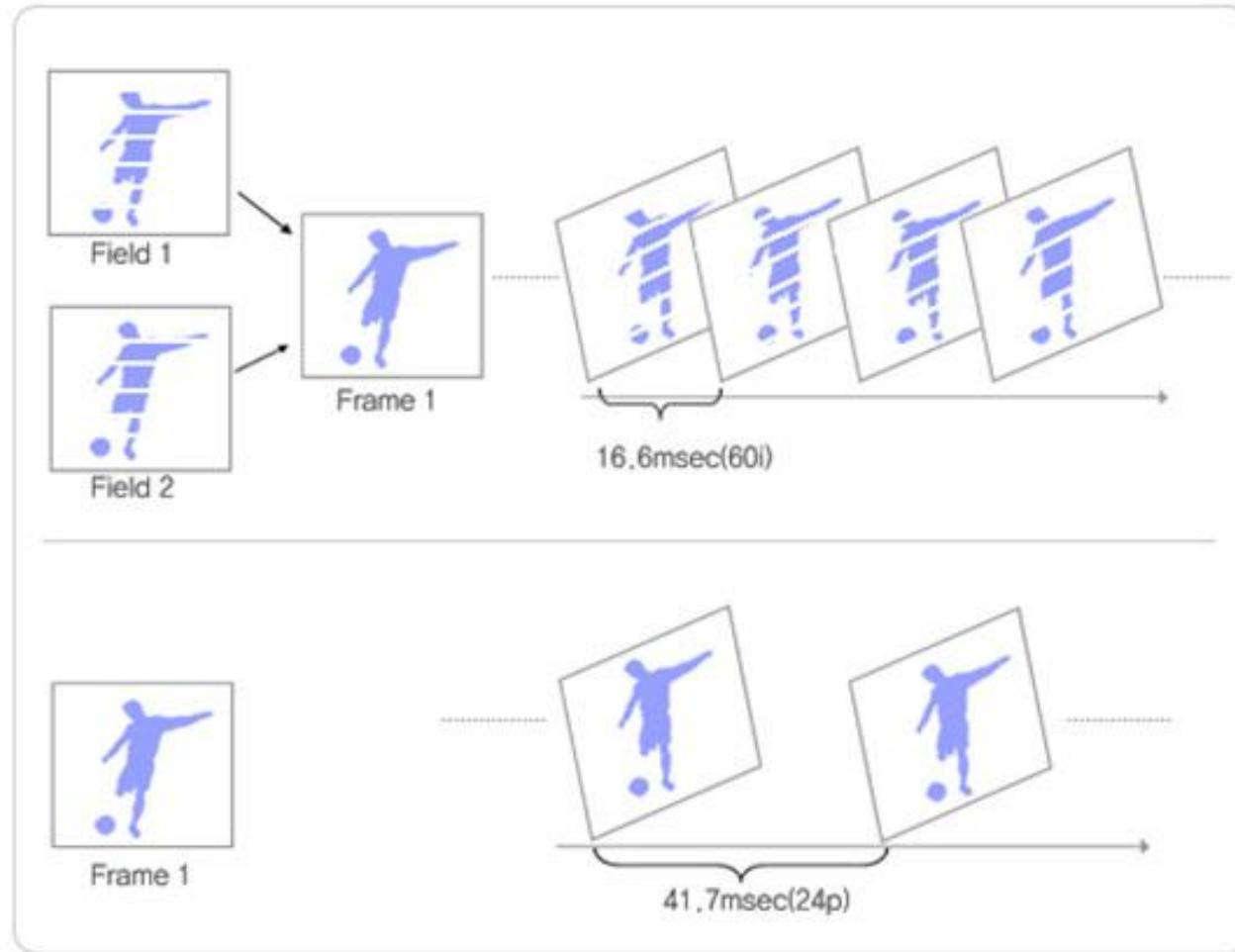


帧内图像I的压缩编码算法





MPEG 2 – Interlace/Progressive



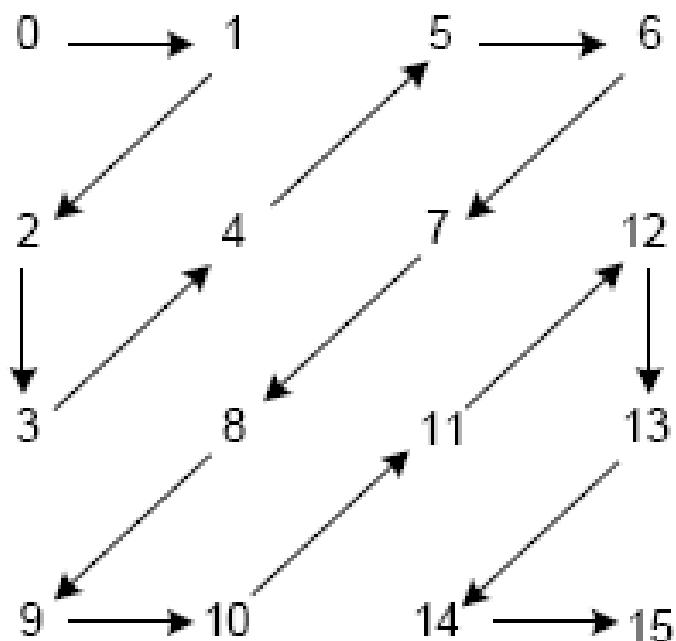
Interlace/Progressive: 隔行扫描/逐行扫描

{yfliu}@ustc.edu.cn

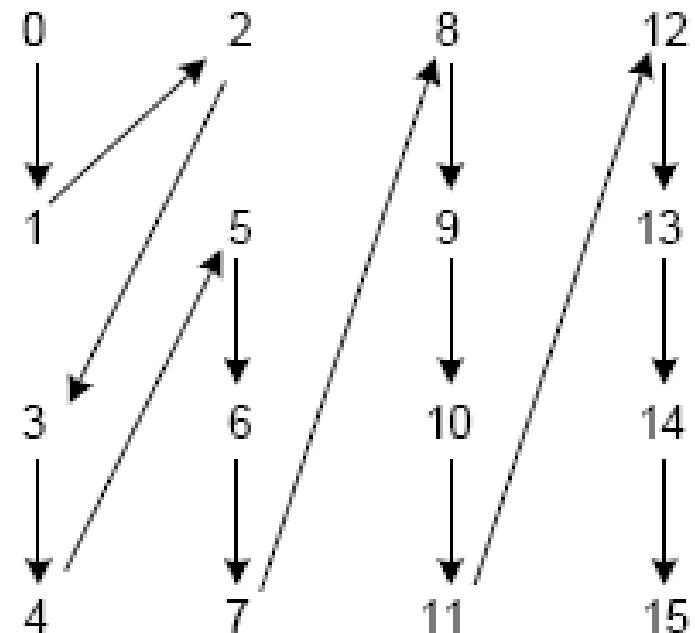




MPEG2 – DCT系数扫描顺序



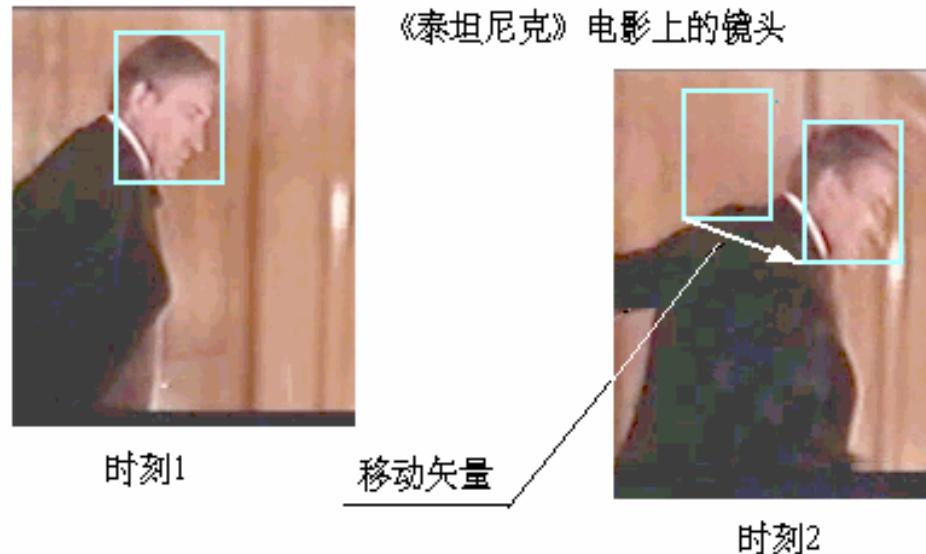
Frame mode: 帧模式



Field mode: 场模式

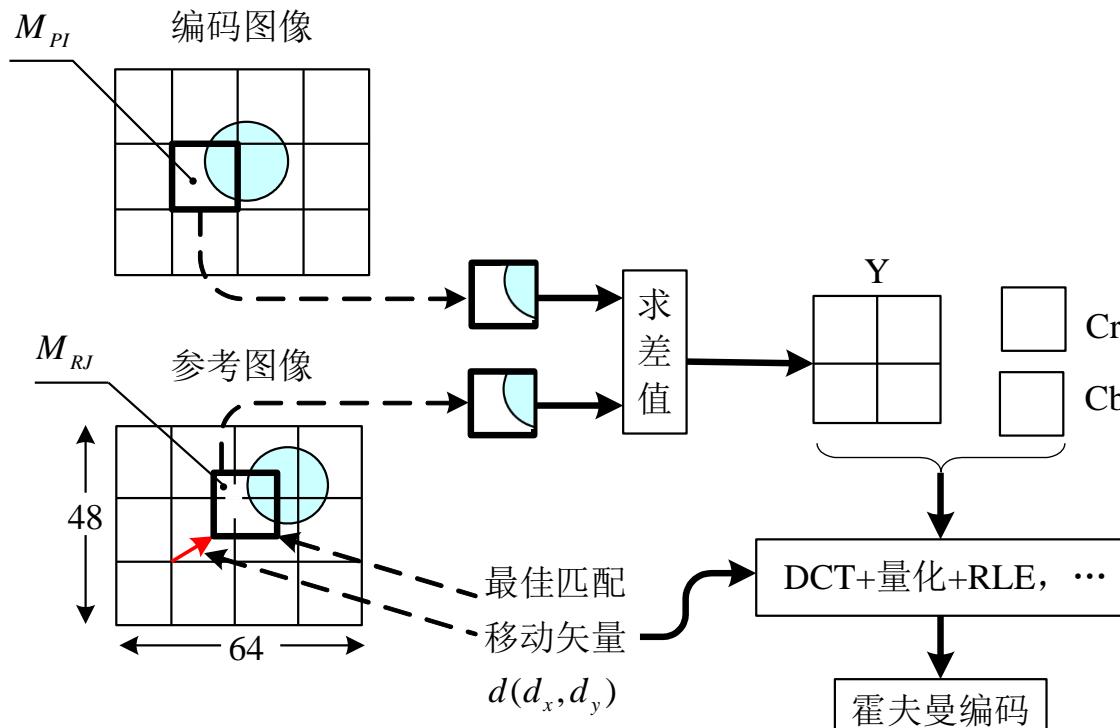
预测图像P的压缩编码算法

- ◆ 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**。



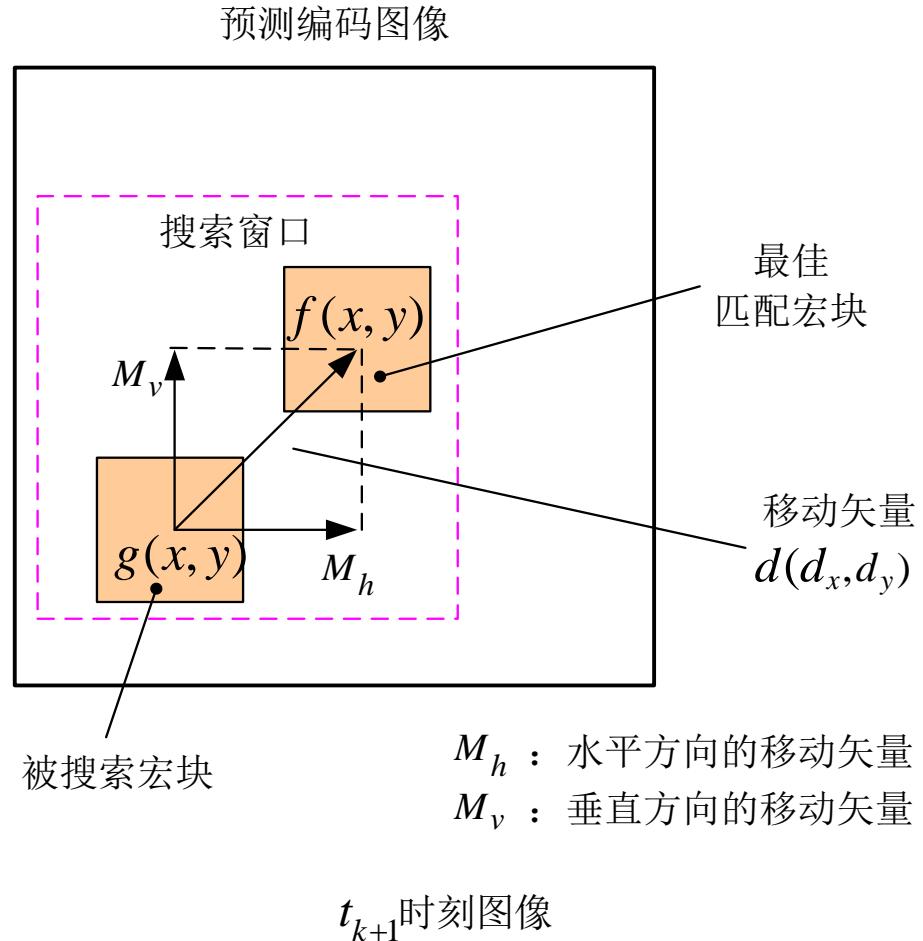
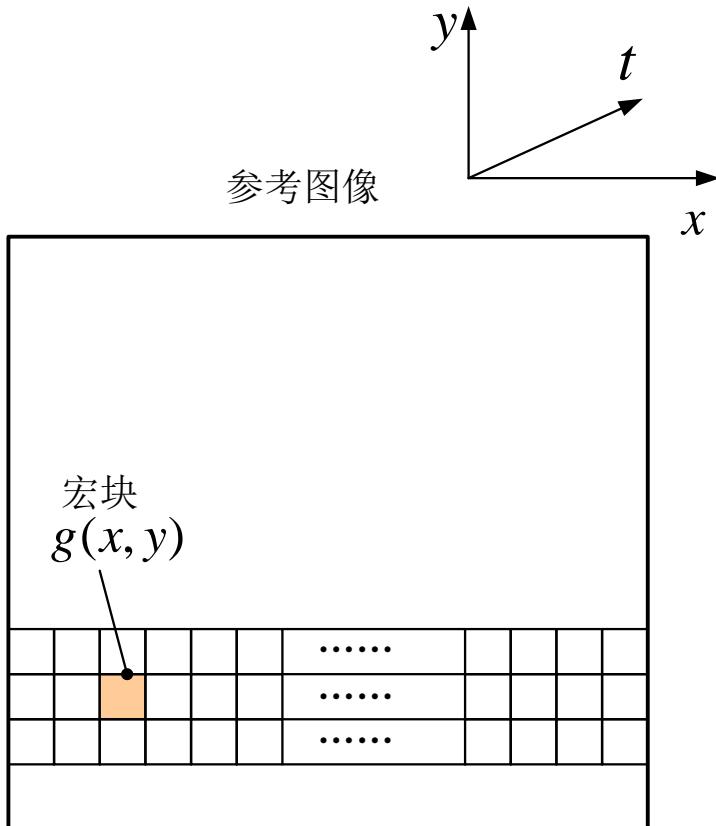
差值的记录

