Research Institute for Future Media Computing
未来媒体技术与研究所
Institute of Computer Vision
计算机视觉研究所

多媒体系统导论

Fundamentals of Multimedia System

授课教师：文嘉俊
邮箱：wenjiajun@szu.edu.cn

2024年春季课程

Outline of Lecture 12&13

- Overview-概述
- MPEG-1
  - Motion Compensation in MPEG-1 –MPEG-1的运动补偿
  - Other Major Differences from H.261-与H.261的其他主要区别
- MPEG-2
  - Supporting Interlaced Video –支持隔行扫描视频
  - MPEG-2 Scalabilities –MPEG-2的可伸缩性
  - Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
- MPEG-4
  - Overview of MPEG-4-MPEG 4概述
  - Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Synthetic Object Coding in MPEG-4-MPEG-4的合成对象编码
  - MPEG-4 Object types, Profile and Levels-MPEG-4部分、规范和层次
- MPEG-7
- Experiments-实验

Overview-概述

- MPEG: Moving Pictures Experts Group, established in 1988 for the development of digital video.- MPEG创立目的
- It is appropriately recognized that proprietary interests need to be maintained within the family of MPEG standards:-要在MPEG标准族中维护自身利益
  - Accomplished by defining only a compressed bitstream that implicitly defines the decoder. .-标准仅定义一个压缩位流（解码器）
  - The compression algorithms, and thus the encoders, are completely up to the manufacturers..-压缩算法及编码器依赖于各生产商

Outline of Lecture 12&13

- Overview-概述
- MPEG-1
  - Motion Compensation in MPEG-1 –MPEG-1的运动补偿
  - Other Major Differences from H.261-与H.261的其他主要区别
- MPEG-2
  - Supporting Interlaced Video –支持隔行扫描视频
  - MPEG-2 Scalabilities –MPEG-2的可伸缩性
  - Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
- MPEG-4
  - Overview of MPEG-4-MPEG 4概述
  - Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Synthetic Object Coding in MPEG-4-MPEG-4的合成对象编码
  - MPEG-4 Object types, Profile and Levels-MPEG-4部分、规范和层次
- MPEG-7
- Experiments-实验

MPEG-1

- MPEG-1 adopts the CCIR601 digital TV format also known as SIF (Source Input Format).-MPEG-1采用数字电视格式
- MPEG-1 supports only non-interlaced video. Normally, its picture resolution is:-MPEG-1仅支持非隔行视频
  - 352 × 240 for NTSC video at 30.-NTSC制式下的分辨率
  - 352 × 288 for PAL video at 25 fps.-PAL制式下的分辨率
  - It uses 4:2:0 chroma subsampling.-采用4:2:0进行色度二次采样
- The MPEG-1 standard is also referred to as ISO/IEC 11172. It has five parts: 11172-1 Systems, 11172-2 Video, 11172-3 Audio, 11172-4 Conformance, and 11172-5 Software.-MPEG-1标准由五部分组成

MPEG-1

- Motion Compensation in MPEG-1-MPEG-1的运动补偿
  - Motion Compensation (MC) based video encoding in H.261works as follows:.-H.261基于运动补偿视频编码工作原理

a) In Motion Estimation (ME), each macroblock (MB) of the Target P-frame is assigned a best matching MB from the previously coded I or P frame - **prediction**.- 在参考帧中对目标帧的宏块进行匹配

b) **prediction error**: The difference between the MB and its matching MB, sent to DCT and its subsequent encoding steps.- 当前宏块与匹配宏块的差为预测误差

c) The prediction is from a previous frame — **forward prediction**.- 预测由前面帧得出，因此叫前向预测

MPEG-1

- Motion Compensation in MPEG-1-MPEG-1的运动补偿

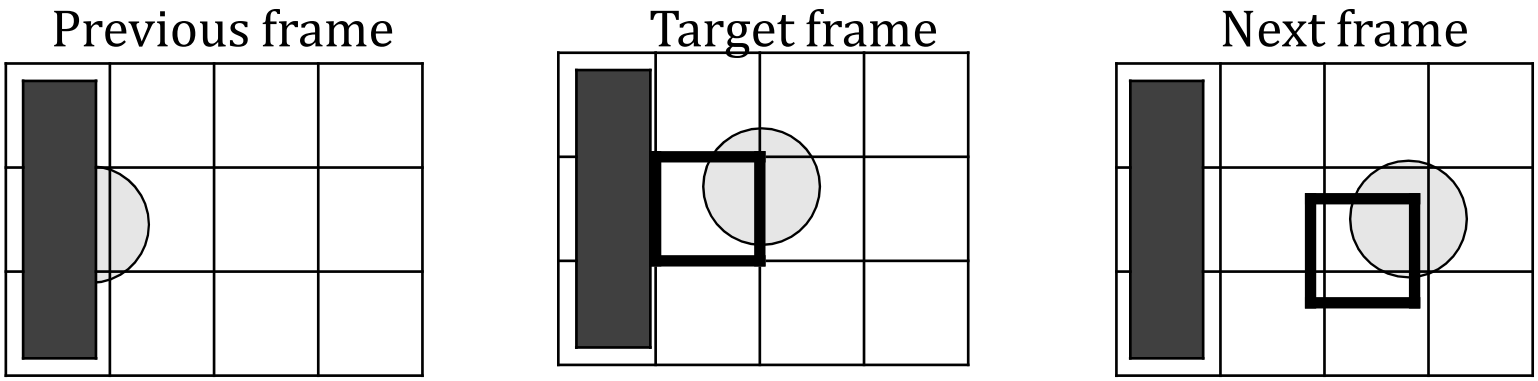


Fig 11.1:The Need for Bidirectional Search

双向搜索的需求

The MB containing part of a ball in the Target frame cannot find a good matching MB in the previous frame because half of the ball was occluded by another object. A match however can readily be obtained from the next frame.-双向搜索在前后帧同时对目标帧的宏块进行匹配

MPEG-1

- Motion Compensation in MPEG-1-MPEG-1的运动补偿
  - MPEG introduces a third frame type—B-frames, and its accompanying bi-directional motion compensation.-MPEG中引入第三类帧（B帧）及双向运动补偿
  - The MC-based B-frame coding idea is illustrated in Fig.11.2.-B帧编码原理图示