- ◆ 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。





移动矢量算法原理





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
-





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

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

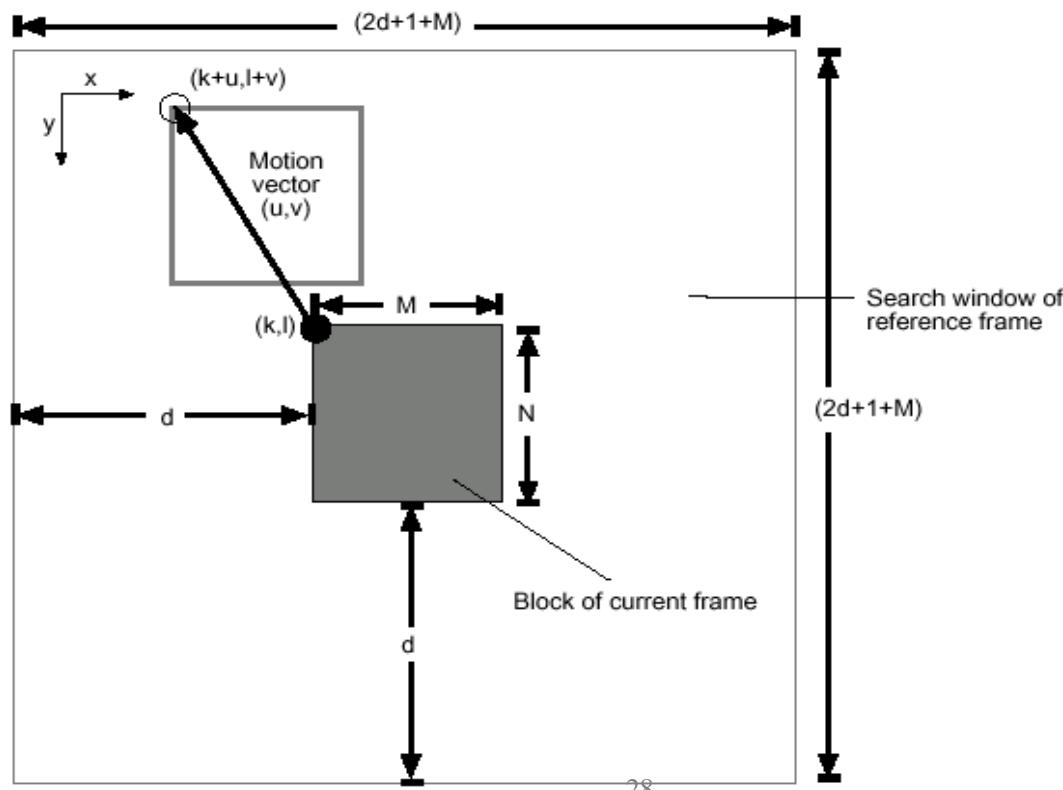
◆ 平均绝对帧差MAD(mean of the absolute frame difference)最小

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

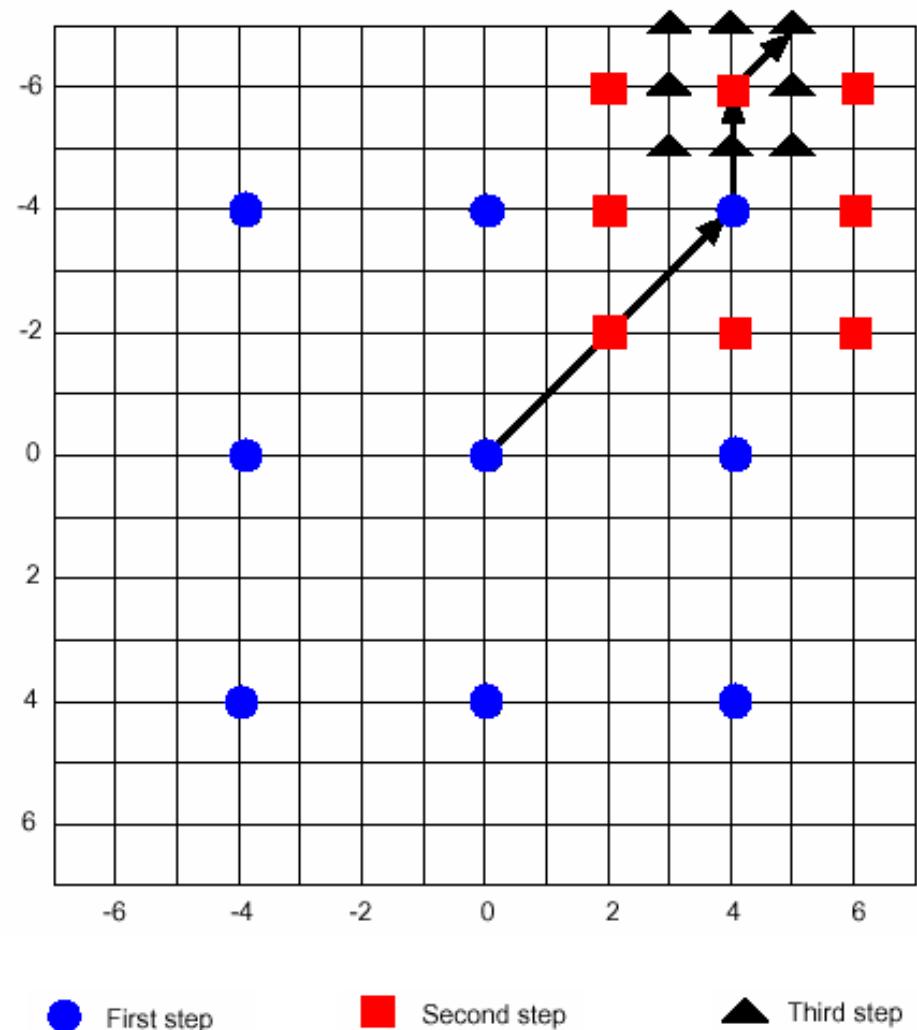
dx和dy分别是参考宏块MRJ的移动矢量d(dx, dy)在X和Y方向上的矢量



- ◆ 最简单的算法
- ◆ 计算量最大

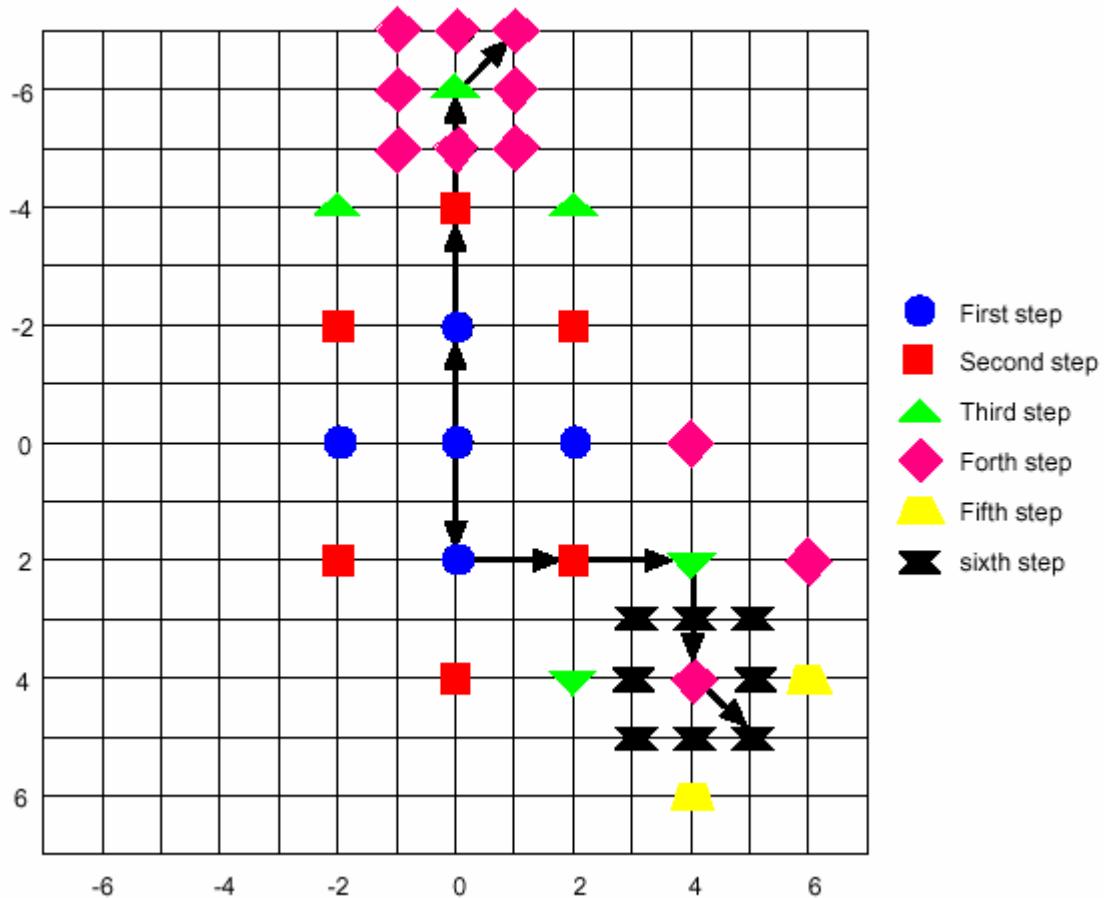


三步搜索法(three-step search)



- ◆ 基于由粗到精的算法，step size如图所示逐步下降
- ◆ 初始步长为最大运动位移的一半

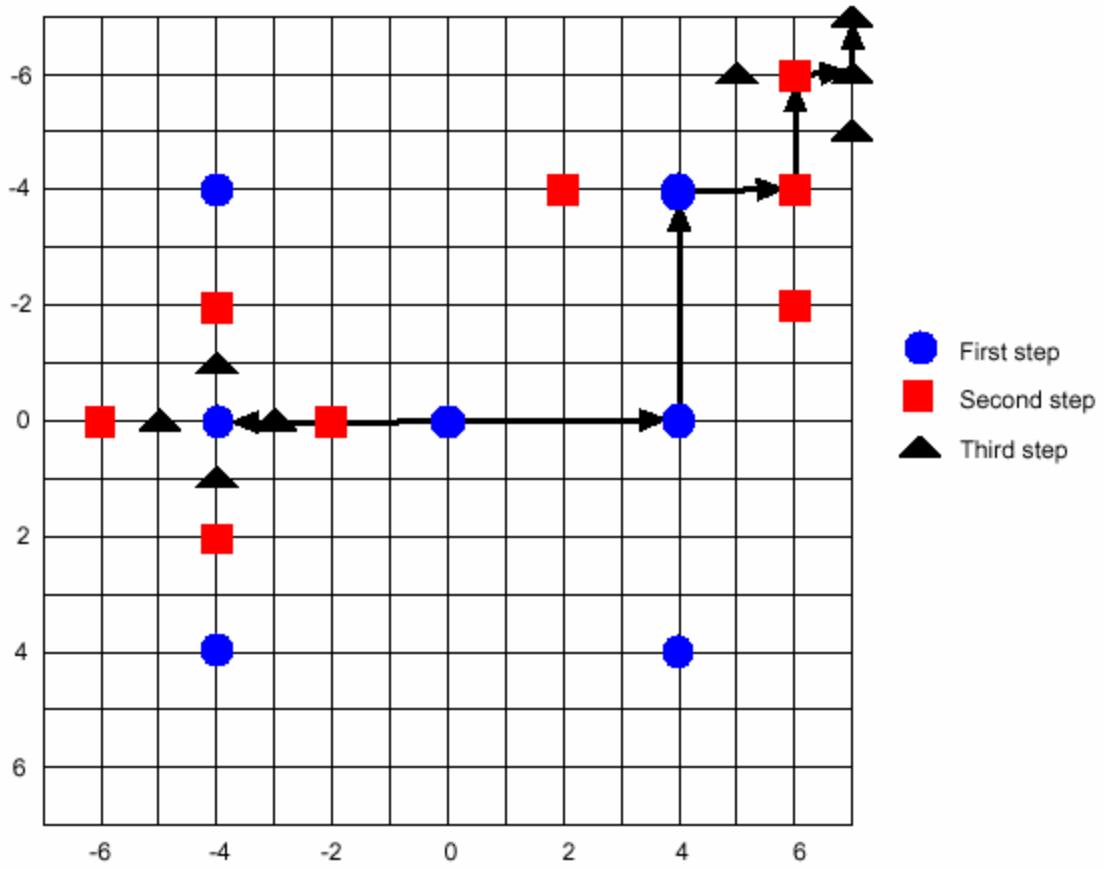
二维对数搜索法(2D-logarithmic search)



- ◆ 从原点开始，以十字形分布的五个点构成每次搜索的点群
- ◆ 初始化步长 $[d/4]$
- ◆ 若最小失真在中央或在边界则减步长
- ◆ 如果步长为1则搜索8个相邻点
- ◆ 例子中显示了两条不同搜索路径

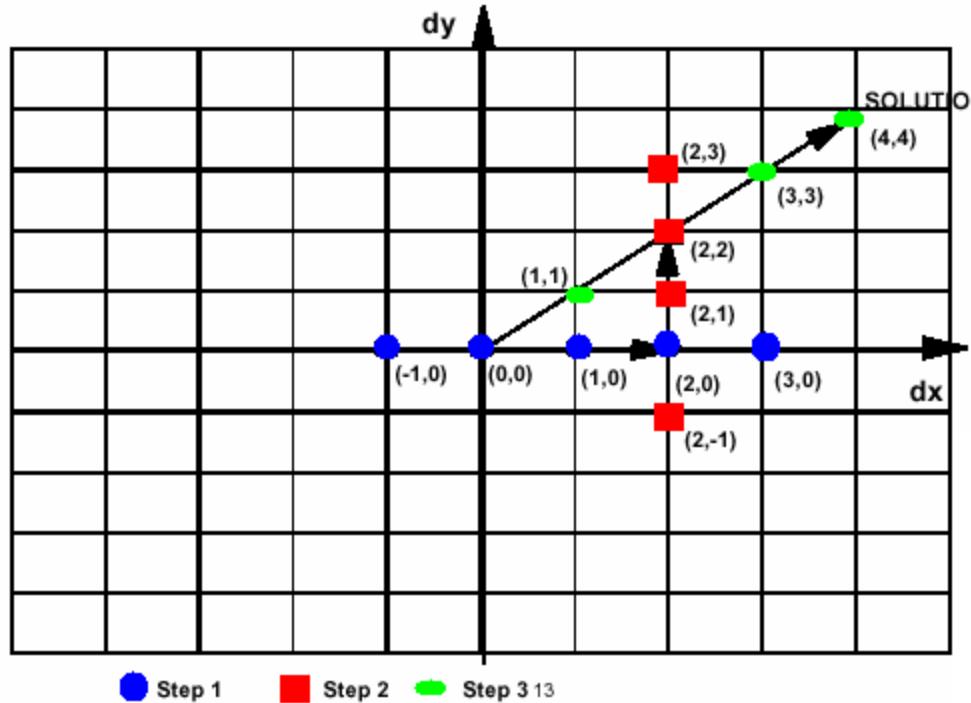
正交搜索算法

OSA: orthogonal search algorithm



- ◆ 包含水平和竖直方向上的搜索步骤，步长（step size）以对数形式递减，初始步长为 $d/2$
- ◆ 搜索由水平方向开始，当前数据块在水平方向上相邻的3个数据块作为 checking points，与待匹配数据块距离最小的数据块作为垂直搜索方向上的中心，重复上面水平方向的步骤（注意步长减半），交替进行水平与竖直方向上的搜索过程，直到搜索步长小于预设的搜索精度

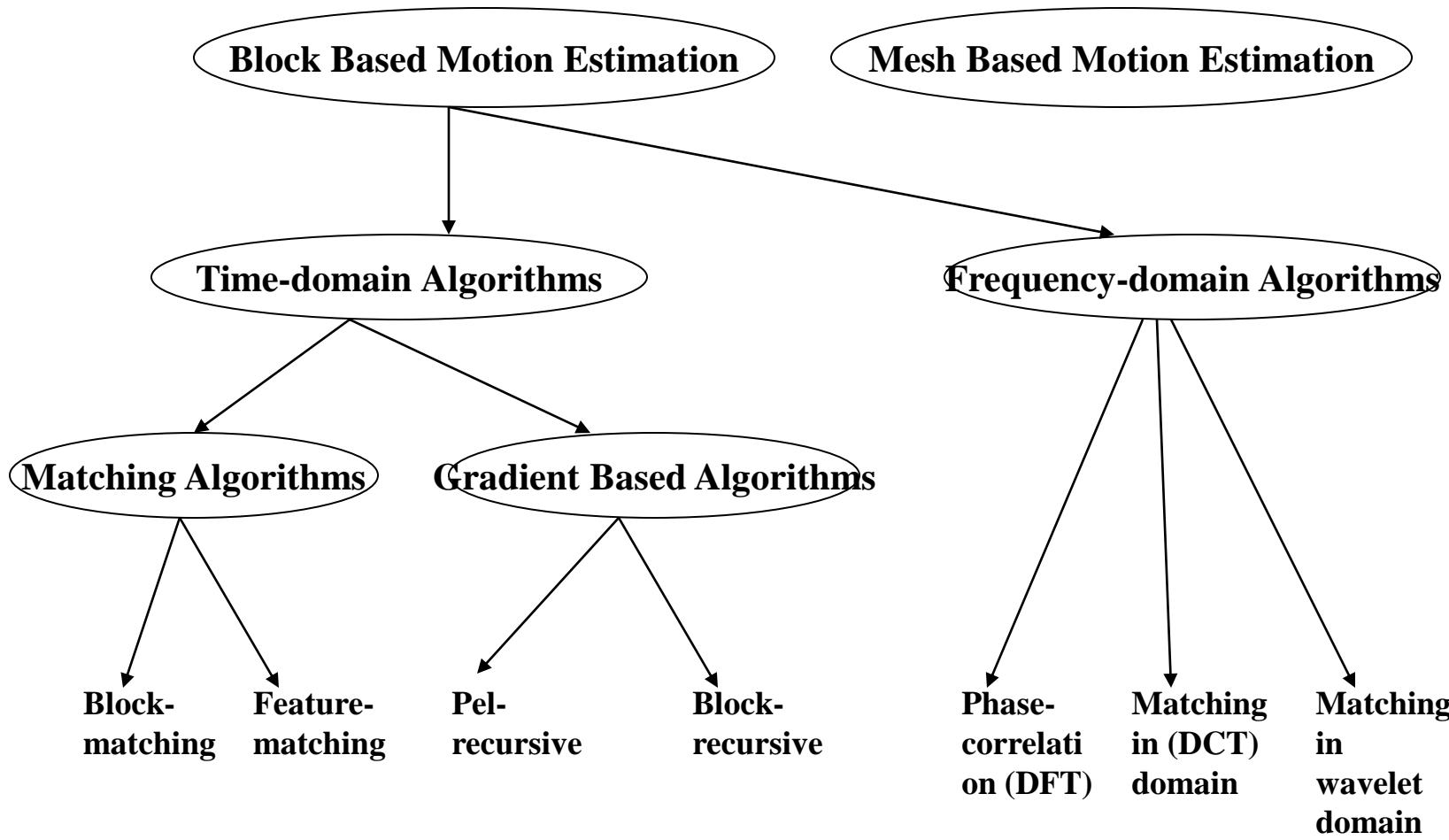
对偶搜索法(conjugate search)