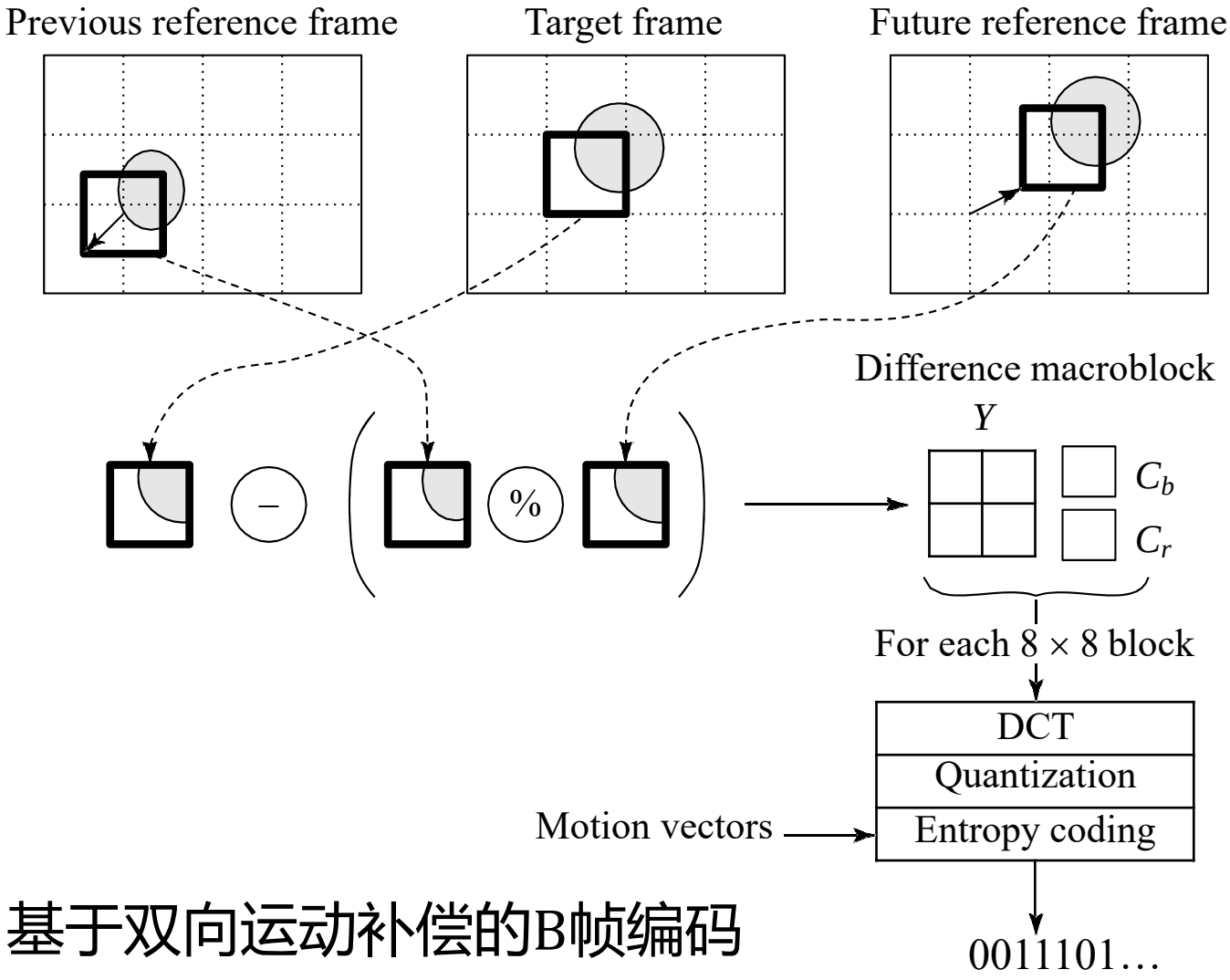
a) Each MB from a B-frame will have up to two motion vectors (MVs) (one from the forward and one from the backward prediction).-B帧的运动补偿有两个运动向量，来自前向预测和后项预测

b) If matching in both directions is successful, then two MVs will be sent and the two corresponding matching MBs are averaged (indicated by '%' in the figure) before comparing to the Target MB for generating the prediction error.-两个方向若匹配均发送，先求对应宏块均值，再求宏块误差

c) If an acceptable match can be found in only one of the reference frames, then only one MV and its corresponding MB will be used from either the forward or backward prediction.-若只有一个参考帧匹配，则采用匹配的宏块及运动向量来预测

MPEG-1

- Motion Compensation in MPEG-1-MPEG-1的运动补偿



基于双向运动补偿的B帧编码

Fig 11.2: B-frame Coding Based on Bidirectional Motion Compensation.

MPEG-1

- Motion Compensation in MPEG-1-MPEG-1的运动补偿

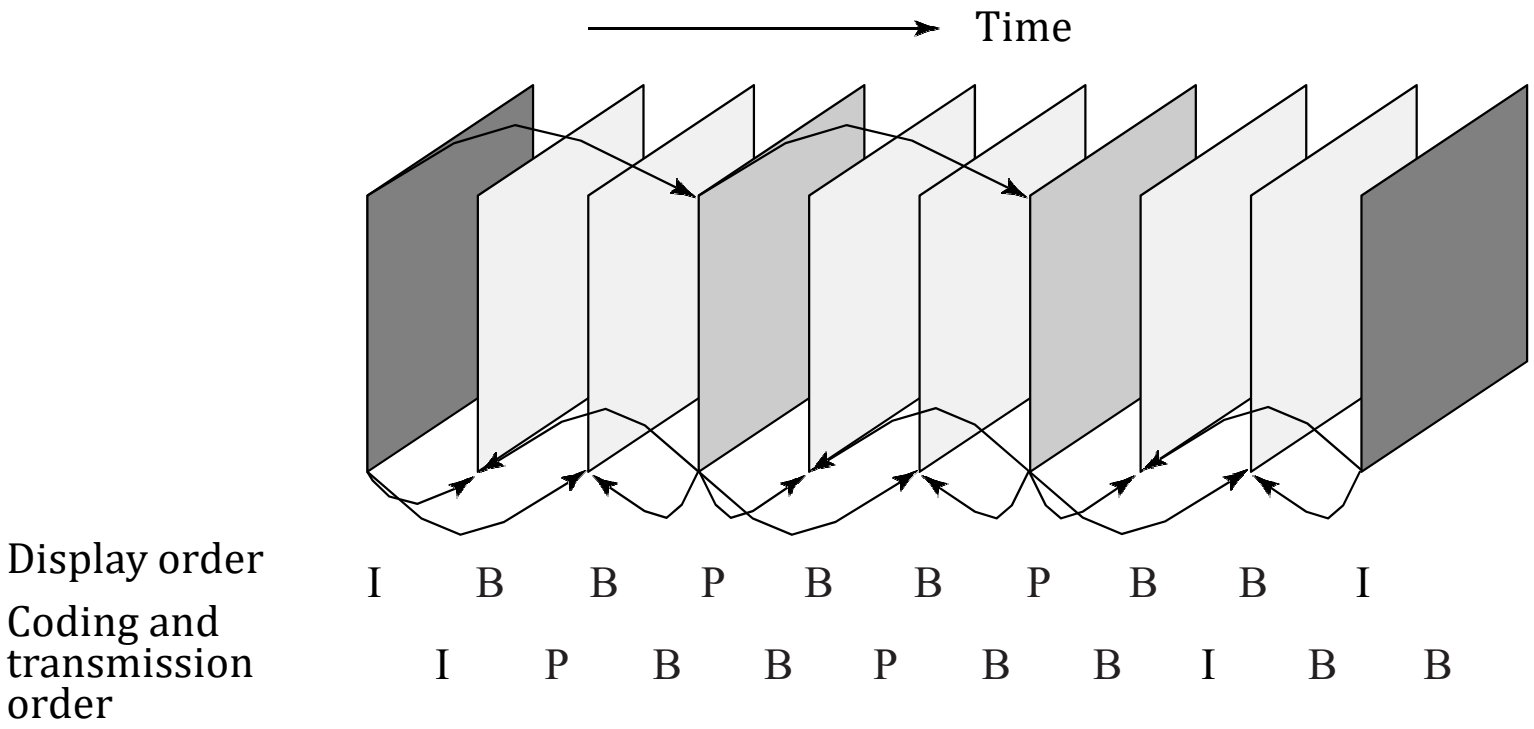


Fig 11.3: MPEG Frame Sequence.

MPEG帧序列

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别
  - Source formats supported:-MPEG1所支持的源格式

a) H.261 only supports CIF (352 × 288) and QCIF (176 × 144) source formats, MPEG-1 supports SIF (352 × 240 for NTSC, 352 × 288 for PAL)-H.261只支持CIF和QCIF两种格式

b) MPEG-1 also allows specification of other formats as long as the Constrained Parameter Set (CPS) as shown in Table 11.1 is satisfied:-MPEG-1要满足表11.1中的约束参数集

Table 11.1: The MPEG-1 Constrained Parameter Set
MPEG-1的约束参数集

Parameter	Value
Horizontal size of picture	≤ 768
Vertical size of picture	≤ 576
No. of MBs / picture	≤ 396
No. of MBs / second	≤ 9,900
Frame rate	≤ 30 fps
Bit-rate	≤ 1, 856 kbps

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别
  - Instead of GOBs as in H.261, an MPEG-1 picture can be divided into one or more slices (Fig.11.4):-MPEG-1图片可分为一个或多个宏块

  a) May contain variable numbers of macroblocks in a single picture.-MPEG-1图片可能包含可变数目的宏块

  b) May also start and end anywhere as long as they fill the whole picture.-MPEG-1宏块可开始或结束于任何位置

  c) Each slice is coded independently — additional flexibility in bit-rate control.-每个宏块单独编码，码率控制具有灵活性

  d) Slice concept is important for error recovery.-宏块片内容对错误率恢复很重要

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别

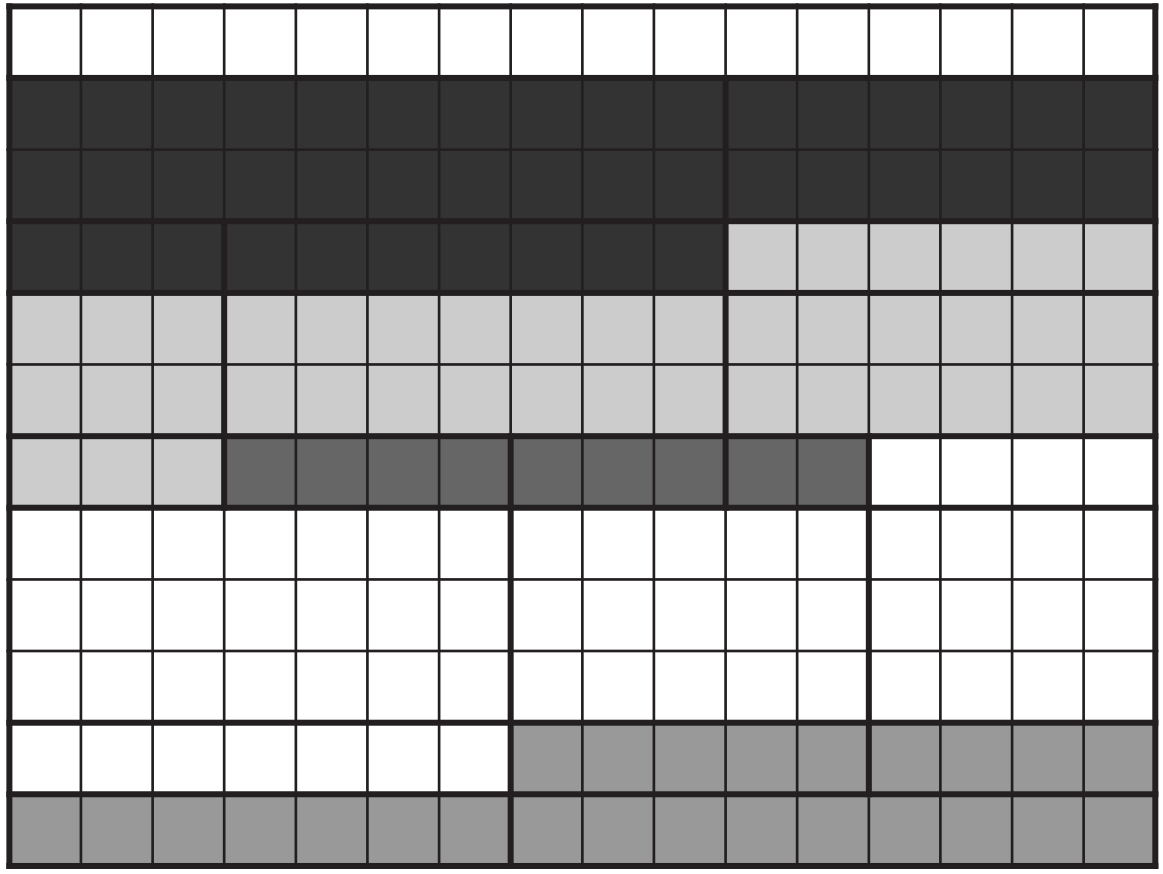


Fig 11.4: Slices in an MPEG-1 Picture.

MPEG-1图片中的宏块片

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别
  - Quantization:-MPEG-1的量化

a) MPEG-1 quantization uses different quantization tables for its Intra and Inter coding (Table 11.2 and 11.3).-不同的量化表

For DCT coefficients in Intra mode:-帧内模式下的DCT系数

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

For DCT coefficients in Inter mode,-帧间模式下的DCT系数

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

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别
  - Quantization:-MPEG-1的量化

Fig 11.2: Default Quantization Table (Q_1) for Intra-Coding

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

帧内编码默认量化表

Fig 11.3: Default Quantization Table (Q_2) for Intra-Coding

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

帧间编码默认量化表

MPEG-1

- Other Major Differences from H.261-与H.261的其他主要区别
  - Quantization:-MPEG-1的量化

a) MPEG-1 allows motion vectors to be of sub-pixel precision (1/2 pixel). The technique of "bilinear interpolation" for H.263 can be used to generate the needed values at half-pixel locations.-MPEG-1允许运动向量具有半像素精度

b) Compared to the maximum range of ±15 pixels for motion vectors in H.261, MPEG-1 supports a range of [−512, 511.5] for half-pixel precision and [−1, 024, 1, 023] for full-pixel precision motion vectors.-支持更大的运动向量范围

c) The MPEG-1 bitstream allows random access — accomplished by GOP layer in which each GOP is time coded.-MPEG-1的位流允许随机访问

MPEG-1

- Typical Sizes of MPEG-1 Frames-MPEG-1帧的大小
  - The typical size of compressed P-frames is significantly smaller than that of I-frames—because temporal redundancy is exploited in inter-frame compression.-MPEG1中压缩的P帧明显小于压缩的I帧
  - B-frames are even smaller than P-frames— because of (a) the advantage of bi-directional prediction and (b) the lowest priority given to B-frames.-B帧比P帧更小

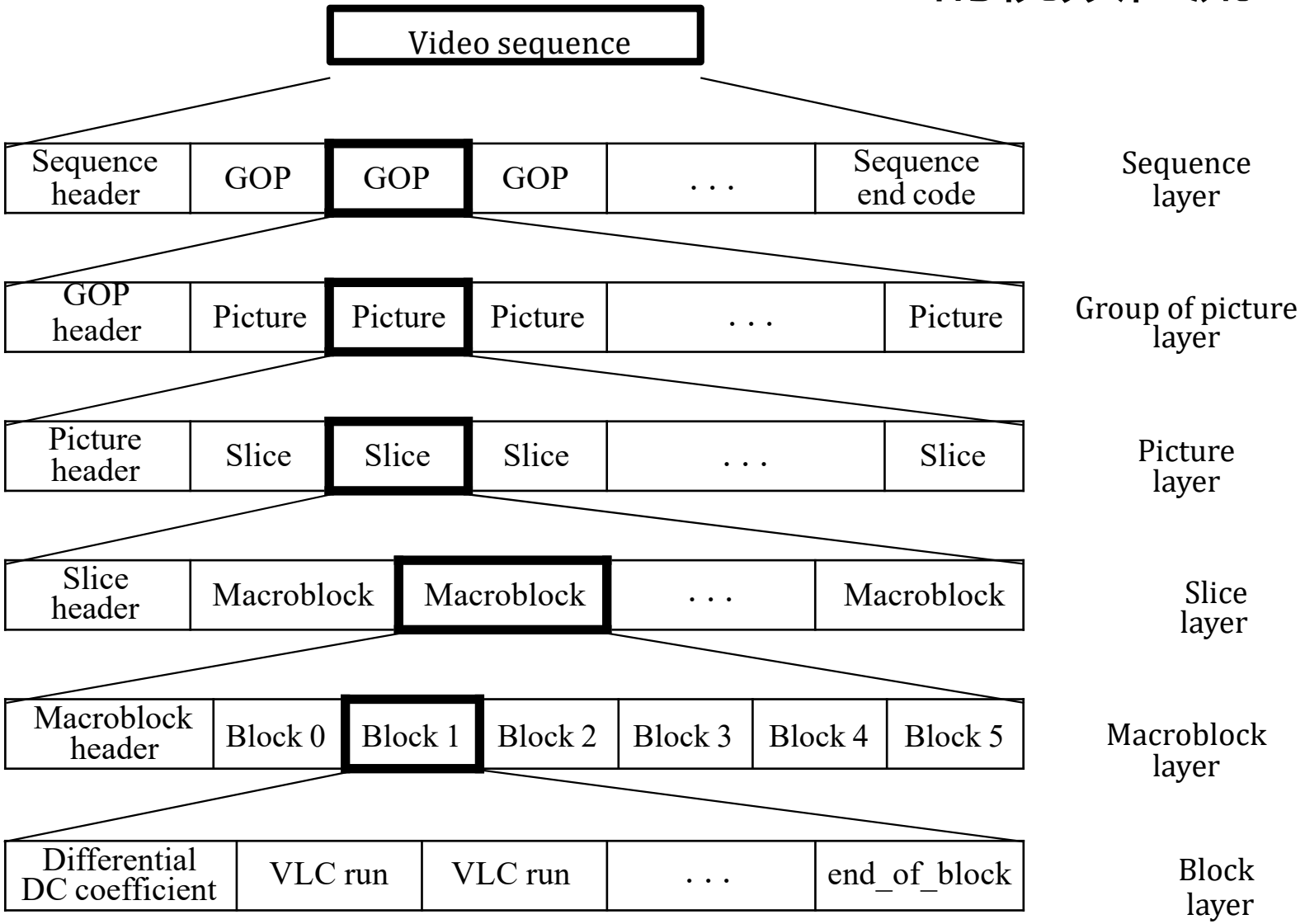
Table 11.4: Typical Compression Performance of MPEG-1 Frames