1. Search in x direction
2. Search in y direction
3. Search in diagonal direction

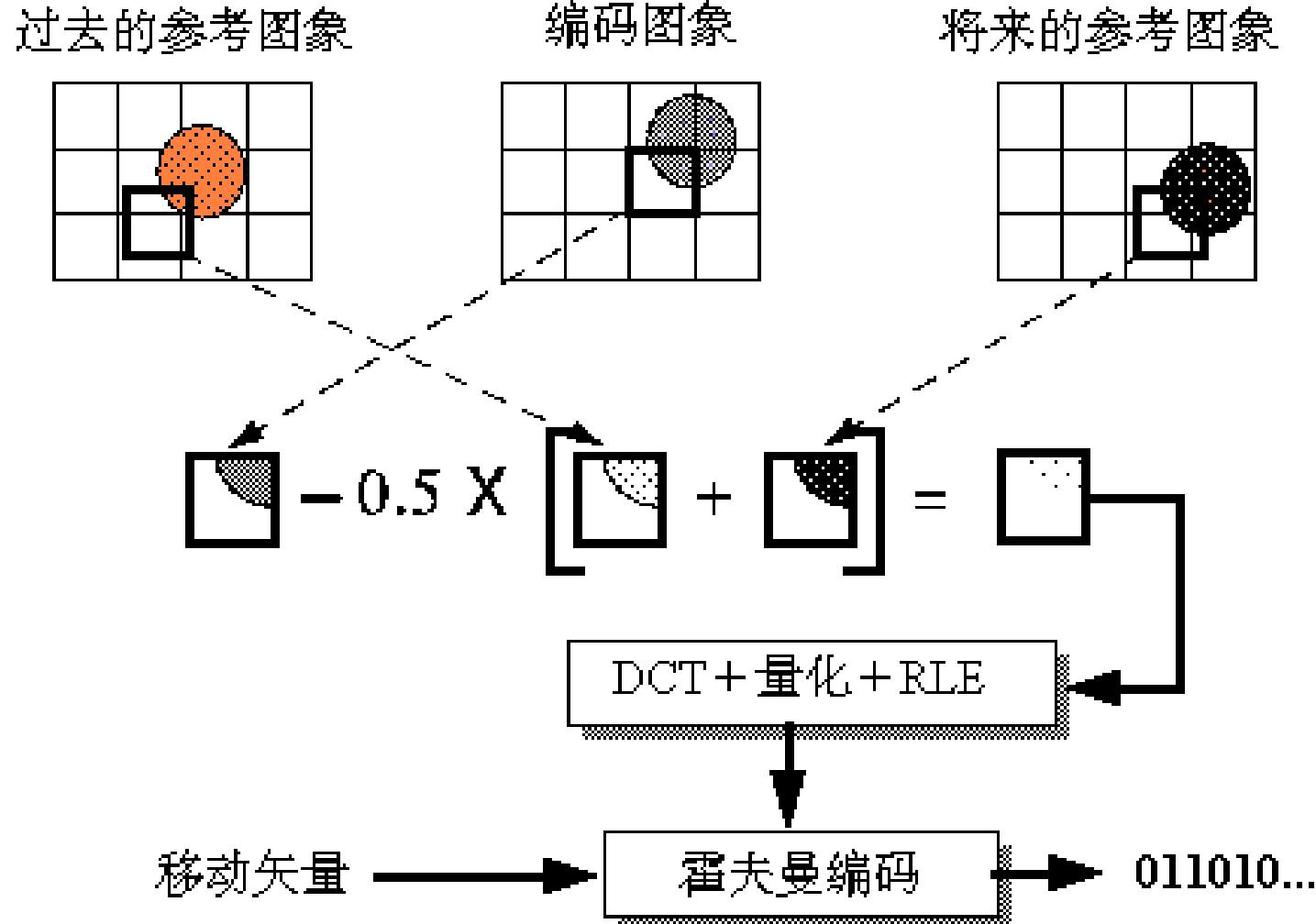


运动估计



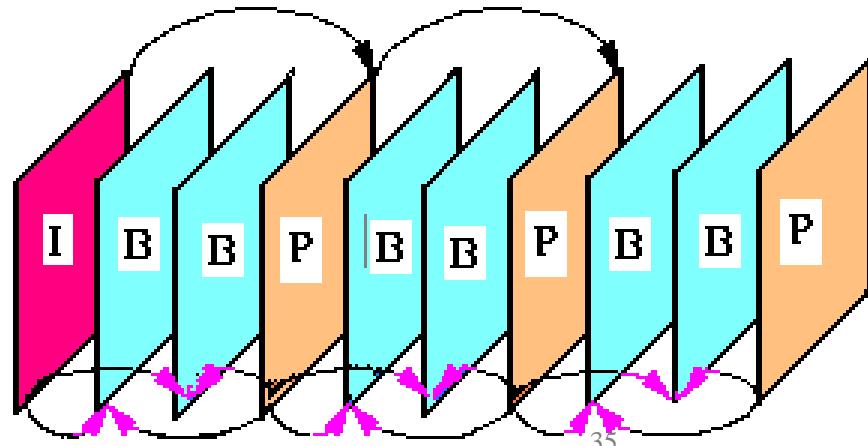
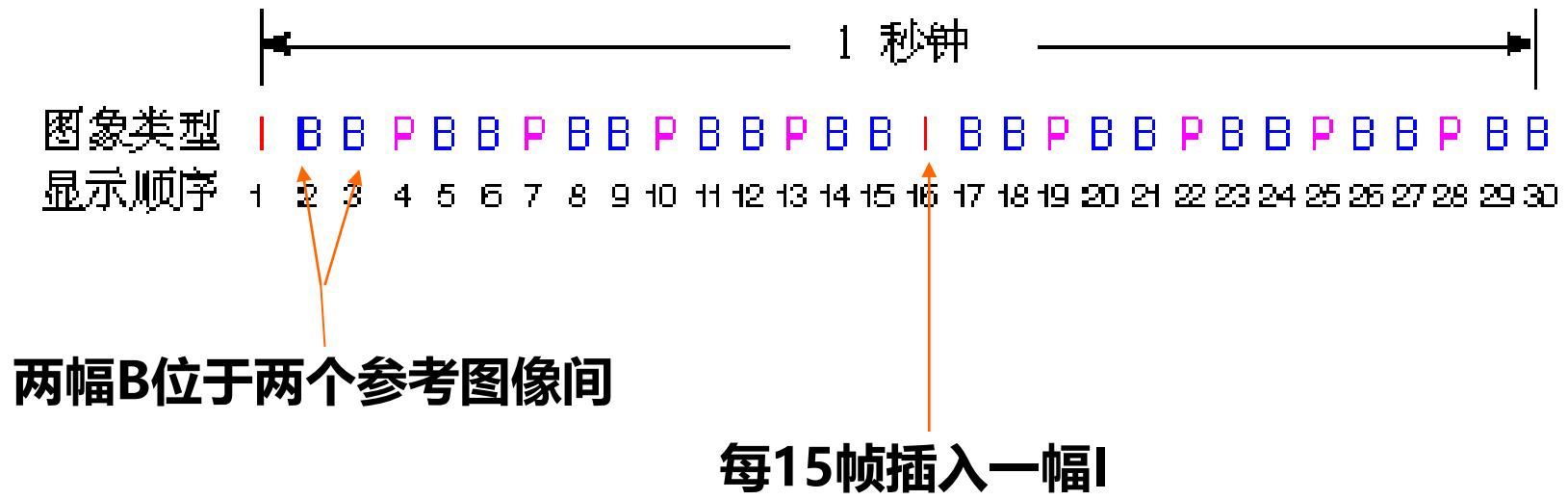


插补图像B的压缩编码算法





MPEG Video 典型流结构





MPEG-2的配置和等级

配置(Profile)	算法(Algorithms)
High (高档)	<ul style="list-style-type: none">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子采样各种: YUV 4:2:2 用于进一步提高图像质量
Spatial scalable (空间分辨率可变)	<ul style="list-style-type: none">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)子采样格式: YUV: 4:2:0
SNR scalable (信噪比可变)	<ul style="list-style-type: none">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)子采样格式: YUV: 4:2:0
Main (基本)	<ul style="list-style-type: none">非可变速率编码算法支持随机存取, B图像预测方式子采样格式: YUV: 4:2:0
Simple (简化)	<ul style="list-style-type: none">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子采样格式: YUV: 4:2:0





MPEG-2等级

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

◆ MP@ML

- (Main Profile, Main Level)
 - 指 $720 \times 576 \times 30$, 4:2:0, 15 Mb/s。

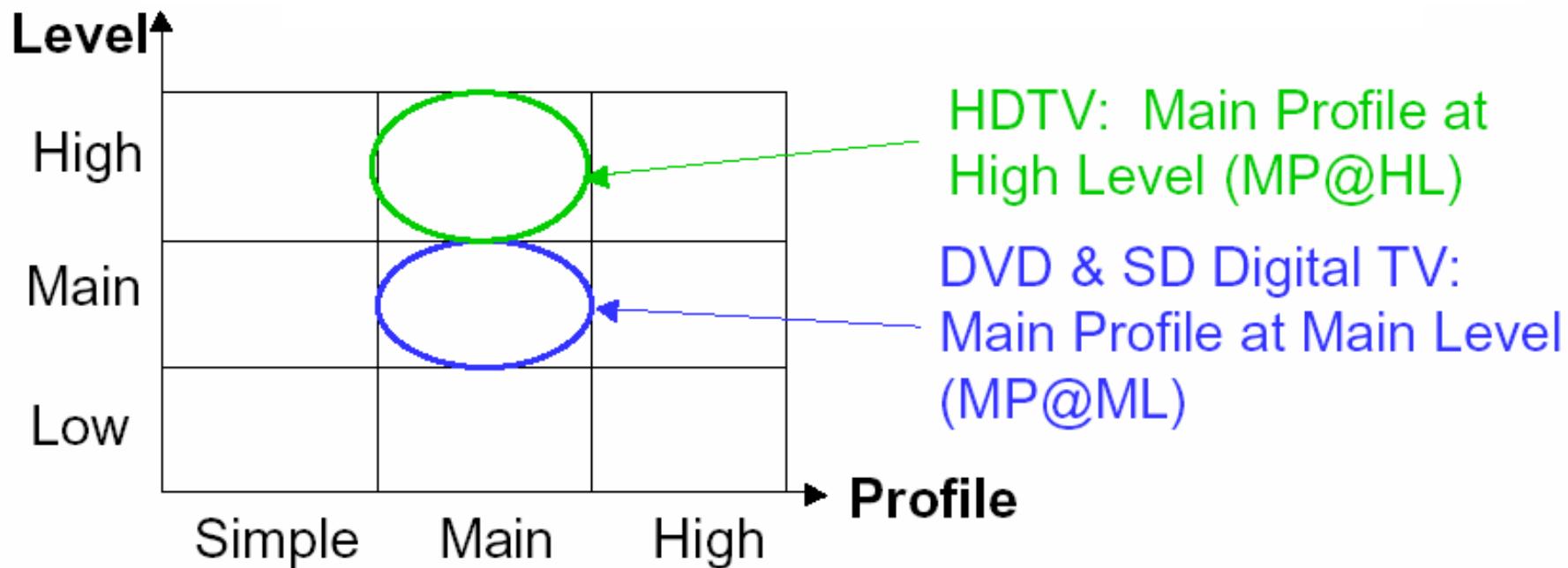
◆ MP@HL

- (Main Profile, High Level)
 - 指 $1920 \times 1152 \times 60$ 、4:2:0、80 Mb/s





MPEG2中常用Profile@Level





第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - MPEG简介
 - MPEG-1、MPEG-2视频压缩原理
 - **MPEG-4视听对象编码简介**
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介





MPEG-4是什么

◆ 视听对象编码(Coding of audio-visual objects)标准

- 版本1：1998年10月，始于1993年
- 版本2：1999年底
- 国际标准：2000年，ISO/IEC 14496
- 不断更新和完善.....

◆ 目标：为通信、广播、存储和其他应用提供数据速率低而视听质量高的数据编码方法和交互播放工具

- 吸收了MPEG-1、MPEG-2和其他相关标准的许多特性
- 引入了**视听对象(audio-visual objects, AVO)**编码的概念
- 扩充了编码类型，由自然对象扩展到合成对象
- 采用了合成对象与自然对象混合编码(Synthetic/Natural Hybrid Coding, SNHC)算法
- 引入了组合、合成和编排等重要概念，以实现交互功能和对象重用





如何理解对象(object)

◆ 文档

- 可作为整体操作的实体，如数据块，文字块，图像

◆ 面向对象程序设计

- 自身包含数据和相关处理指令的可重用模块，通常由函数(通过名称调用的一组指令)和变量(函数使用的有名称的存储区)组成

◆ 计算机图形

- 2-D或3-D信息块，如mesh(网格)，curve(曲线)、cube(立方体)，包含位置、旋转、大小和变换矩阵等信息，可被连接到其他对象

◆ 计算机安全

- 受到控制的实体，如文件、程序或存储器等

◆ 计算机通信

- 严格定义的信息、定义或规范，有自己的名称，以便在通信实例中能够被识别

◆ 人工智能

- 物理或概念上的实体，有一个或多个属性

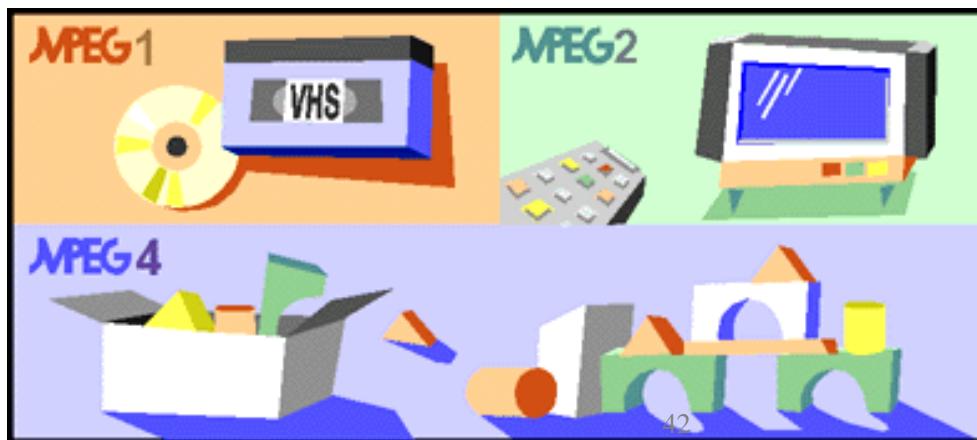




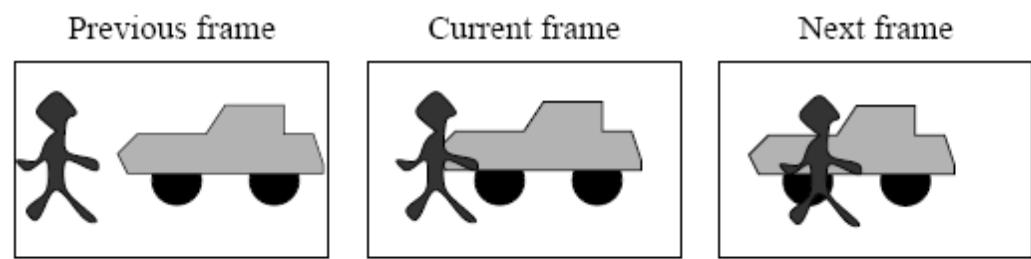
Audio Visual Object

◆ AV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。

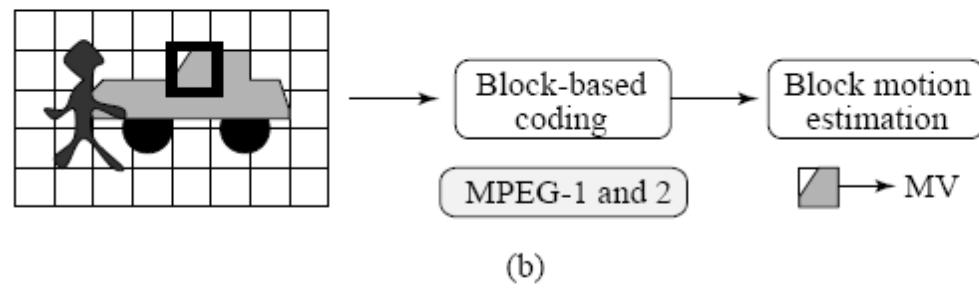
□ 对象是指在一个场景中能够访问和操纵的实体，对象的划分可根据其独特的纹理、运动、形状、模型和高层语义为依据。在MPEG-4中所见的视音频已不再是过去MPEG-1、MPEG-2中图像帧的概念，而是一个个视听场景（AV场景），不同的AV场景由不同的AV对象组成。