MPEG-1帧的典型压缩性能

Type	Size	Compression
I	18 kB	7:1
P	6 kB	20:1
B	2.5 kB	50:1
Avg	4.8 kB	27:1

MPEG-1

◆ MPEG-1 Video Bitstream-MPEG-1的视频位流



MPEG-1的视频位流

Fig 11.5: Layers of MPEG-1 Video Bitstream.

Outline of Lecture 12&13

- Overview-概述
- MPEG-1
  - Motion Compensation in MPEG-1 –MPEG-1的运动补偿
  - Other Major Differences from H.261-与H.261的其他主要区别
- MPEG-2
  - Supporting Interlaced Video –支持隔行扫描视频
  - MPEG-2 Scalabilities –MPEG-2的可伸缩性
  - Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
- MPEG-4
  - Overview of MPEG-4-MPEG 4概述
  - Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Synthetic Object Coding in MPEG-4-MPEG-4的合成对象编码
  - MPEG-4 Object types, Profile and Levels-MPEG-4部分、规范和层次
- MPEG-7
- Experiments-实验

MPEG-2

- MPEG-2: For higher quality video at a bit-rate of more than 4 Mbps.-MPEG-2视频质量和码率更高
- Defined seven profiles aimed at different applications-为不同应用定义了7种规格
  - Simple, Main, SNR scalable, Spatially scalable, High, 4:2:2, Multiview.-MPEG2的特点
  - Within each profile, up to four levels are defined (Table 11.5).-每种规格定义了4个等级
  - The DVD video specification allows only four display resolutions: 720 × 480, 704 × 480, 352 × 480, and 352 × 240-DVD视频仅有4种显示分辨率
  - a restricted form of the MPEG-2 Main profile at the Main and Low levels-DVD标准采用MPEG-2中Main级别和Low级别的限制形式

MPEG-2

- Defined seven profiles aimed at different applications-为不同应用定义了7种规格

Table 11.5: Profiles and Levels in MPEG-2
MPEG-2中的规格和等级

Level	Simple Profile	Main Profile	SNR Scalable Profile	Spatially Scalable Profile	High Profile	4:2:2 Profile	Multiview Profile
High		*			*		
High 1440		*		*	*		
Main	*	*	*		*	*	*
Low		*	*				

Table 11.6: Four Levels in the Main Profile of MPEG-2
MPEG-2中主要规格的4个等级

Level	Max Resolution	Max fps	Max Pixels/sec	Max coded Data Rate (Mbps)	Application
High	1,920 × 1,152	60	62.7×10^6	80	film production
High 1440	1,440 × 1,152	60	47.0×10^6	60	consumer HDTV
Main	720 × 576	30	10.4×10^6	15	studio TV
Low	352 × 288	30	3.0×10^6	4	consumer tape equiv.

MPEG-2

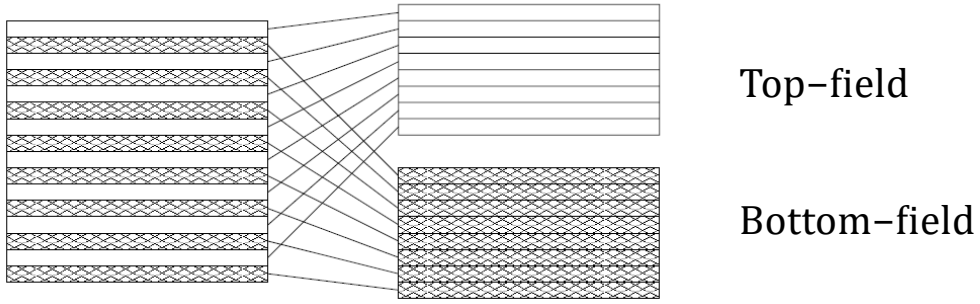
- Supporting Interlaced Video-支持隔行扫描视频
  - MPEG-2 must support interlaced video as well since this is one of the options for digital broadcast TV and HDTV.-MPEG-2支持隔行扫描视频
  - In interlaced video each frame consists of two fields, referred to as the top-field and the bottom-field.-MPEG2中每帧由两个域组成，即顶域和底域

a) In a Frame-picture, all scanlines from both fields are interleaved to form a single frame, then divided into 16 × 16 macroblocks and coded using MC.-交错成帧，宏块划分

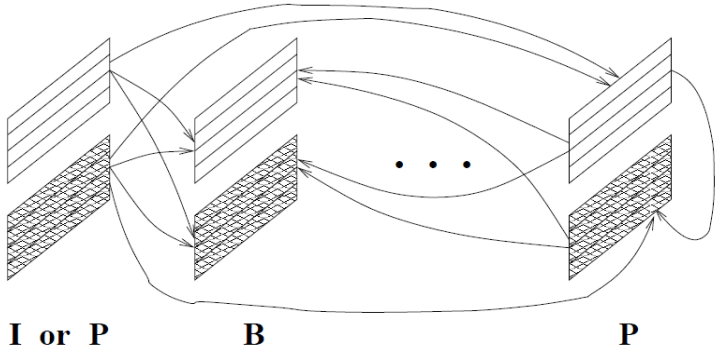
b) If each field is treated as a separate picture, then it is called Field-picture.-把每个域看成独立的图片，即域图

MPEG-2

- Supporting Interlaced Video-支持隔行扫描视频



(a) Frame–picture vs. Field–pictures



(b) Field Prediction for Field–pictures

Fig 11.6: Field pictures and Field-prediction for Field-pictures in MPEG-2.

MPEG-2中的域图和域图的域预测

MPEG-2

- Five Modes of Predictions-支持隔行扫描视频
  - MPEG-2 defines **Frame Prediction** and **Field Prediction** as well as five prediction modes:-MPEG2定义帧预测和域预测，以及五种预测模式

a) **Frame Prediction for Frame-pictures**: Identical to MPEG-1 MC-based prediction methods in both P-frames and B- frames.-与MPEG-1中P帧和B帧的预测方法相同

b) **Field Prediction for Field-pictures**: A macroblock size of 16 × 16 from Field-pictures is used. For details, see Fig. 11.6(b).-域图中16 × 16 大小的宏块将被使用

MPEG-2

- Five Modes of Predictions-支持隔行扫描视频

  c) **Field Prediction for Frame-pictures**: The top-field and bottom-field of a Frame-picture are treated separately. Each 16 × 16 macroblock (MB) from the target Frame- picture is split into two 16 × 8 parts, each coming from one field. Field prediction is carried out for these 16 × 8 parts in a manner similar to that shown in Fig. 11.6(b).-顶域与底域分开处理

  d) **16 × 8 MC for Field-pictures**: Each 16 × 16 macroblock (MB) from the target Field-picture is split into top and bottom 16 × 8 halves. Field prediction is performed on each half. This generates two motion vectors for each 16 × 16 MB in the P-Field-picture, and up to four motion vectors for each MB in the B-Field-picture. -把16 × 16宏块分为顶部和底部，单独预测

  This mode is good for a finer MC when motion is rapid and irregular.-对速度快，无规律的目标补偿效果好