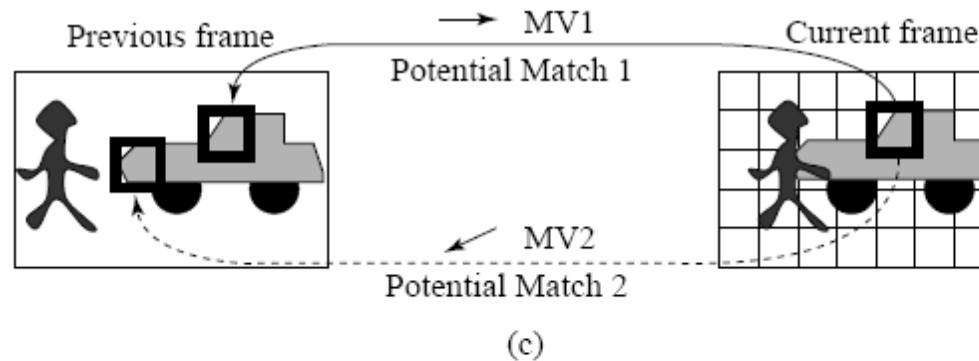
Block based & Object based



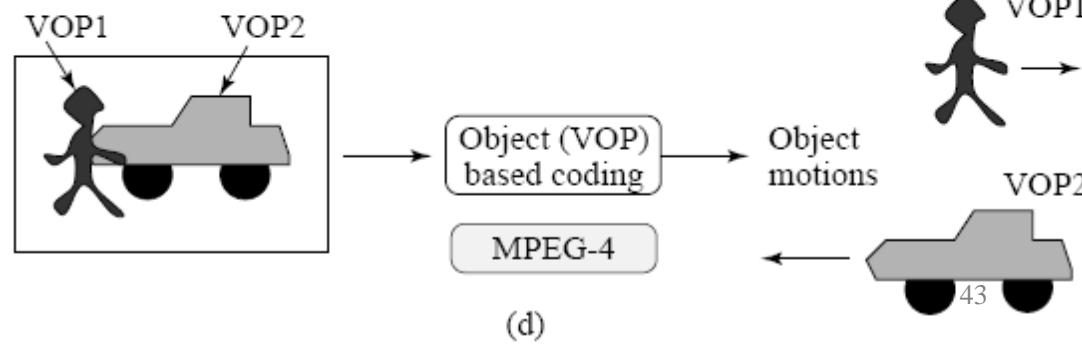
(a)



(b)



(c)



(d)

◆ A

□ 原始图像

◆ B

□ 基于宏块的

◆ C

□ 引起宏块方式中
预测错误

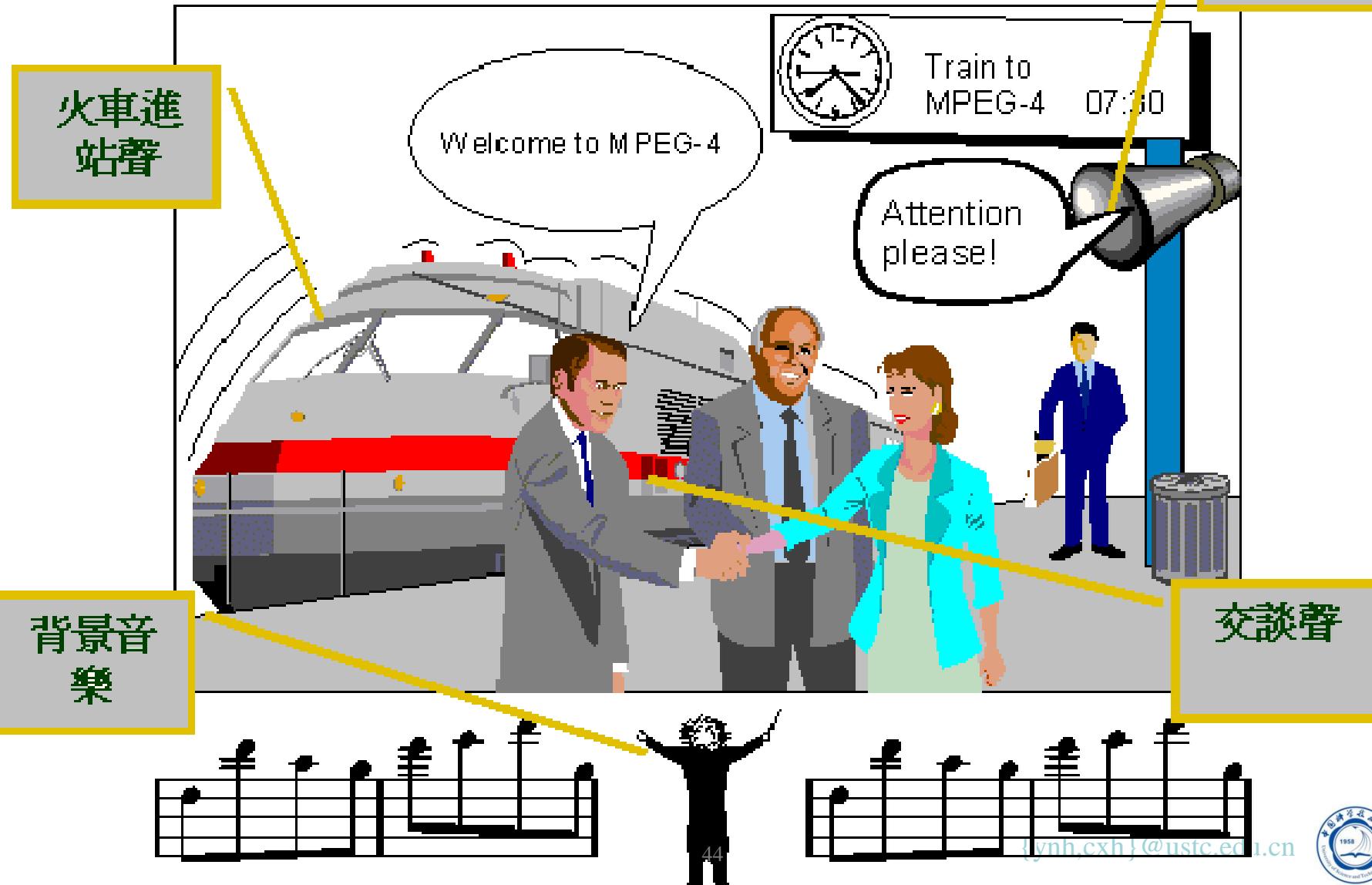
◆ D

□ 基于对象的



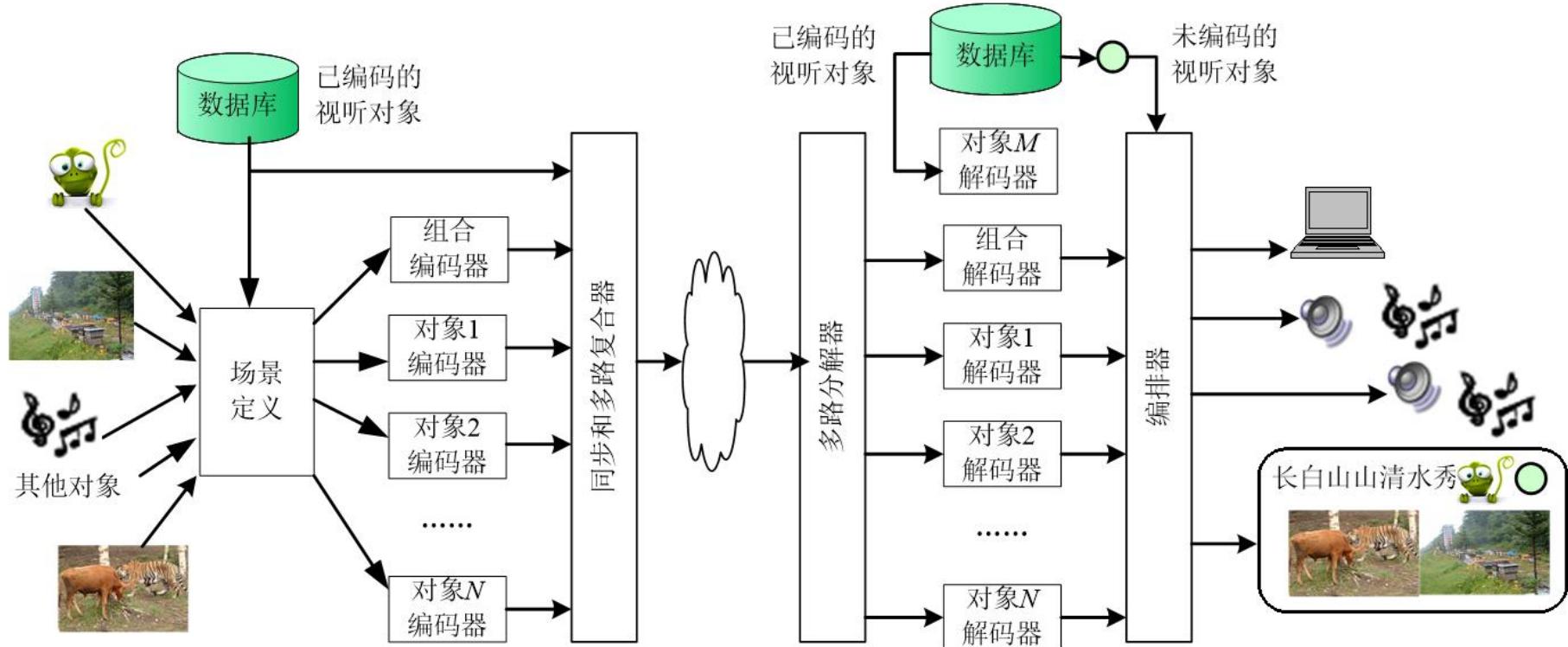
MPEG-4音频对象

廣播聲



MPEG-4基于对象的系统模型

- ◆ 基于对象的系统模型由编码和解码两大部分组成
- ◆ 主要功能：用于互动视听场景通信





MPEG-4由一系列的子标准组成

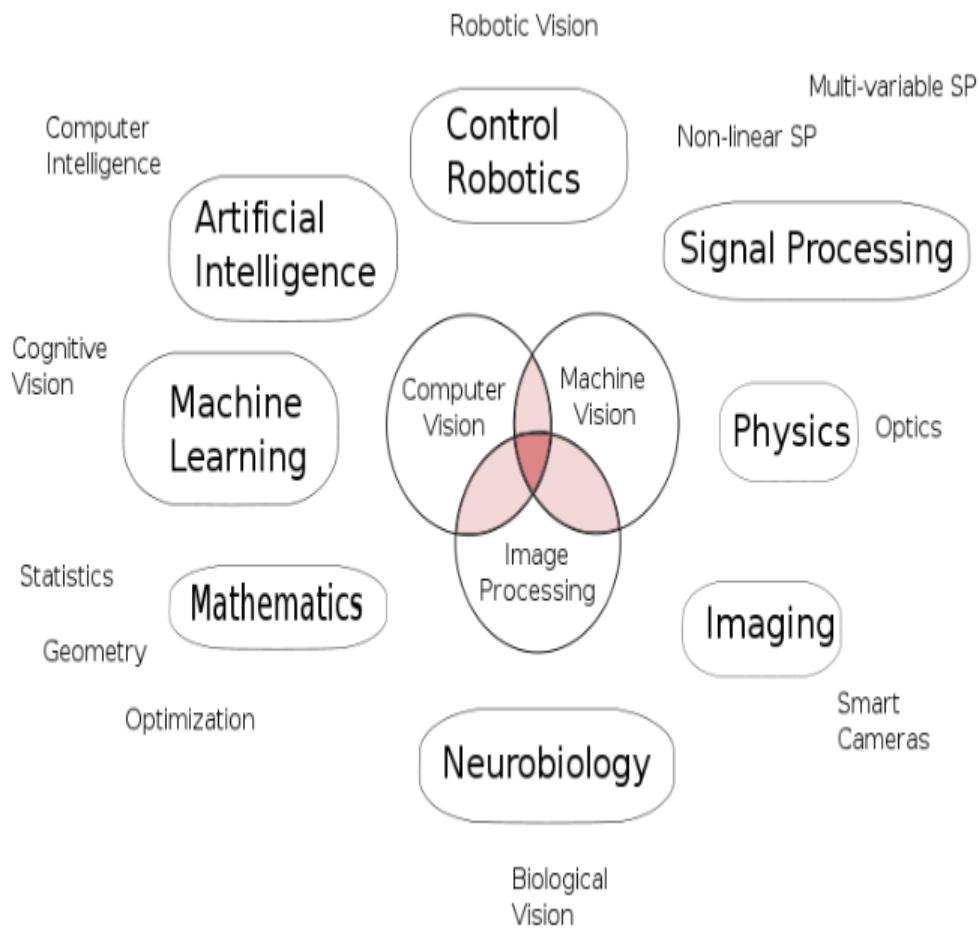
- ◆ part1 (ISO/IEC 14496-1): 系统: 描述视频和音频的同步以及混合方式(multiplexing)。
- ◆ part2(ISO/IEC 14496-2): 视频: 定义了对视觉信息的编解码器。
- ◆ part3(ISO/IEC 14496-3): 音频: 定义了一个对各种音频信号进行编码的编解码器的集合。
- ◆ part4(ISO/IEC 14496-4): 一致性: 定义了对本标准其他的部分进行一致性测试的程序。
- ◆ part5(ISO/IEC 14496-5): 参考软件。
- ◆ part6(ISO/IEC 14496-6): 多媒体传输集成框架。
- ◆ 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IP网络上传输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。



◆ MPEG4的困难在于标准之外的技术：计算机视觉

◆ Computational Vision

◆ 用各种成象系统代替视觉器官作为输入敏感手段，由计算机来代替大脑完成处理和解释。计算机视觉的最终研究目标就是使计算机能象人那样通过视觉观察和理解世界，具有自主适应环境的能力。有不少学科的研究目标与计算机视觉相近或与此有关。这些学科中包括图像处理、模式识别或图像识别、景物分析、图象理解等。





第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - MPEG简介
 - MPEG-1、MPEG-2视频压缩原理
 - MPEG-4视听对象编码简介
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介





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。H.26L这个名称虽然不太常见，但是一直被使用着。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

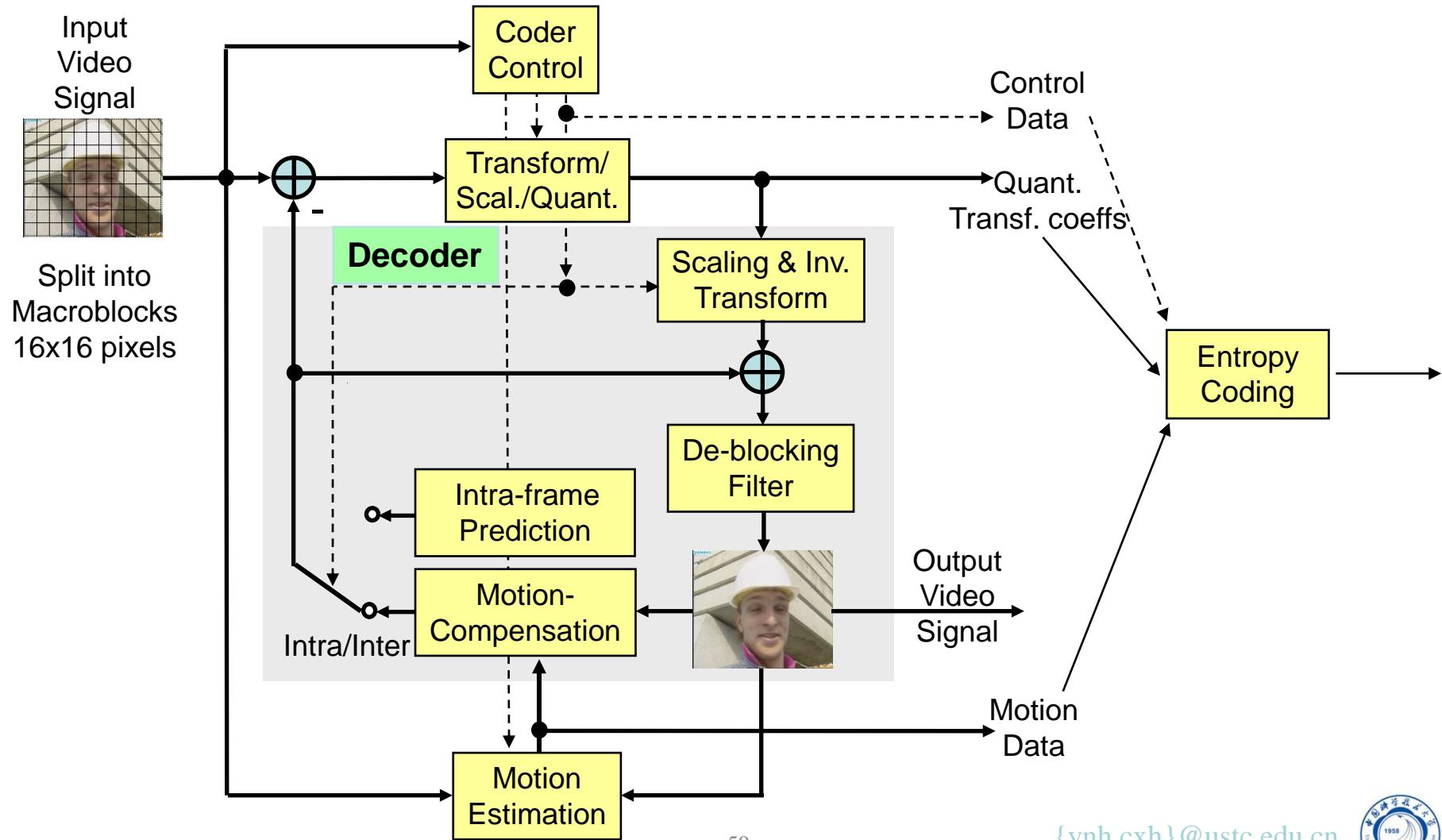
□ JVT (Joint Video Team, 2001)