MPEG-2

- Five Modes of Predictions-支持隔行扫描视频

e) Dual-Prime for P-pictures: First, Field prediction from each previous field with the same parity (top or bottom) is made. Each motion vector mv is then used to derive a calculated motion vector cv in the field with the opposite parity taking into account the temporal scaling and vertical shift between lines in the top and bottom fields. For each MB the pair mv and cv yields two preliminary predictions. Their prediction errors are averaged and used as the final prediction error. .-P图的对偶素数预测模式

This mode mimics B-picture prediction for P-pictures without adopting backward prediction (and hence with less encoding delay).-利用不使用后向预测的P图来模仿B图预测

This is the only mode that can be used for either Frame-pictures or Field-pictures.-既可用于帧图，又可用于域图

MPEG-2

- Alternate Scan and Field_DCT-交替扫描和域离散余弦变换
  - Techniques aimed at improving the effectiveness of DCT on prediction errors, only applicable to Frame-pictures in interlaced videos:-域DCT是为了提高DCT在预测误差上的有效性

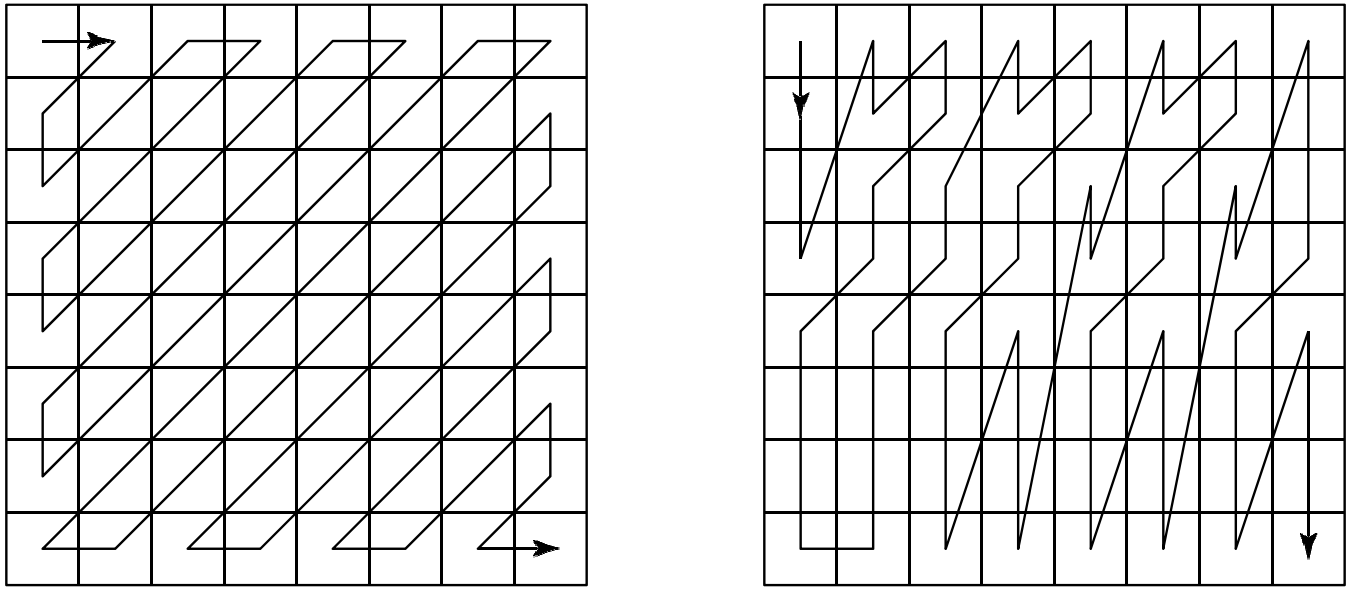
  a) Due to the nature of interlaced video the consecutive rows in the 8×8 blocks are from different fields, there exists less correlation between them than between the alternate rows.-由于隔行视频的特性，连续行的相关性比隔行之间的相关性要小

  b) Alternate scan recognizes the fact that in interlaced video the vertically higher spatial frequency components may have larger magnitudes and thus allows them to be scanned earlier in the sequence.-先对具有较高垂直空间频率的分量进行扫描

  - In MPEG-2, Field_DCT can also be used to address the same issue.-域DCT同样适用于MPEG-2

MPEG-2

- Alternate Scan and Field_DCT-交替扫描和域离散余弦变换



(a) (b)

Fig 11.7: Zigzag and Alternate Scans of DCT Coefficients for Progressive and Interlaced Videos in MPEG-2.

MPEG-2中视频的DCT系数的Z字（逐行）扫描和交替（隔行）扫描

MPEG-2

- MPEG-2 Scalabilities-MPEG-2的可伸缩性
  - The MPEG-2 **scalable coding**: A base layer and one or more enhancement layers can be defined — also known as **layered coding**.-MPEG-2的层次编码

  a) The base layer can be independently encoded, transmitted and decoded to obtain basic video quality.-基层独立编码、传输和解码，以获得基本视频质量

  b) The encoding and decoding of the enhancement layer is dependent on the base layer or the previous enhancement layer.-增强层编解码与基层及之前增强层的关系

  - Scalable coding is especially useful for MPEG-2 video transmitted over networks with following characteristics:-可伸缩编码的适用性

  a) Networks with very different bit-rates.-不同码率

  b) Networks with variable bit rate (VBR) channels.-可变码率信道

  c) Networks with noisy connections.-低质量的连接

MPEG-2

- MPEG-2 Scalabilities-MPEG-2的可伸缩性
  - MPEG-2 supports the following scalabilities:-MPEG-2支持以下可伸缩性

a) SNR Scalability — enhancement layer provides higher SNR.-SNR可伸缩性

b) Spatial Scalability — enhancement layer provides higher spatial resolution-空间可伸缩性

c) Temporal Scalability — enhancement layer facilitates higher frame rate.-时间可伸缩性

d) Hybrid Scalability — combination of any two of the above three scalabilities.-混合可伸缩性

e) Data Partitioning — quantized DCT coefficients are split into partitions.-数据划分

MPEG-2

- SNR Scalability-MPEG-2的SNR可伸缩性
  - SNR scalability: Refers to the enhancement /refinement over the base layer to improve the Signal-Noise-Ratio (SNR):-对基层进行增强/精华以改进信噪比(SNR)
  - The MPEG-2 SNR scalable encoder will generate output bit-streams Bits_base and Bits_enhance at two layers:-SNR可伸缩性编码器在两个层上输出位流