□ ITU-T和MPEG两个机构合作开发

● 之前的视频编码标准MPEG-2也是由MPEG和ITU-T两方合作开发，因此MPEG-2在ITU-T的命名规范中被称之为H.262)



基本宏块编码结构





主要特点对比

性能	标准		
	MPEG-2	MPEG-4 Visual	H.264/AVC
块变换	8×8 DCT	8×8 DCT/ 小波变换	4×4, 8×8 整数 DCT, 4×4, 2×2 哈达玛变换
帧内预测	帧内 DC 预测	变换域预测	空间域预测
双向预测	向前/向后	向前/向后	向前/向后, 向前/向前, 向后/向后
加权预测	-	-	有
移动估算	16×16	16×16 或 8×8	16×16~4×4
量化	HVS 加权 均匀量化	HVS 加权 均匀量化	量化阶增量 12.5%
熵编码	VLC	VLC	VLC, CAVLC, CABAC
像素精度	1/2 像素	1/4 像素	1/4 像素
图像类型	I, P, B	I, P, B	I, P, B, SI, SP
消块滤波	-	选择	有
参考图像	一帧	一帧	多帧
编码器	复杂度中等	复杂度中等	复杂度高
后向兼容	与先前标准 可兼容	与先前标准可兼容	与先前标准不兼容
传输速率	2~15Mbps	64kbps~4Mbps	64kbps~150Mbps

H.264/AVC继承了先前视频压缩标准的许多优点，虽然在结构上没有明显改变，只是在各个主要功能模块内部做了“小打小闹”和“精雕细刻”的改进，但正是这些改进使编码效率有了明显提高。





创新点

◆ 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 4×4 亮度块的预测、4种 16×16 亮度块的预测和4种色度块的预测

◆ De-blocking Filter

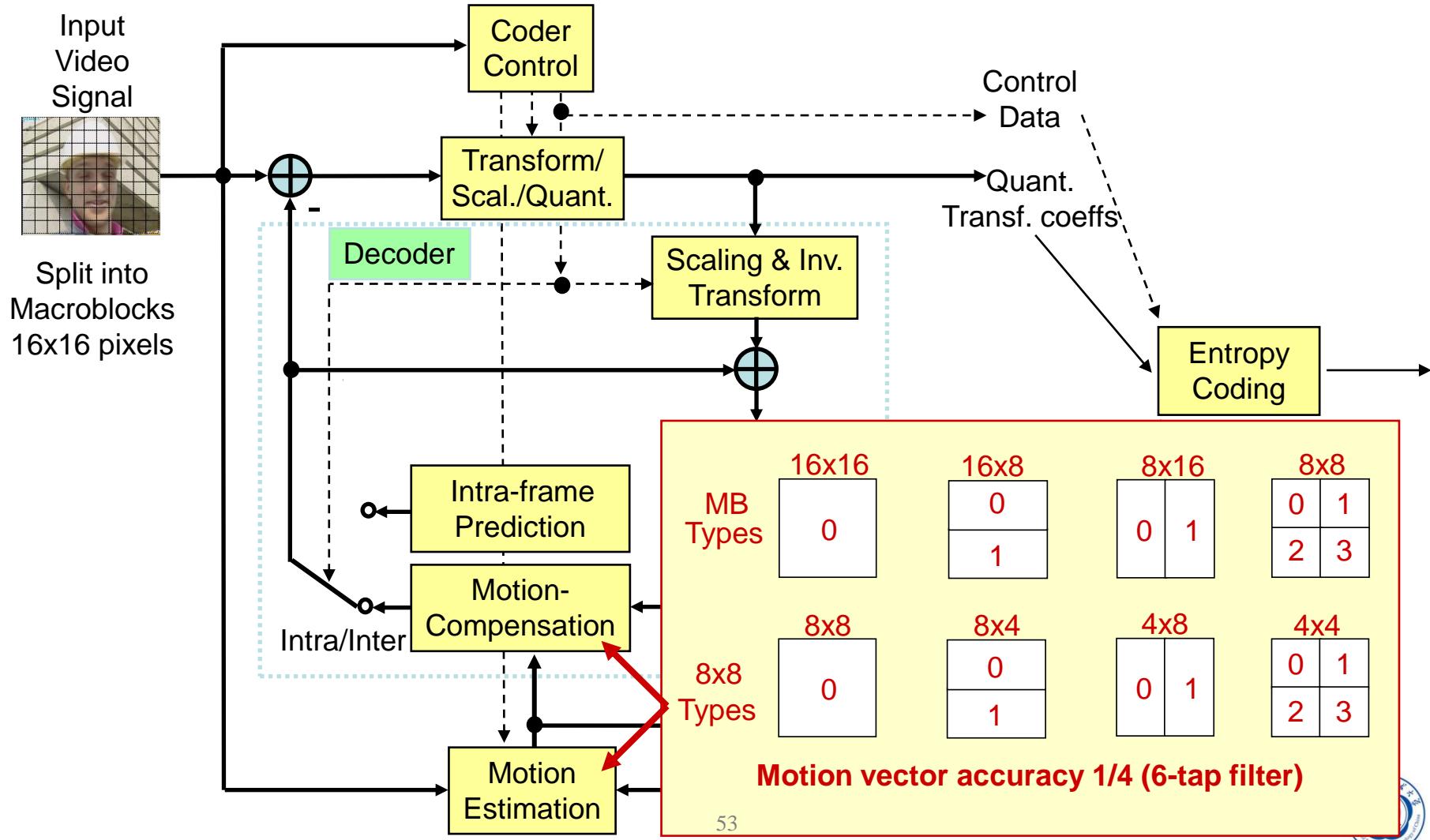
- 优秀的运动补偿滤波器

◆ Transform/Scal./Quant.

◆ Entropy Coding



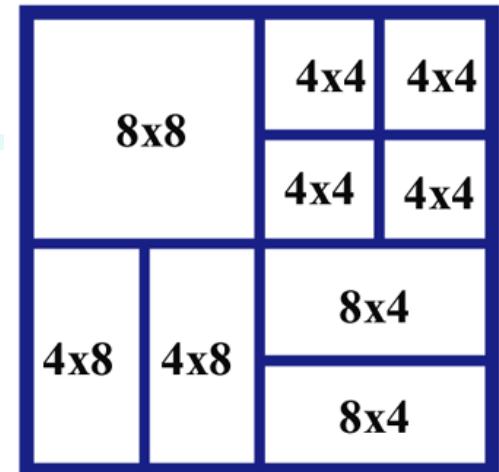
运动补偿精度





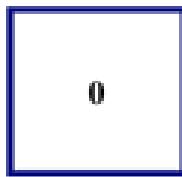
宏块大小

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



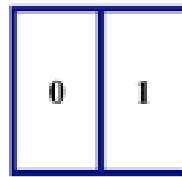
Mode 1

One 16x16 block
One motion vector



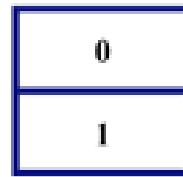
Mode 2

Two 8x16 blocks
Two motion vectors



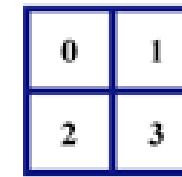
Mode 3

Two 16x8 blocks
Two motion vectors



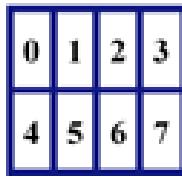
Mode 4

Four 8x8 blocks
Four motion vectors



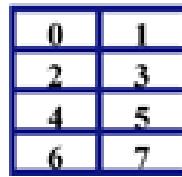
Mode 5

Eight 4x8 blocks
Eight motion vectors



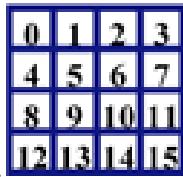
Mode 6

Eight 8x4 blocks
Eight motion vectors



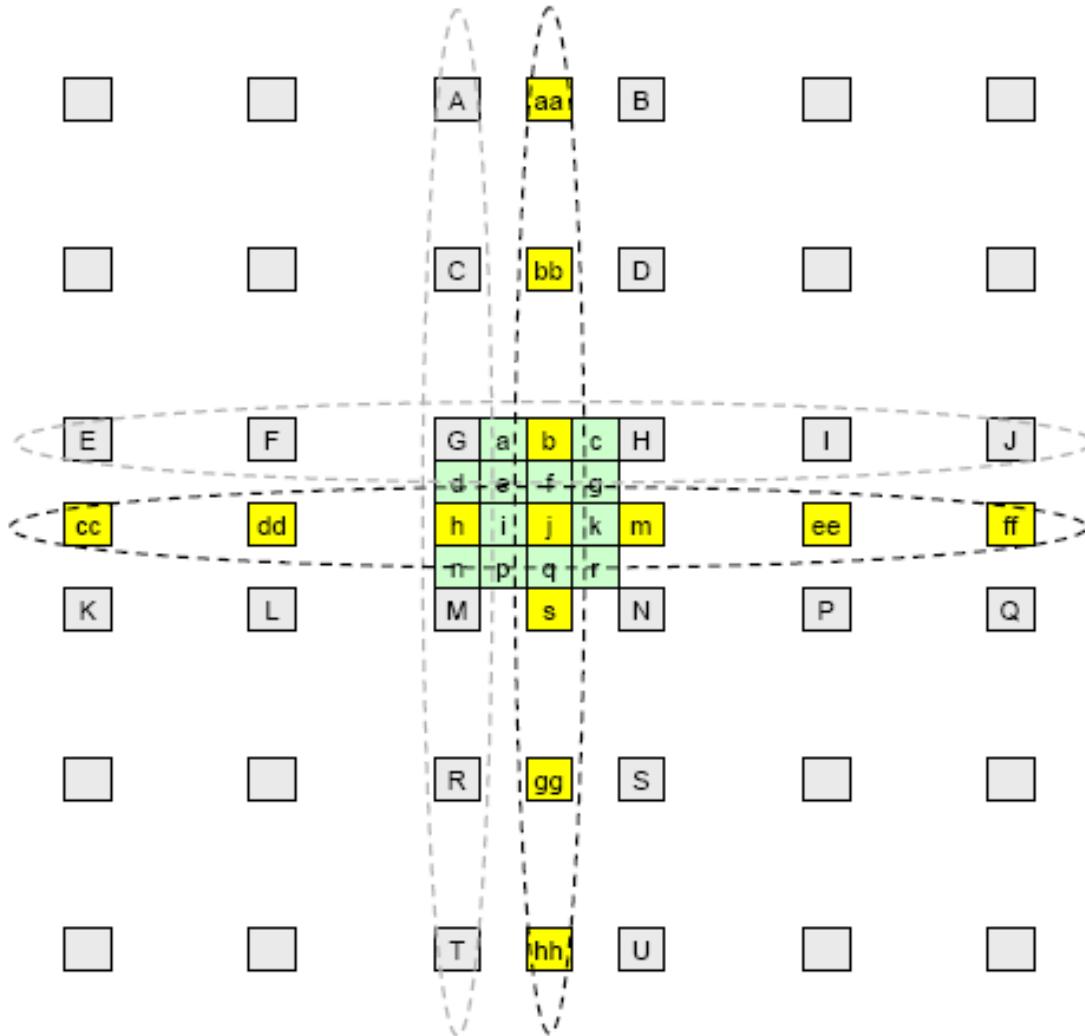
Mode 7

Sixteen 4x4 blocks
Sixteen motion vectors



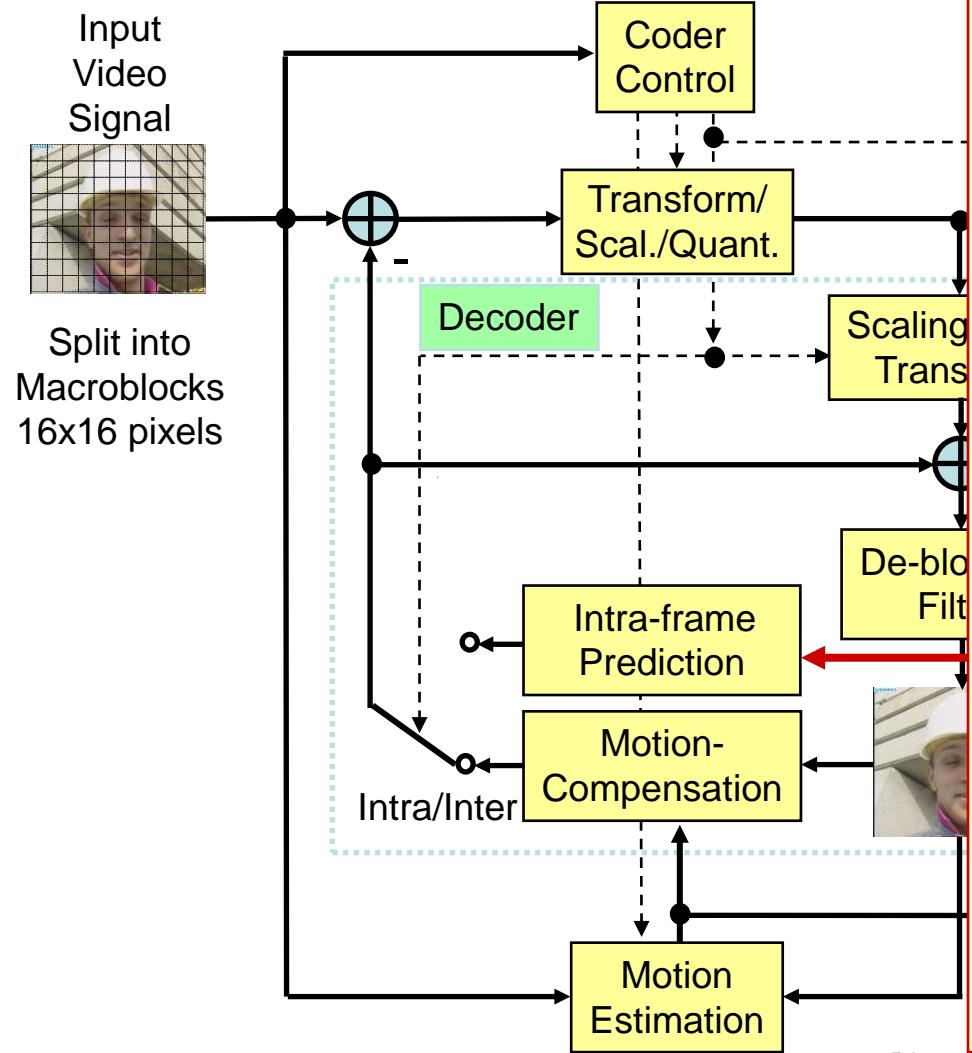


高精度的运动估计策略



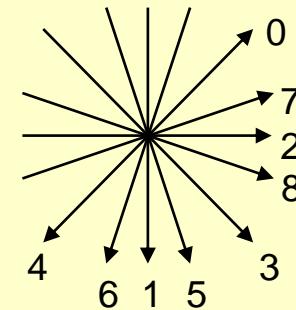
- ◆ $\frac{1}{2}$ -pel pixel grid values are interpolated using a separable 6-tap filter
- ◆ $\frac{1}{4}$ -pel pixel grid values are interpolated using the average of two $\frac{1}{2}$ -pel pixels

帧内预测 (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								



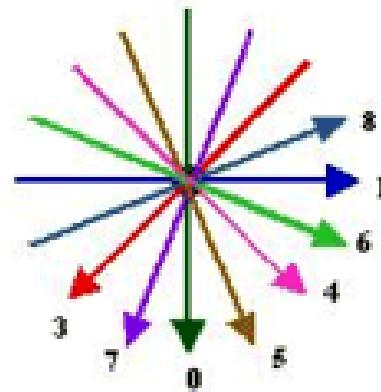
- e.g., Mode 3:
diagonal down/right prediction
a, f, k, p are predicted by
 $(A + 2Q + I + 2) \gg 2$



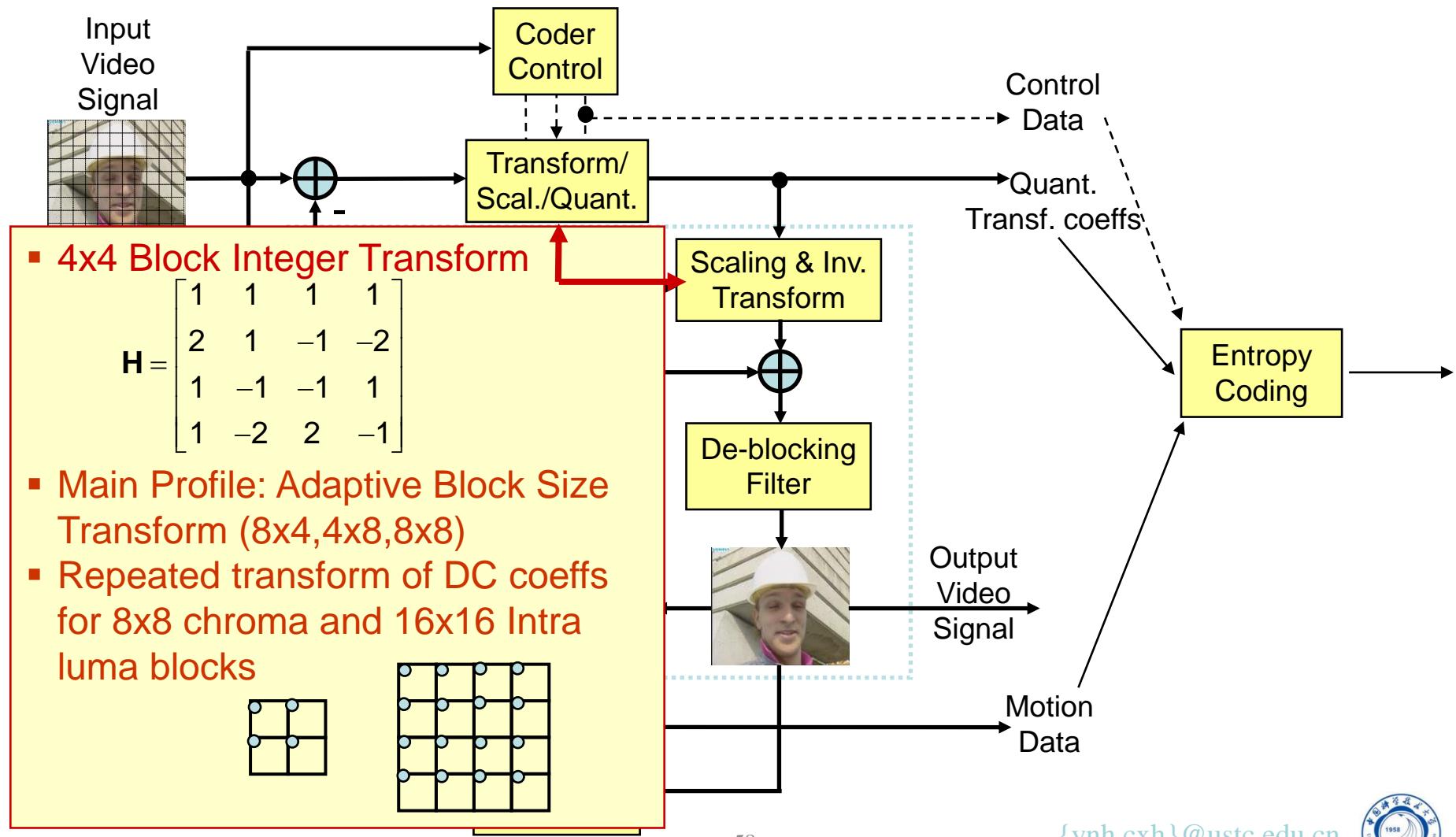
帧内预测的本质

- ◆ JPEG中每个宏块都进行编码
- ◆ 相邻宏块之间通常具有很强的相关性（空间冗余）
- ◆ 后编码的宏块若和已编码过的宏块相似，则只需要记录预测值和实际值之间的差值
- ◆ 对宏块的编码从左上角开始

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				



变换编码 (Transform Coding)



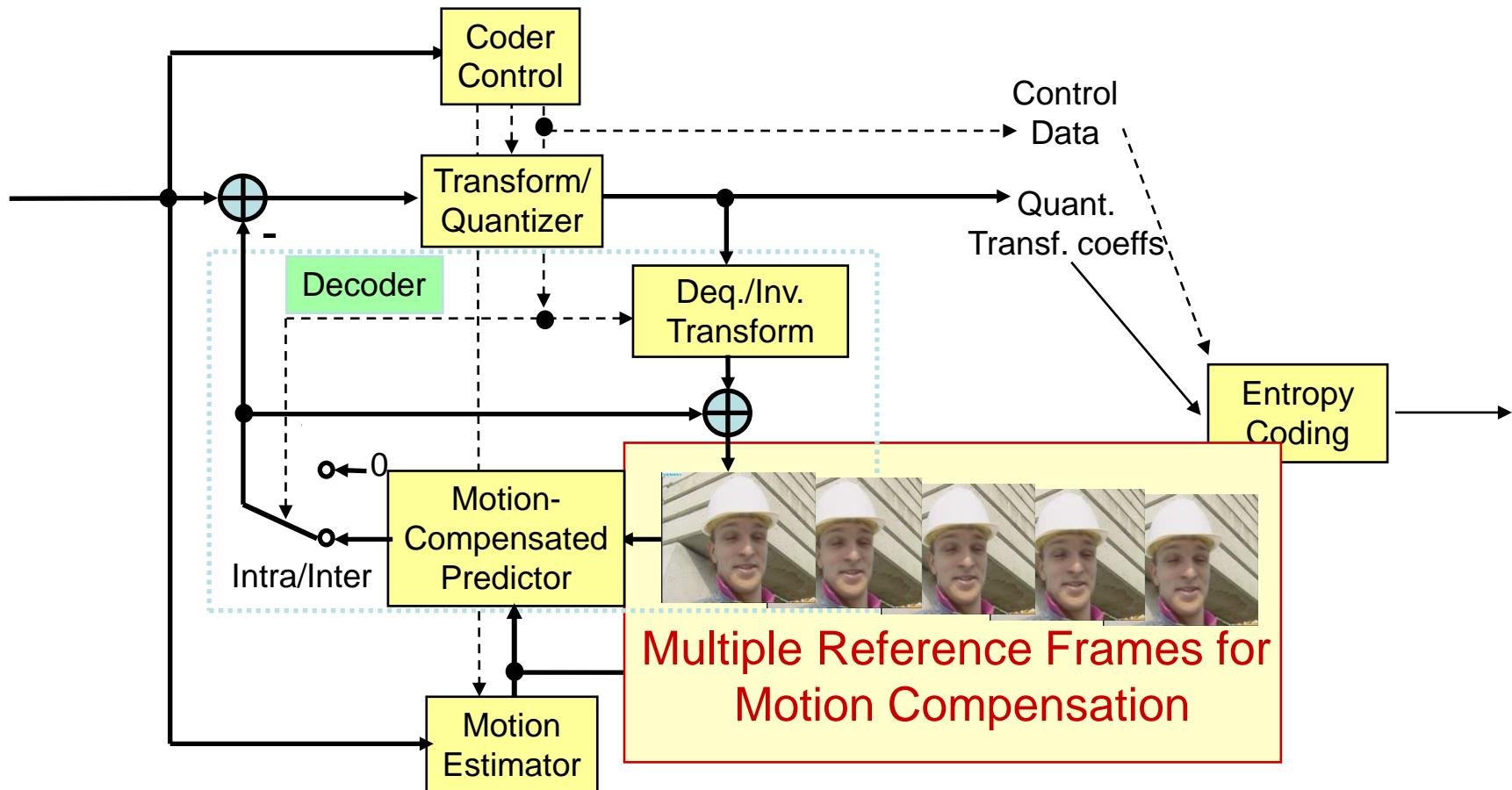


整数变换和量化

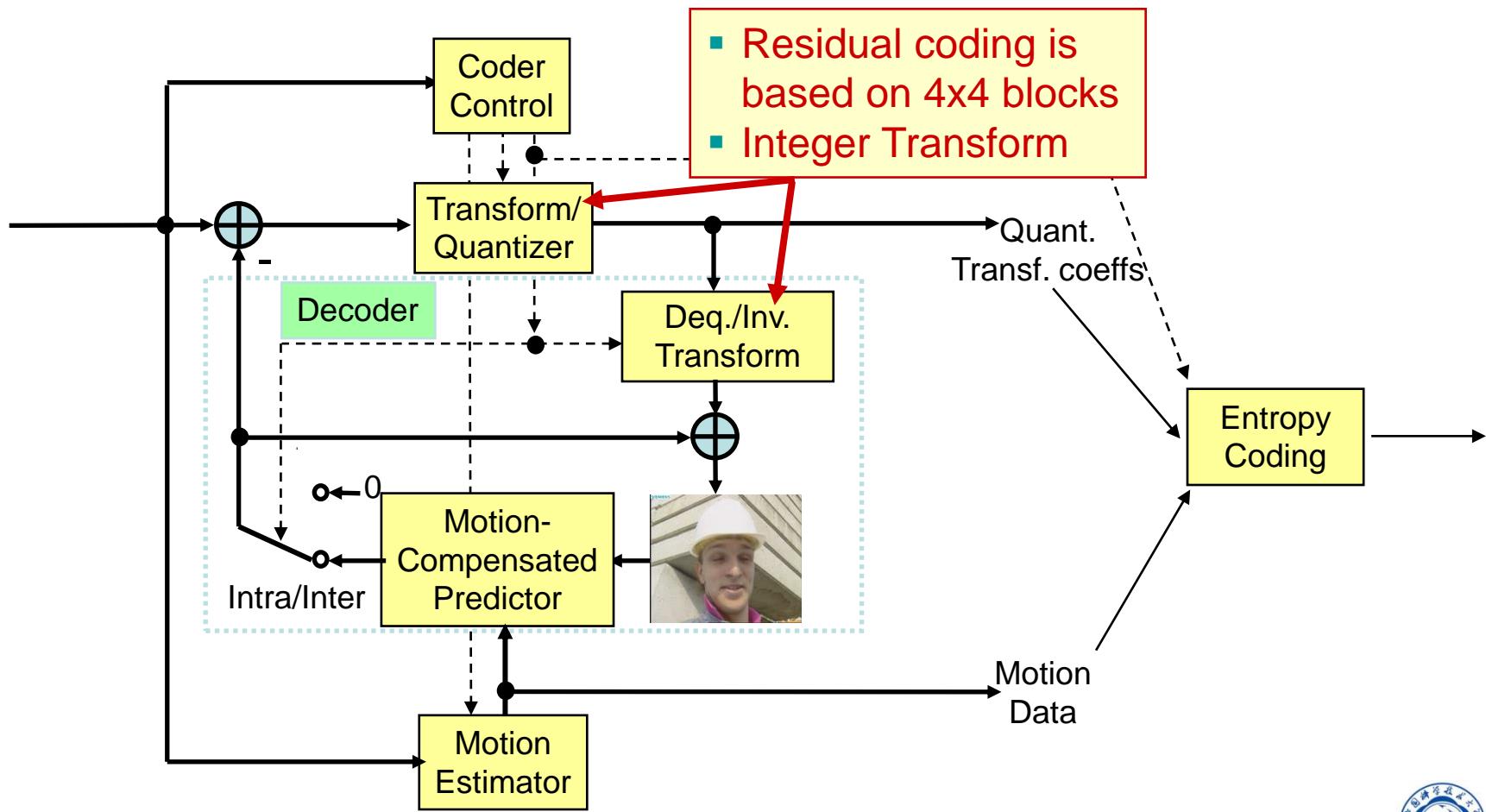
- ◆ 将**变换**和**变换后系数的量化**一体化设计
- ◆ Transform
 - H.264/MPEG-4 AVC使用整数DCT 4×4 变换；而MPEG-2和MPEG-4使用浮点DCT 8×8 变换。更小块减少了块效应和明显的人工痕迹。
- ◆ Quantization
 - 量化使用了52级步长的量化器，而H.263标准只有31级。量化步长以12.5%递增，量化步长范围的扩大使得编码器能够更灵活、精确地进行控制，在比特率和图像质量之间达到折中。



多参考帧



残差编码

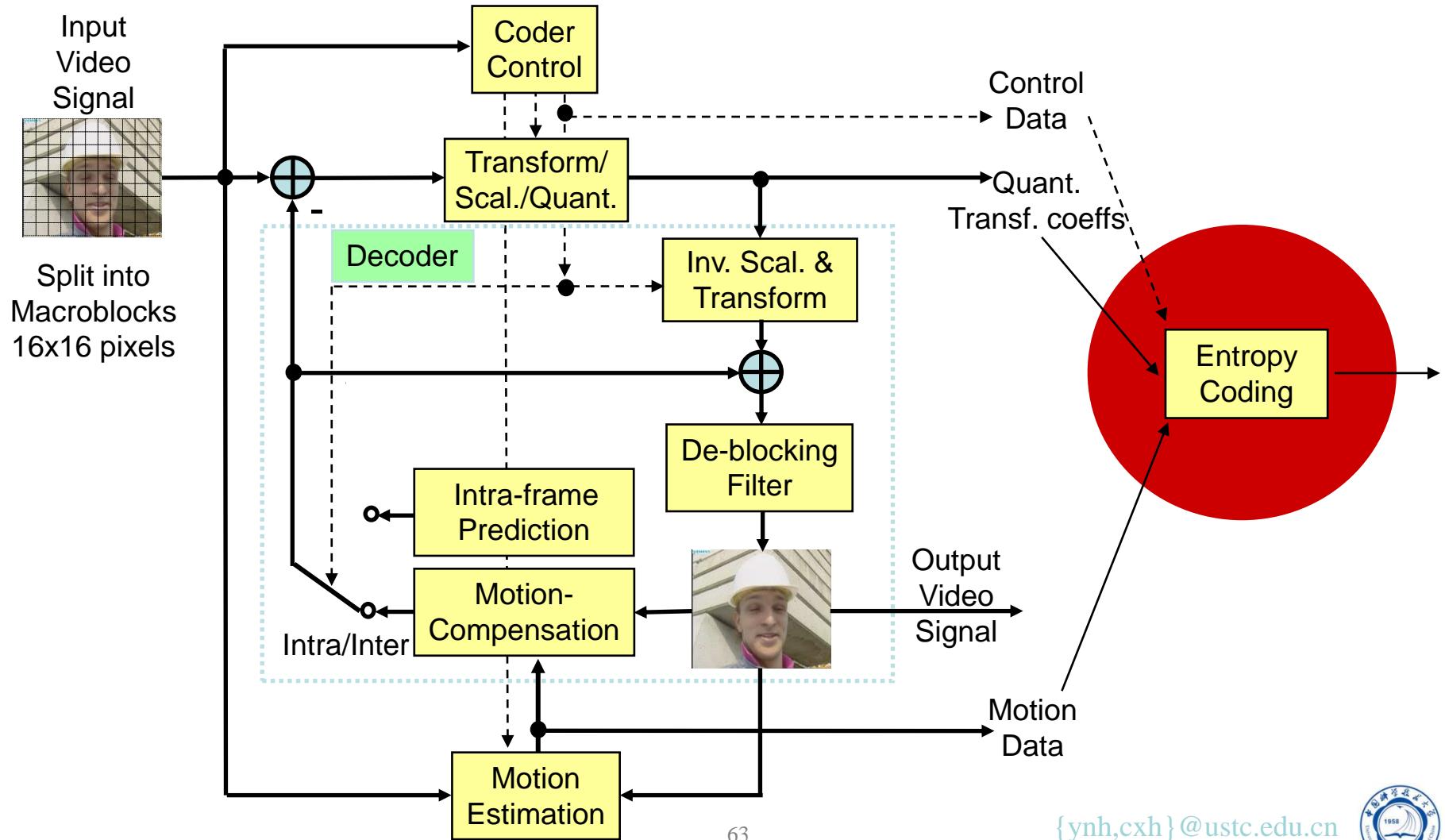


去块效应滤波器(Deblocking filter)

- ◆ 滤波的目的是消除由于相邻宏块有不同的运动估计类型（比如运动估计和内部估计）或者不同的量化参数导致的人工痕迹。在块边界这种情况下，过滤的目的是消除可能由于变换/量化和来自于相邻块运动矢量的差别引起的人工痕迹。环路滤波通过一个内容自适应的非线性算法修改在宏块/块边界的同一边的两个像素。



熵编码





熵编码的改进

H.264/AVC推荐使用3种熵编码器

- ◆ VLC(Variable Length Coding)
 - Huffman Coding → **Golomb Coding**
- ◆ CA VLC
 - Context-Adaptive Variable Length Coding
 - 上下文自适应变长编码
- ◆ CABAC
 - Context-Adaptive Binary Arithmetic Coding
 - 上下文自适应二值算数编码