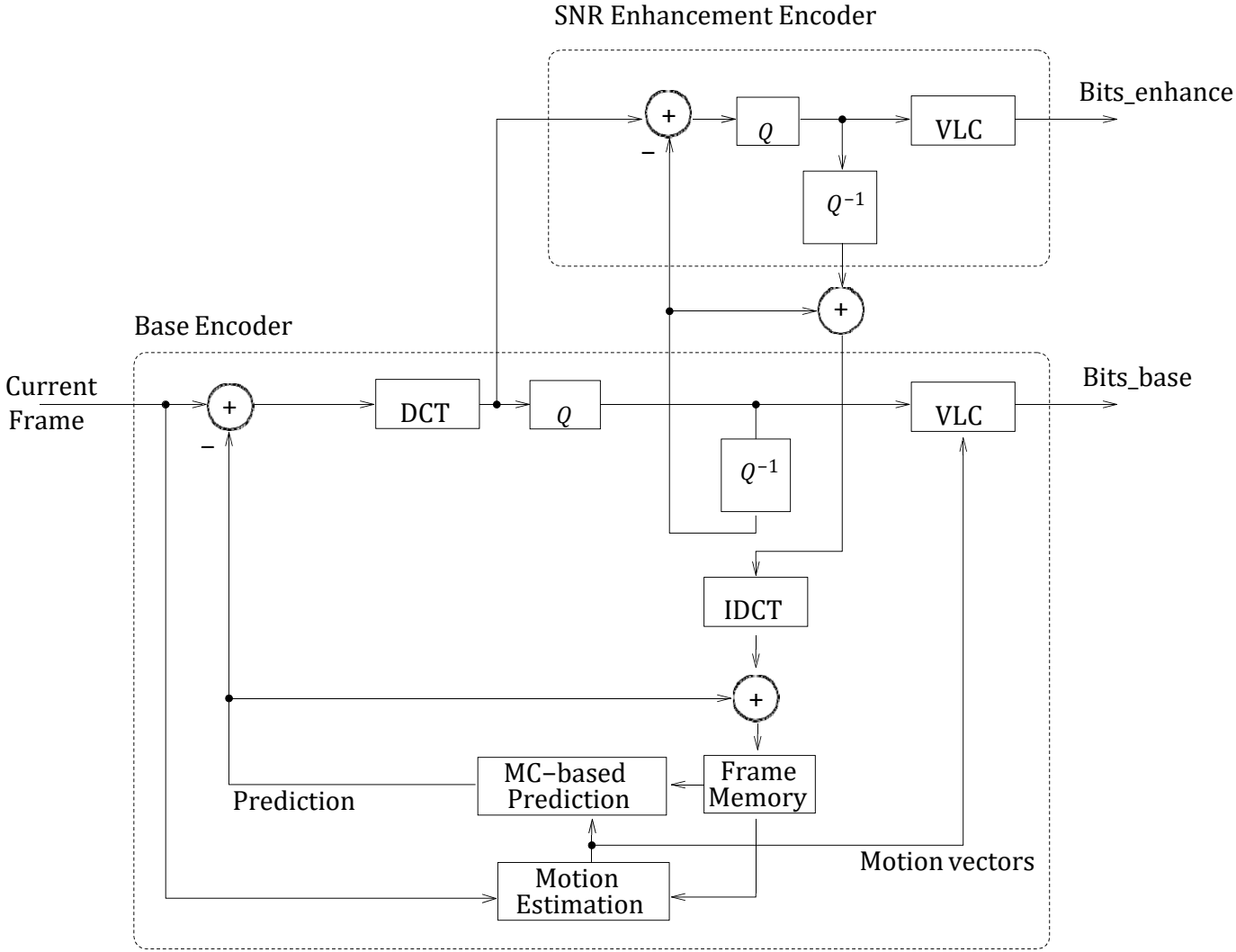
a) At the Base Layer, a coarse quantization of the DCT coefficients is employed which results in fewer bits and a relatively low quality video.-在基层对DCT系数进行粗略量化

b) The coarsely quantized DCT coefficients are then inversely quantized (Q−1) and fed to the Enhancement Layer to be compared with the original DCT coefficient.-对量化后的DCT系数进行逆量化反馈给增强层

c) Their difference is finely quantized to generate a DCT coefficient refinement, which, after VLC, becomes the bitstream called Bits_enhance.-对以上两种DCT系数的差进行量化、编码，产生位流

MPEG-2

◆ MPEG-2 Scalabilities-MPEG-2的可伸缩性



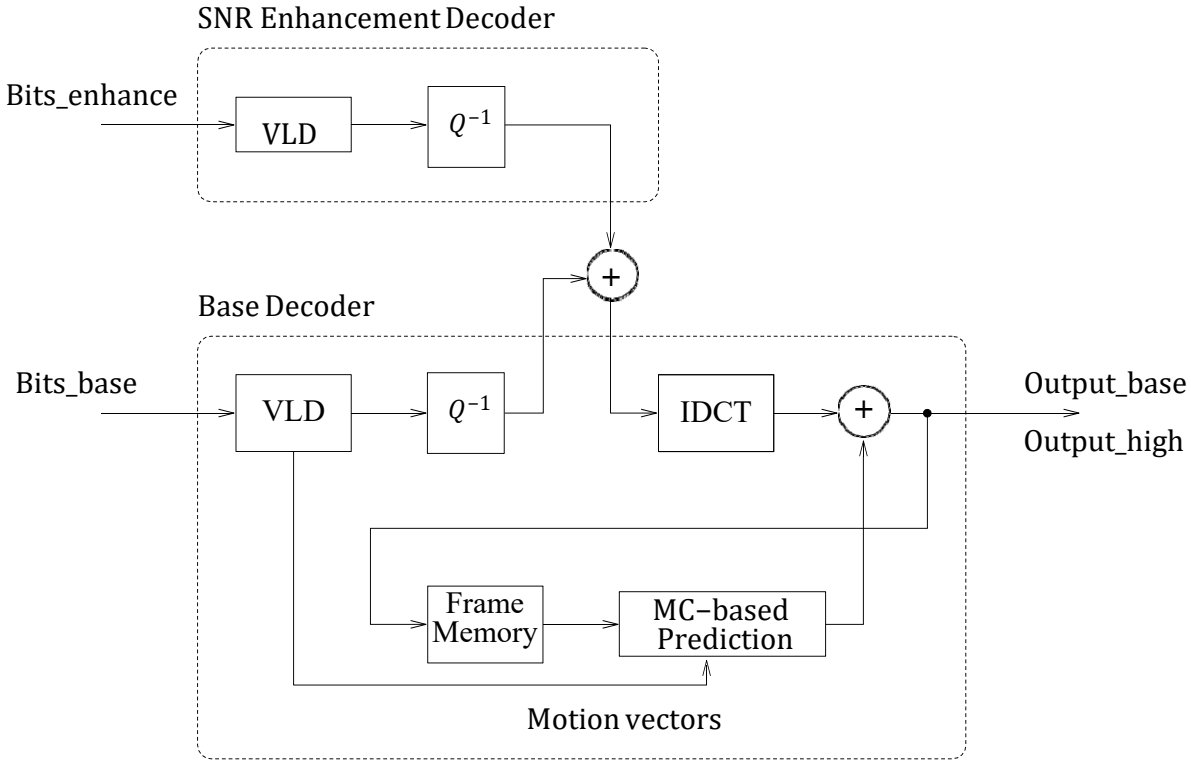
(a) Encoder

Fig 11.8 (a) MPEG-2 SNR Scalability (Encoder).

MPEG-2的编码器

MPEG-2

- MPEG-2 Scalabilities-MPEG-2的可伸缩性



(b) Decoder

Fig 11.8 (b) MPEG-2 SNR Scalability (Decoder).

MPEG-2的解码器

MPEG-2

- Spatial Scalability-MPEG-2的空间可伸缩性
  - The base layer is designed to generate bitstream of reduced resolution pictures. When combined with the enhancement layer, pictures at the original resolution are produced.-基层用来为低分辨率的图像产生位流
  - The Base and Enhancement layers for MPEG-2 spatial scalability are not as tightly coupled as in SNR scalability.-MPEG-2可伸缩性的基层和增强层联系不紧密
  - Fig. 11.9(a) shows a typical block diagram. Fig. 11.9(b) shows a case where temporal and spatial predictions are combined.-时间和空间预测结合的例子