上下文自适应 (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根据相临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



Value Quotient Remainder Code

0	0	0	1 00
1	0	1	1 01
2	0	2	1 10
3	0	3	1 11
4	1	0	0 1 00
5	1	1	0 1 01
6	1	2	0 1 10
7	1	3	0 1 11
8	2	0	00 1 00
9	2	1	00 1 01
10	2	2	00 1 10
11	2	3	00 1 11
12	3	0	000 1 00
13	3	1	000 1 01
14	3	2	000 1 10
15	3	3	000 1 11

Golomb 编码

http://urchin.earth.li/~twic/Golomb-Rice_Coding.html

◆ 规则：“商+1+余数”

◆ 首先将需要编码的正数除以给定参数m（例中m=4）

◆ 商用一元编码(Unary code)

◆ 商表示0的个数

◆ 余数用固定长度的二进制编码

◆ 余数需要的位长由参数m决定
(m=4对应长度2bit)

◆ 商和余数之间插入1作为停止位

◆ 参数m为2的整数次幂时称为
Golomb-Rice 编码





指数Golomb码

参数m与待编码数值相关，呈指数变化

Value	Value + 1 in Binary	Exponent	Remainder	Code
0	1	0	0	1
1	10	1	0	010
2	11	1	1	011
3	100	2	0	00100
4	101	2	1	00101
5	110	2	2	00110
6	111	2	3	00111
7	1000	3	0	0001000
8	1001	3	1	0001001
9	1010	3	2	0001010
10	1011	3	3	0001011
11	1100	3	4	0001100
12	1101	3	5	0001101
13	1110	3	6	0001110
14	1111	3	7	0001111
15	10000	4	0	000010000

◆ 规则：“指数+1+(余数+1)”

◆ 首先将需要编码的正数+1后表示为给定参数（例如base=2）的幂形式：如

$$\begin{aligned}0 &\rightarrow 1 = 2^0 + 1, \\1 &\rightarrow 2 = 2^1 + 0, \\2 &\rightarrow 3 = 2^1 + 1\end{aligned}$$

◆ 指数用来表示0的个数

◆ “余数+1”表示为二进制数

◆ 指数和余数之间插入1作为停止位





H.264 编码效果

PSNR(Peak signal-to-noise ratio)

$$\text{SNR(dB)} = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right) = 20 \log_{10} \left(\frac{A_{\text{signal}}}{A_{\text{noise}}} \right)$$

通常用峰值信噪比来衡量编码后图像或视频的质量

$$SNR_{dB} = (6.02N + 1.76)dB$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right)$$

MAX_I 表示图像像素的最大值

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2$$

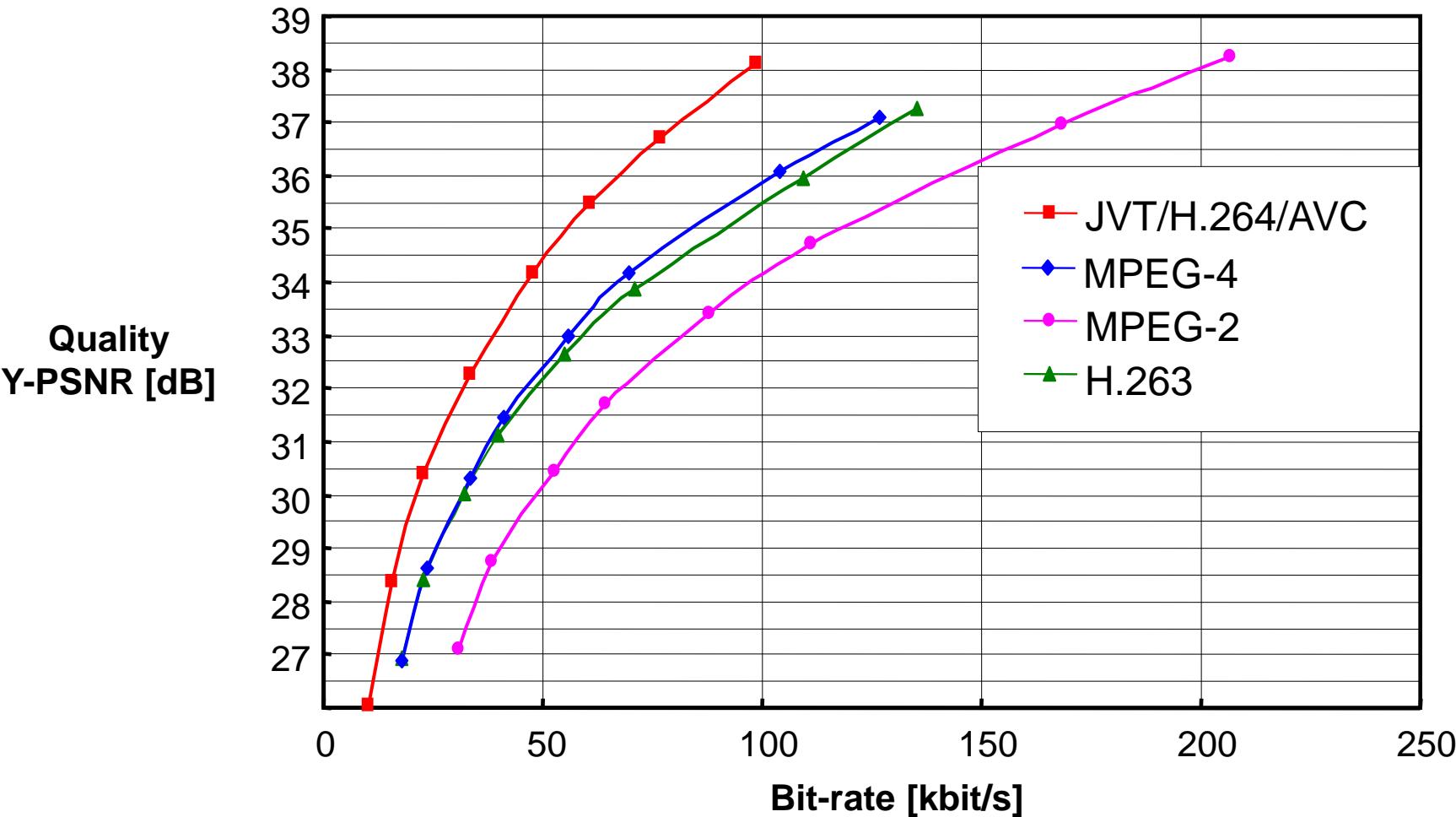




H.264 编码效果示例

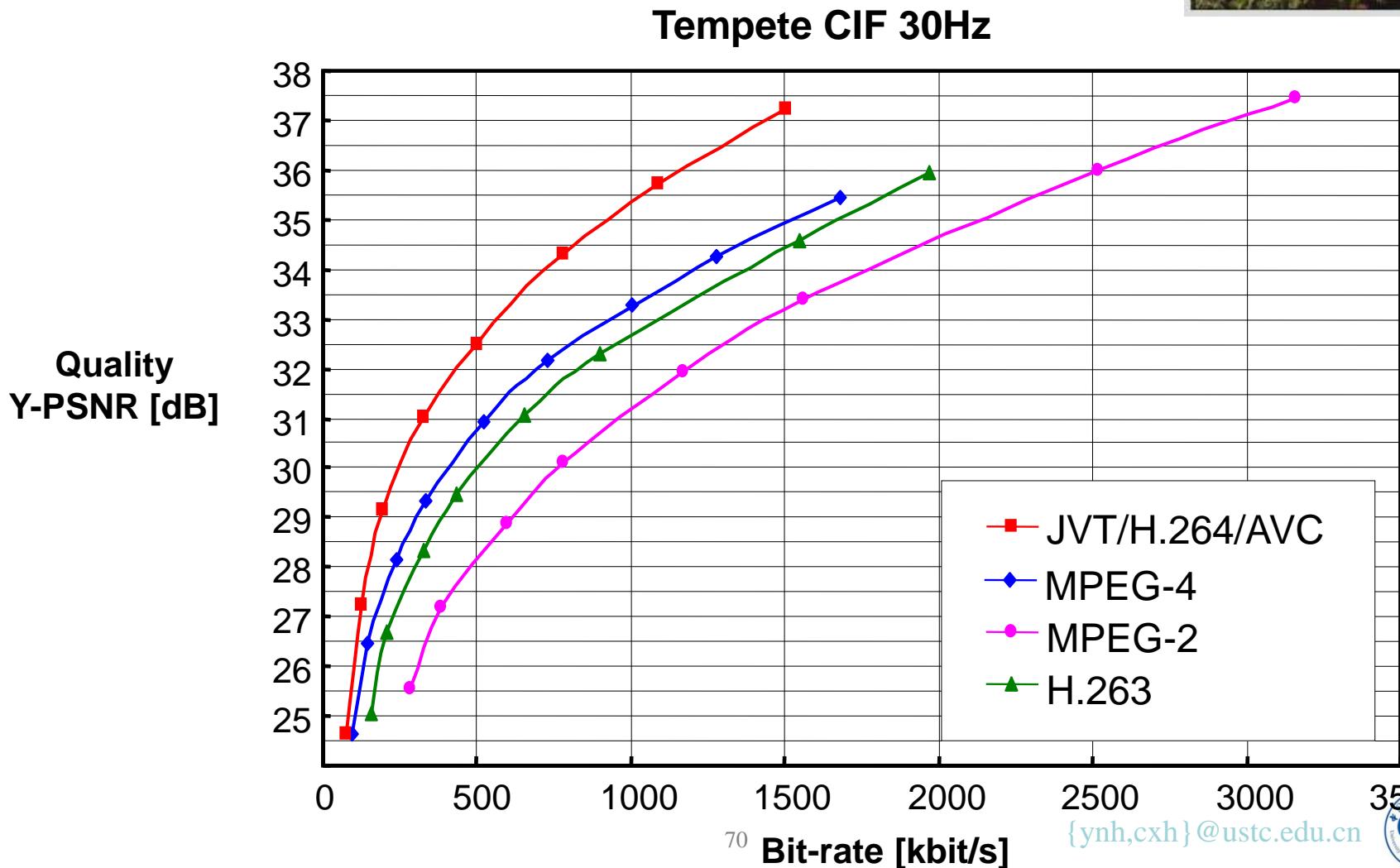


Foreman QCIF 10Hz





H.264 编码效果示例





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→(4×4整数变换 & 量化)
- 52级步长的量化器，量化步长以12.5%递增

◆ 熵编码

- CAVLC/CABAC





第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - MPEG简介
 - MPEG-1、MPEG-2视频压缩原理
 - MPEG-4视听对象编码简介
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介



◆ AVS工作组

- 数字音视频编解码技术标准工作组由国家原信息产业部科学技术司于2002年6月批准成立。工作组的任务是：面向我国的信息产业需求，联合国内企业和科研机构，制（修）订数字音视频的压缩、解压缩、处理和表示等共性技术标准，为数字音视频设备与系统提供高效经济的编解码技术，服务于高分辨率数字广播、高密度激光数字存储媒体、无线宽带多媒体通讯、互联网宽带流媒体等重大信息产业应用。

◆ AVS标准

- AVS是我国具备自主知识产权的信源编码标准。顾名思义，“信源”是信息的“源头”，信源编码技术解决的重点问题是数字音视频海量数据（即初始数据、信源）的编码压缩问题，故也称数字音视频编解码技术。





AVS技术特征

- ◆ AVS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：
 - 8×8 的整数变换与64级量化；
 - 亮度和色度帧内预测都是以 8×8 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；
 - 16×16 、 16×8 、 8×16 和 8×8 4种块模式的运动补偿；
 - 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；
 - 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。





AVS→AVS2→AVS3

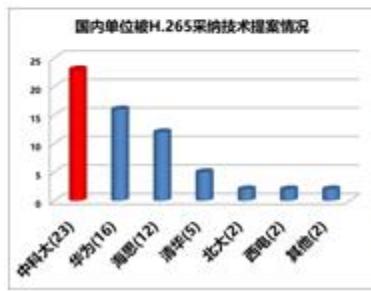
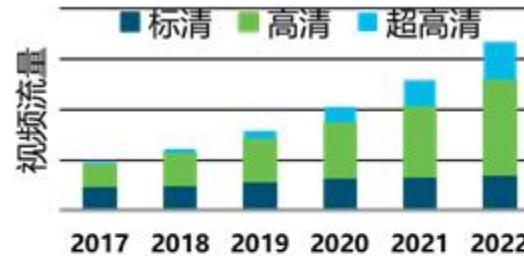
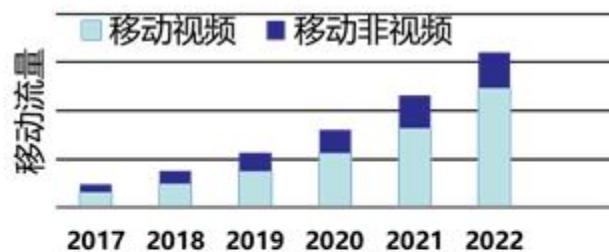
- ◆ 2006年2月，《先进音视频编码》的“视频”部分，获国家标准化管理委员会正式批准成为国家标准，并于2006年3月1日起实施。
- ◆ 2009年，我国的AVS视频标准成为国际视频标准之一。与MPEG-2、ITU-T MPEG-4 AVC（等同于ITU-T H.264）、美国微软牵头制定的行业标准SMPTE VC-1同列为国际视频标准。
- ◆ **2016年5月17日**，我国第二代数字视频编码标准AVS2颁布成为广播电视台行业标准，**AVS2**标准颁布标志着我国视频技术和产业正式进入“超高清”和“超高效”的“双超时代”。根据国家新闻出版广电总局广播电视台规划院进行的严格测试，AVS2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)。
- ◆ **2021年11月4日**，中关村视听产业技术创新联盟（AVS产业联盟）批准发布团体标准《信息技术 智能媒体编码 第2部分：视频》（标准号T/AI 109.2-2021），即日实施，该标准将替代T/AI 109.2-2020《信息技术 智能媒体编码 第2部分：视频》，即AVS3标准。





视频压缩编码的中国贡献

视频压缩技术是计算机处理视频的前提。视频信号数字化后数据量巨大，计算机很难对之进行保存和处理。压缩技术就是将数据中的冗余信息去掉。国际电联的H.264、H.265、AVS.....



华为麒麟970芯片 华为监控芯片Hi3519



中科大共有47项专利技术被国际国内标准采纳（23项专利技术被H.265国际标准采纳），**国内排名第一**。相关技术**在麒麟芯片中使用**。



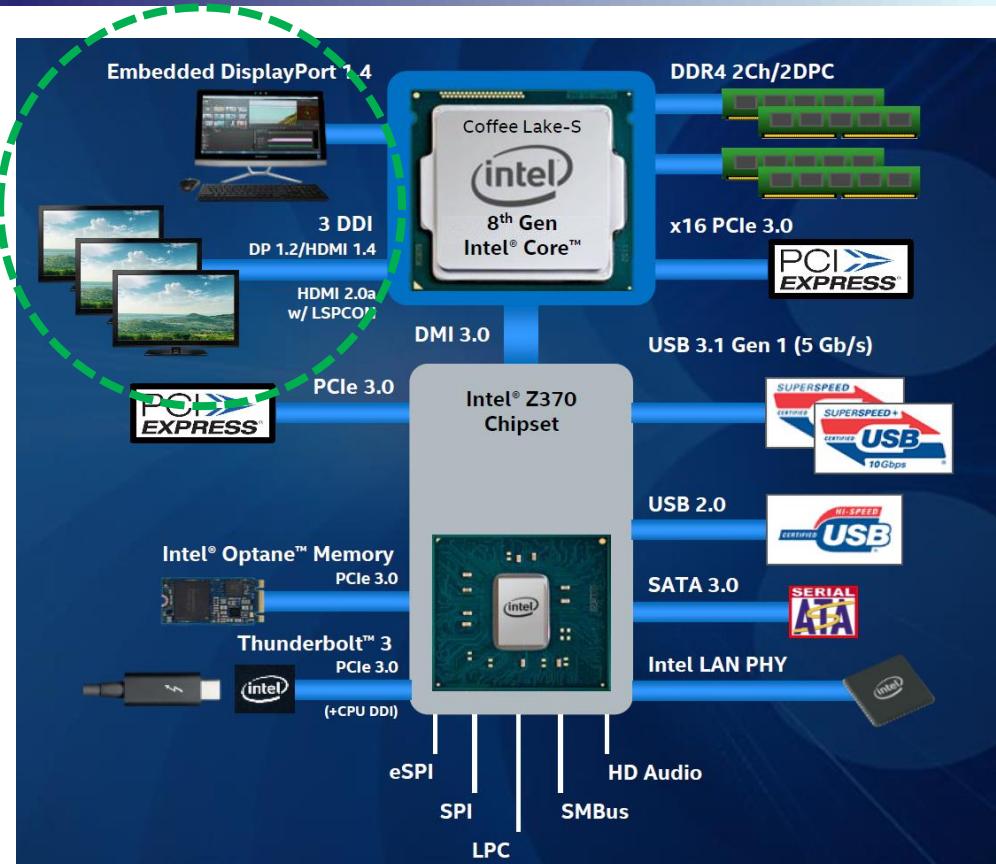
第三章 多媒体数据压缩

- ◆ § 3. 1 无损数据压缩
- ◆ § 3. 2 音频数据的压缩标准
- ◆ § 3. 3 图像数据的压缩标准
- ◆ § 3. 4 视频数据的压缩标准
 - MPEG简介
 - MPEG-1、MPEG-2视频压缩原理
 - MPEG-4视听对象编码简介
 - H.264/MPEG-4 AVC
 - AVS简介
 - H.265/MPEG-H、H.266/MPEG-I简介