MPEG-2

- Spatial Scalability-MPEG-2的空间可伸缩性

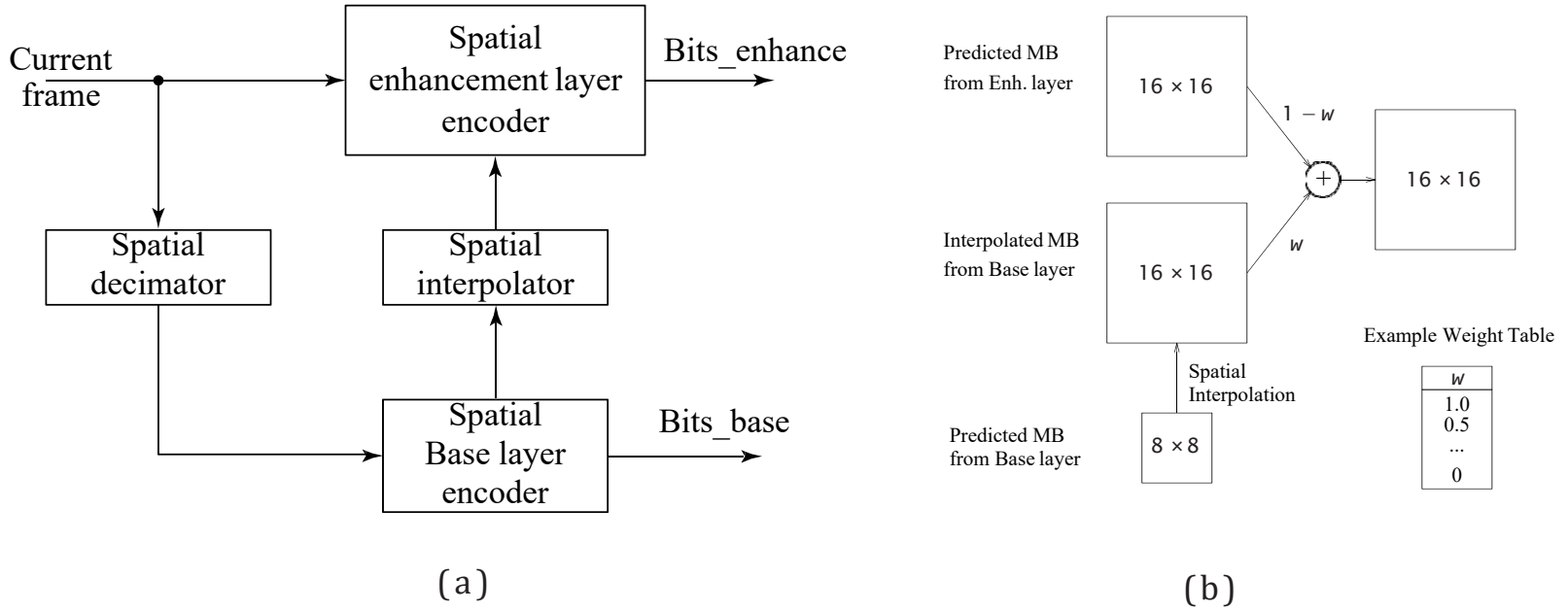


Fig 11.9: Encoder for MPEG-2 Spatial Scalability. (a) Block Diagram. (b) Combining Temporal and Spatial Predictions for Encoding at Enhancement Layer.

MPEG-2空间可伸缩性编码器

MPEG-2

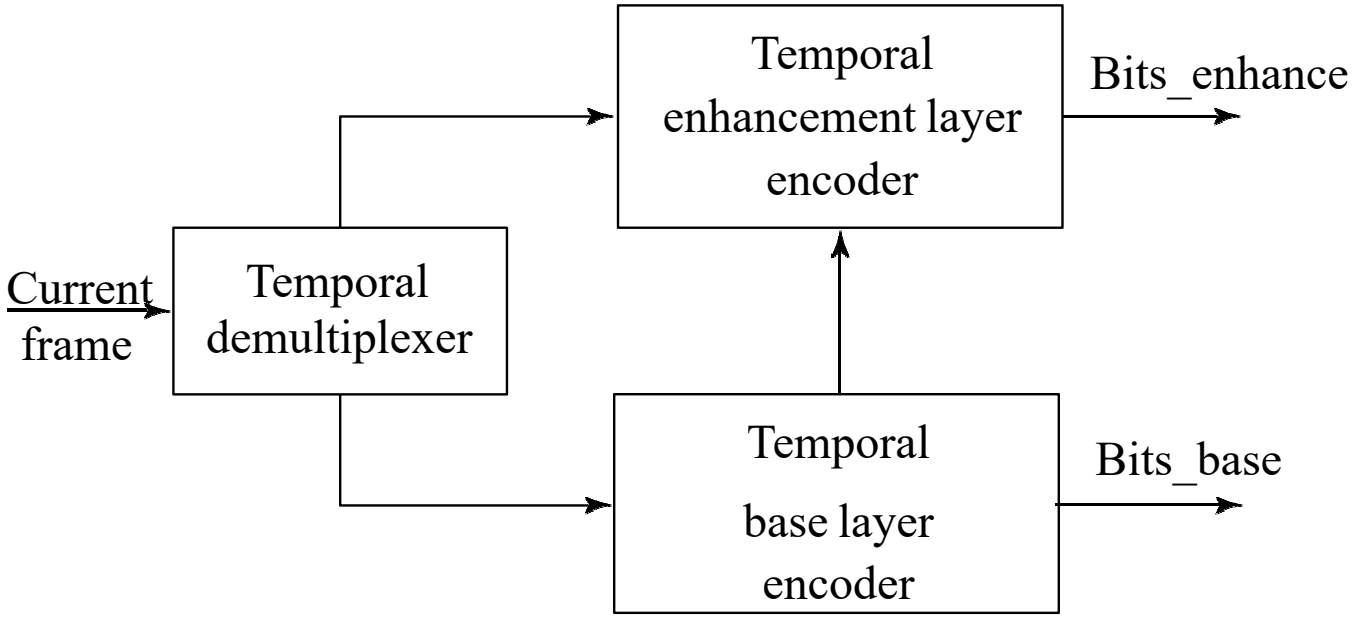
- Temporal Scalability-时间可伸缩性
  - The input video is temporally demultiplexed into two pieces, each carrying half of the original frame rate.-视频在时间上分成两半，帧率减半
  - Base Layer Encoder carries out the normal single-layer coding procedures for its own input video and yields the output bitstream Bits_base.-基层编码器对视频单层编码产生位流
  - The prediction of matching MBs at the Enhancement Layer can be obtained in two ways:-对增强层宏块的预测有两种方式

a) Interlayer MC (Motion-Compensated) Prediction (Fig. 11.10(b))-隔层运动补偿预测

b) Combined MC Prediction and Interlayer MC Prediction (Fig. 11.10(c))-运动补偿预测与隔层运动补偿预测相结合

MPEG-2

◆ Temporal Scalability-时间可伸缩性



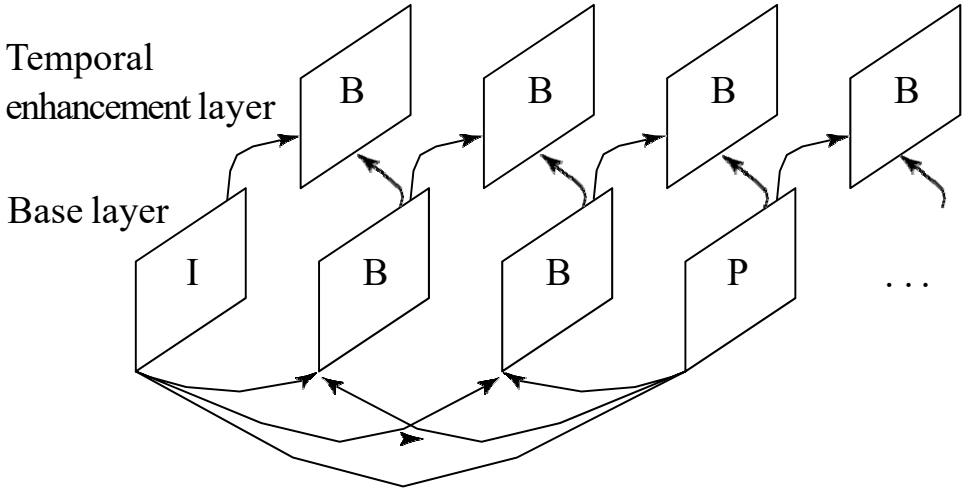
(a) Block Diagram

Fig 11.10: Encoder for MPEG-2 Temporal Scalability.

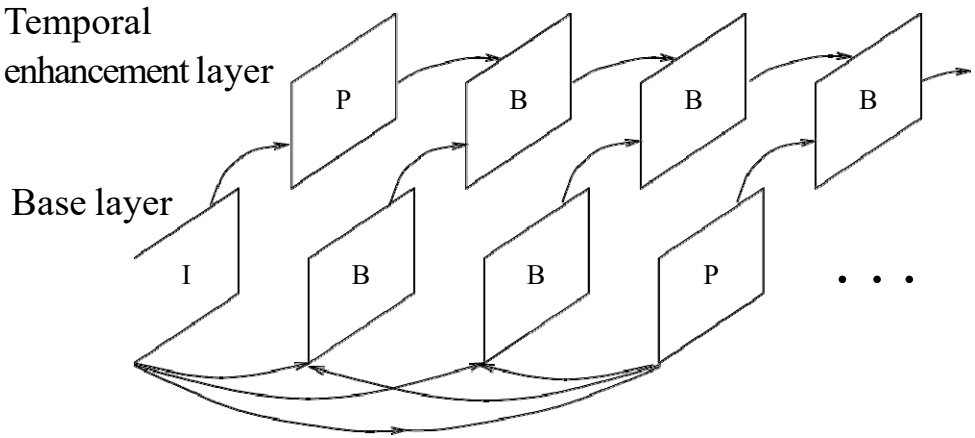
MPEG-2时间可伸缩性编码器

MPEG-2

◆ Temporal Scalability-时间可伸缩性



(a) Interlayer Motion-Compensated (MC) Prediction.