HEVC(H.265)产生的背景：4K时代来临 Intel UHD Graphic 630



CPU内置的GPU核心开始支持高分辨率，并支持DisplayPort、HDMI等接口

Processor Graphics

Processor Graphics <small>?</small>	Intel® UHD Graphics 630
Graphics Base Frequency <small>?</small>	350 MHz
Graphics Max Dynamic Frequency <small>?</small>	1.20 GHz
Graphics Video Max Memory <small>?</small>	64 GB
4K Support <small>?</small>	Yes, at 60Hz
Max Resolution (HDMI 1.4) [‡] <small>?</small>	4096x2304@24Hz
Max Resolution (DP) [‡] <small>?</small>	4096x2304@60Hz
Max Resolution (eDP - Integrated Flat Panel) [‡] <small>?</small>	4096x2304@60Hz
DirectX* Support <small>?</small>	12
OpenGL* Support <small>?</small>	4.5
Intel® Quick Sync Video <small>?</small>	Yes
Intel® InTru™ 3D Technology <small>?</small>	Yes
Intel® Clear Video HD Technology <small>?</small>	Yes
Intel® Clear Video Technology <small>?</small>	Yes
# of Displays Supported [‡]	
3	



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 4×4 或 8×8 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。





H.265/HEVC是什么

- ◆ H.265/HEVC(High Efficiency Video Coding)
 - 提供的视频数据压缩率大约是H.264/AVC的2倍
 - 支持分辨率高达 8192×4320 的8K超高清电视
- ◆ H.265/HEVC是在H.264/AVC基础上的改进， 其结构和原理基本相同
- ◆ 标准开发者
 - JCT-VC(Joint Collaborative Team on Video Coding)
 - JCT-VC是ITU-T VCEG (Video Coding Experts Group)/视频编码专家组和ISO/IEC MPEG(Moving Picture Experts Group)/动态图像专家组联合组建的团队
- ◆ H.265/HEVC标准名
 - H.265
 - MPEG-H HEVC(MPEG-H Part 2, ISO/IEC 23008-2)





在H.264十年之后

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.265
(04/2013)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS
Infrastructure of audiovisual services – Coding of moving
video

High efficiency video coding

Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)
IEEE Transactions on Circuit and System for Video Technology, vol. 22, NO. 12, Dec 2012

Recommendation ITU-T H.265

HEVC Complexity and Implementation Analysis
IEEE Transactions on Circuit and System for Video Technology, vol. 22, NO. 12, [Dec 2012]@ustc.edu.cn

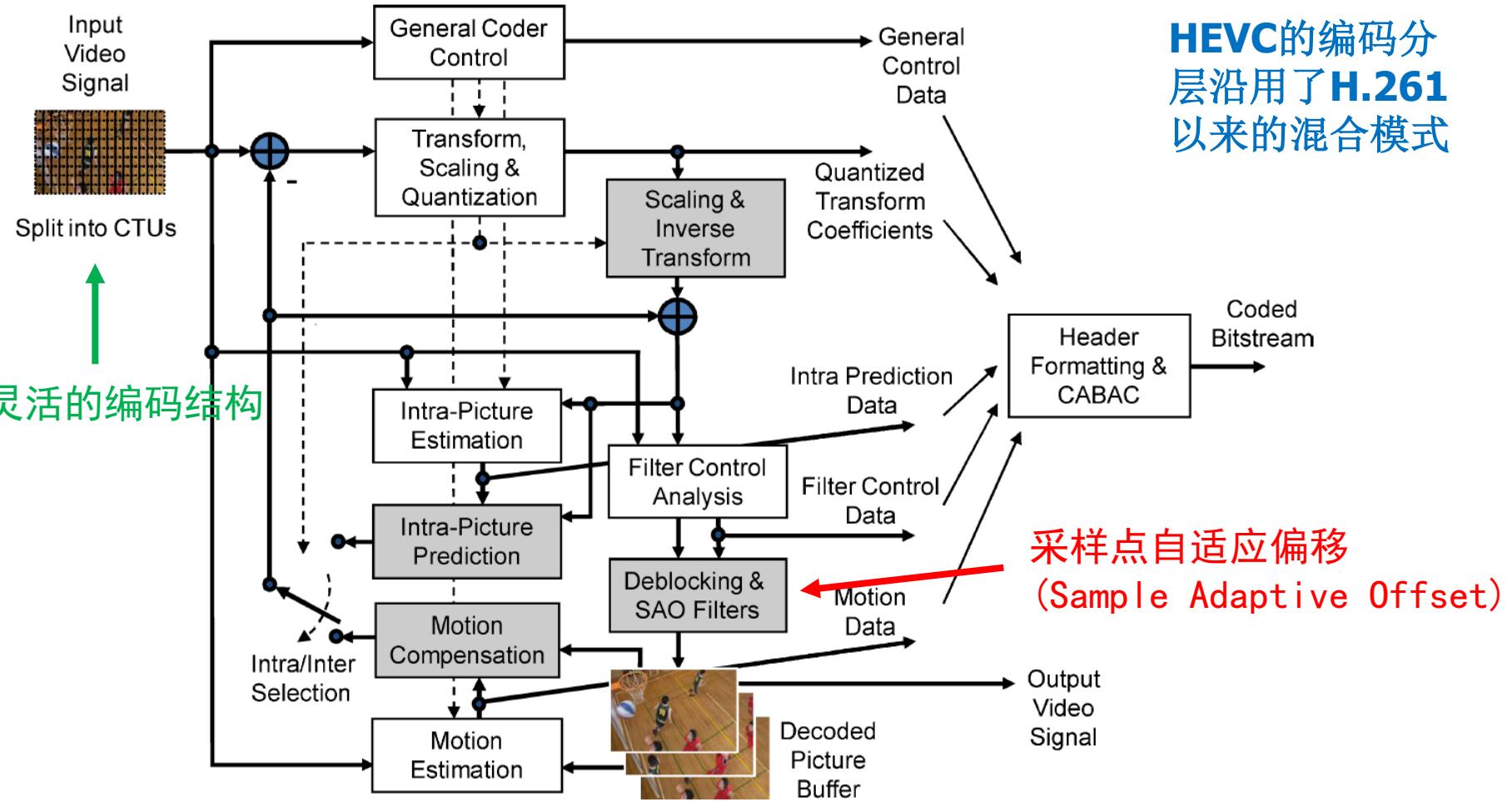
HEVC: High Efficiency Video Coding, 是一种新的视频压缩标准。可以替代H.264/ AVC编码标准。2013年1月26号正式成为国际标准。

HEVC 编码器可以达到与H.264/MPEG-4/AVC标准的编码器相同的主观重放质量，使用的比特率平均可以减少 **50%**.

总体上，HEVC解码器的复杂度与H.264/AVC解码器的相比没有太大差异.而 HEVC 的编码器复杂度则是H.264/AVC编码器的数倍之多——未来的研究方向.



HEVC(H.265)编码器框架



HEVC的编码分层沿用了H.261以来的混合模式



推荐阅读的文献

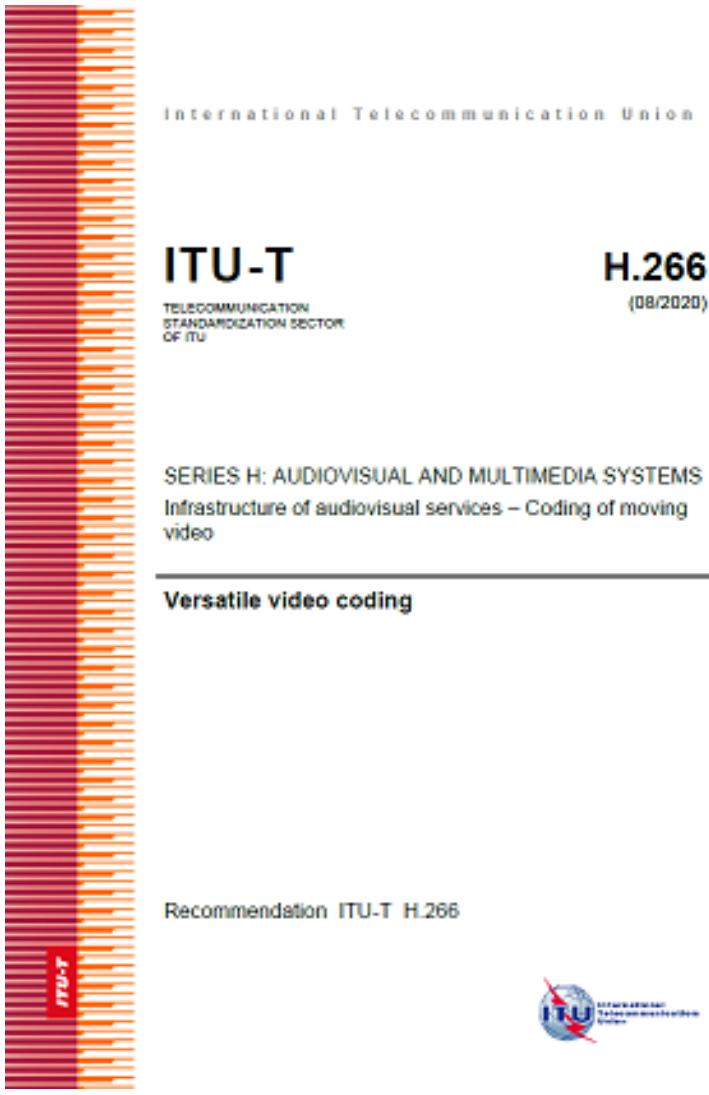
- ◆ 朱秀昌等, 新一代视频编码标准——HEVC, 《南京邮电大学学报》, 2013年6月
- ◆ IEEE Transactions on Circuit and System for Video Technology, vol. 22, N0. 12, Dec 2012
 - Overview of the High Efficiency Video Coding (HEVC) Standard
 - Sample Adaptive Offset in the HEVC Standard
 - Block Partitioning Structure in the HEVC Standard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





H.266

Versatile Video Coding (VVC)

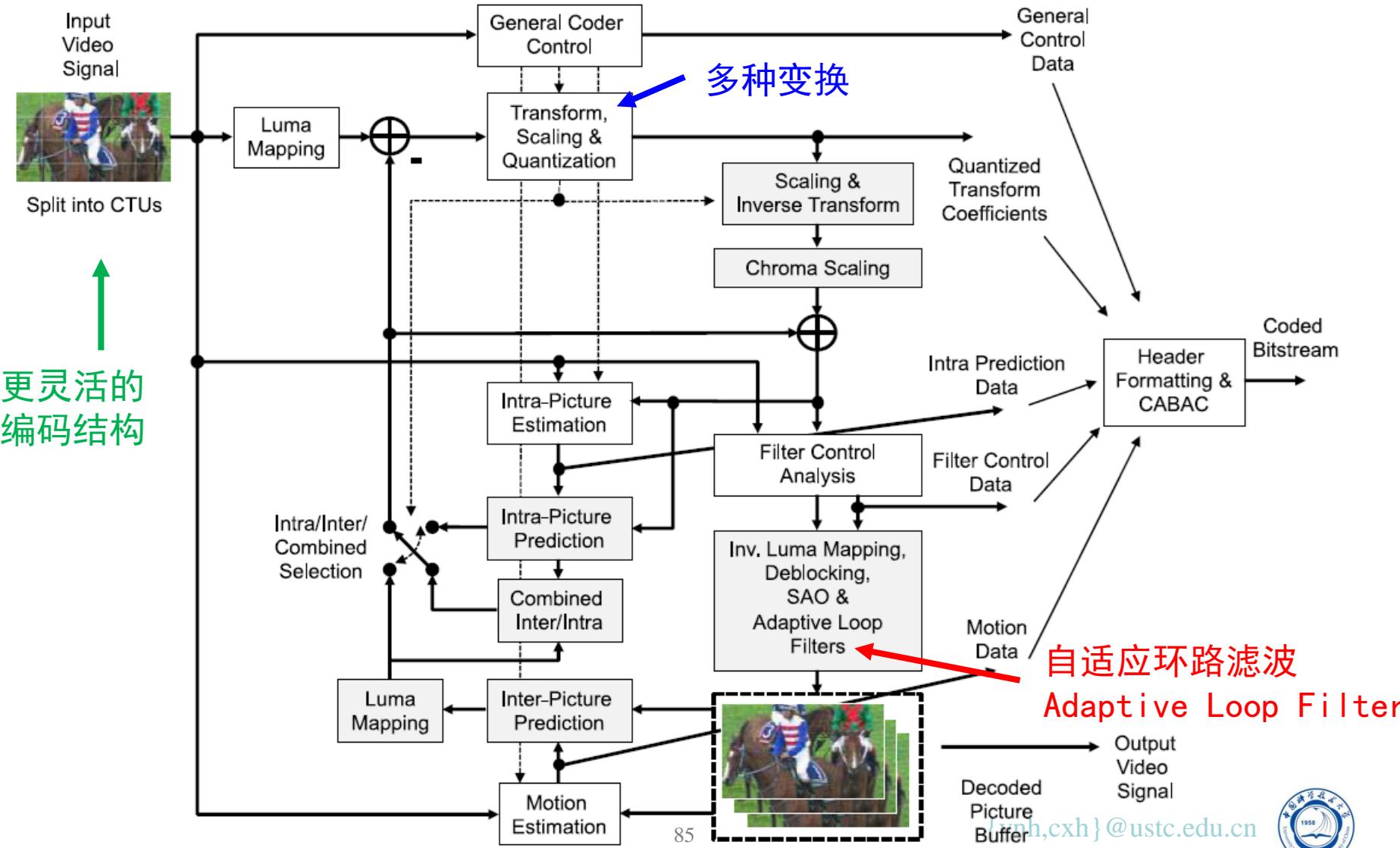


- ◆ Versatile Video Coding (VVC) was finalized in July 2020. It was developed by the Joint Video Experts Team (JVET) of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG).
- ◆ The VVC achieves significant **bit rate reductions** in the neighborhood of **50%** over its predecessor for equal video quality, the HEVC standard, and 75% over the currently most-used format, the AVC standard.



VVC(H.266)编码器框架

视频编码层的结构仍然是传统的基于块的混合视频编码模式





推荐阅读的文献

- ◆ 朱秀昌等, H.266/VVC:新一代通用视频编码国际标准,《南京邮电大学学报》, 2021年4月
- ◆ IEEE Transactions on Circuits and Systems for Video Technology (Vol: 31, Issue: 10, Oct 2021)
 - Overview of the Versatile Video Coding (VVC) Standard and its Applications
 - The High-Level Syntax of the Versatile Video Coding (VVC) Standard
 - VVC Complexity and Software Implementation Analysis
 - Overview of the Screen Content Support in VVC: Applications, Coding Tools, and Performance
 - Block Partitioning Structure in the VVC Standard
 - Intra Prediction and Mode Coding in VVC
 - Motion Vector Coding and Block Merging in the Versatile Video Coding Standard
 - Subblock-Based Motion Derivation and Inter Prediction Refinement in the Versatile Video Coding Standard
 - Transform Coding in the VVC Standard
 - Quantization and Entropy Coding in the Versatile Video Coding (VVC) Standard
 - VVC In-Loop Filters



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
 - ◆ H.264 Intra Frame
- ◆ HEVC, 2013年4月13号
 - ◆ HEVC Intra Frame
- VVC, 2020-08

◆ China AVS

- AVS, 2003
- AVS2, 2016
- AVS3





小结(1)

- ◆ 视频编解码器原理
- ◆ MPEG1/2/4
 - MPEG1/2中空间冗余和时间冗余的消除
 - MPEG4中Audio Visual Object
- ◆ H.264
 - 运动估计/帧内预测/去块效应滤波器/整数变换/熵编码.....
- ◆ H.265
 - 灵活的编码结构/采样点自适应偏移/适用于并行计算.....
- ◆ H.266
 - 更灵活的编码结构/自适应环路滤波/多种变换.....





小结(2)

- ◆ H.261/H.262/H.263/...H.264/H.265/H.266
- ◆ MPEG1/MPEG2/MPEG4
- ◆ GOP/Picture/Slice/MacroBlock/Block
- ◆ I/P/B Frame
- ◆ Progressive/Interlace/Frame mode/Field mode
- ◆ Exhaustive/three-step/2D-logarithmic/Orthogonal/conjugate Search
- ◆ MPEG-2/Profile @ Level/MP@ML/MP@HL
- ◆ Motion Compensation/Intra Prediction/Integer Transform/Deblocking filter/VLC/CAVLC/CABAC