MPEG-2时间可伸缩性编码器

(b) Combined MC Prediction and Interlayer MC Prediction

Fig 11.10 (Cont'd): Encoder for MPEG-2 Temporal Scalability.

MPEG-2

- Hybrid Scalability-混合可伸缩性
  - Any two of the above three scalabilities can be combined to form hybrid scalability:-以上伸缩性可任意组合

  a) Spatial and Temporal Hybrid Scalability.-空间时间混合可伸缩性

  b) SNR and Spatial Hybrid Scalability.-SNR空间混合可伸缩性

  c) SNR and Temporal Hybrid Scalability.-SNR时间混合可伸缩性

  - Usually, a three-layer hybrid coder will be adopted which consists of Base Layer, Enhancement Layer 1, and Enhancement Layer 2.-三层混合编码器

MPEG-2

- Data Partitioning-数据划分
  - Base partition contains lower-frequency DCT coefficients, enhancement partition contains high-frequency DCT coefficients.-压缩视频流划分为两个区
  - Strictly speaking, data partitioning is not layered coding, since a single stream of video data is simply divided up and there is no further dependence on the base partition in generating the enhancement partition.-数据划分不产生相同类型层次编码
  - Useful for transmission over noisy channels and for progressive transmission.-在噪声信道及渐进传输上有效

MPEG-2

- Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
  - Better resilience to bit-errors: In addition to Program Stream, a Transport Stream is added to MPEG-2 bit streams.-更好地从位错误中恢复
  - Support of 4:2:2 and 4:4:4 chroma subsampling.-支持4:2:2和4:4:4色度二次采样
  - More restricted slice structure: MPEG-2 slices must start and end in the same macroblock row. In other words, the left edge of a picture always starts a new slice and the longest slice in MPEG-2 can have only one row of macroblocks.-更严格的宏块片结构
  - More flexible video formats: It supports various picture resolutions as defined by DVD, ATV and HDTV.-更灵活的视频格式

MPEG-2

- Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
  - Nonlinear quantization — two types of scales are allowed:-非线性量化

a) For the first type, scale is the same as in MPEG-1 in which it is an integer in the range of [1, 31] and scalei is equal to i.-一种是与MPEG-1相同的尺度

b) For the second type, a nonlinear relationship exists, i.e., scalei is not equal to i. The i th scale value can be looked up from Table 11.7.-另一种存在一个非线性关系

Table 11.7： Possible Nonlinear Scale in MPEG-2

MPEG-2中可能的非线性尺度

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

Outline of Lecture 12&13

- Overview-概述
- MPEG-1
  - Motion Compensation in MPEG-1 –MPEG-1的运动补偿
  - Other Major Differences from H.261-与H.261的其他主要区别
- MPEG-2
  - Supporting Interlaced Video –支持隔行扫描视频
  - MPEG-2 Scalabilities –MPEG-2的可伸缩性
  - Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
- MPEG-4
  - Overview of MPEG-4-MPEG 4概述
  - Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Synthetic Object Coding in MPEG-4-MPEG-4的合成对象编码
  - MPEG-4 Object types, Profile and Levels-MPEG-4部分、规范和层次
- MPEG-7
- Experiments-实验

MPEG-4

◆ Overview of MPEG-4-MPEG-4概述

- MPEG-4: a newer standard. Besides compression, pays great attention to issues about user interactivities.-MPEG-4除了压缩，更关注用户交互
- MPEG-4 departs from its predecessors in adopting a new object-based coding:-MPEG-4与以前版本不同的是采用基于对象编码的方式

a) offering higher compression ratio, also beneficial for digital video composition, manipulation, indexing, and retrieval.-基于对象编辑的好处

b) illustrates how MPEG-4 videos can be composed and manipulated by simple operations on the visual objects.-下图展示MPEG-4如何对可是对象进行简单的操作

- The bit-rate for MPEG-4 video now covers a large range between 5 kbps to 10 Mbps.-MPEG-4的码率更高

MPEG-4

- Overview of MPEG-4-MPEG-4概述

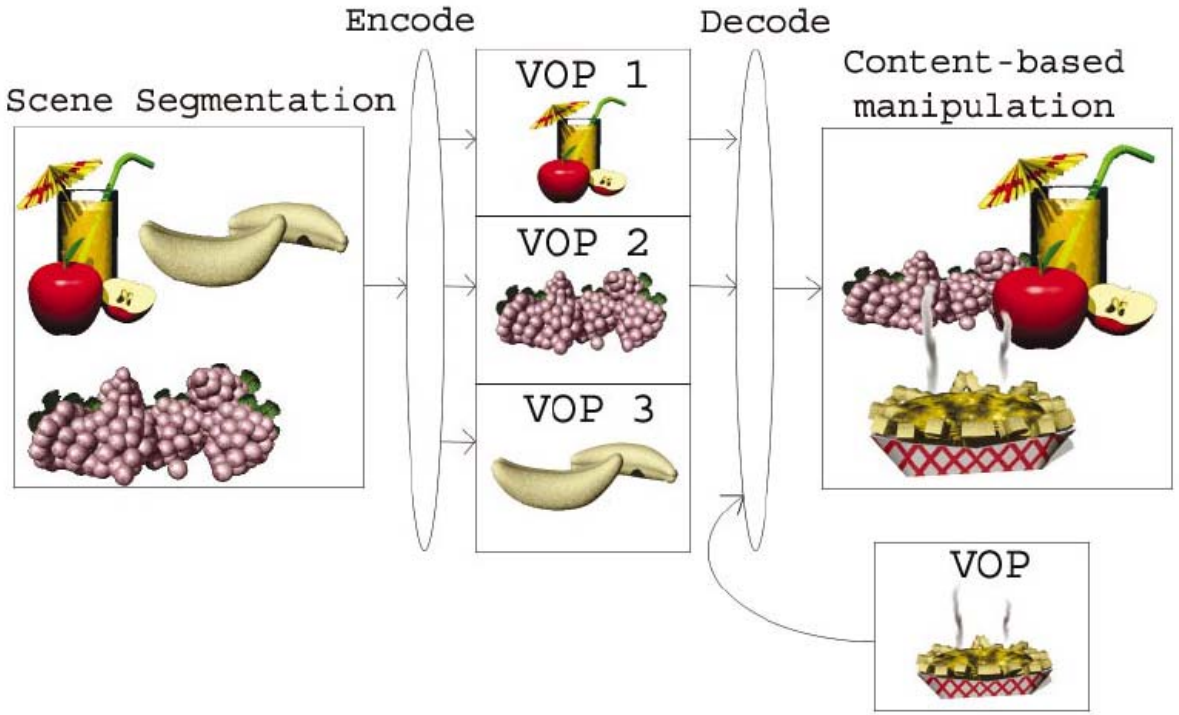


Fig. 13.1: Composition and Manipulation of MPEG-4 Videos.

MPEG-4视频的合成和操作

MPEG-4

- Overview of MPEG-4-MPEG-4概述
  - MPEG-4 (Fig. 13.2(b)) is an entirely new standard for:-全新的标准

a) Composing media objects to create desirable audiovisual scenes.-通过合成媒体创建视听场景

b) Multiplexing and synchronizing the bitstreams for these media data entities so that they can be transmitted with guaranteed Quality of Service (QoS).-通过多路传播和同步媒体数据位流保证传输服务质量

c) Interacting with the audiovisual scene at the receiving end — provides a toolbox of advanced coding modules and algorithms for audio and video compressions.-可在接收端与视听场景进行交互

MPEG-4

- Overview of MPEG-4-MPEG-4概述

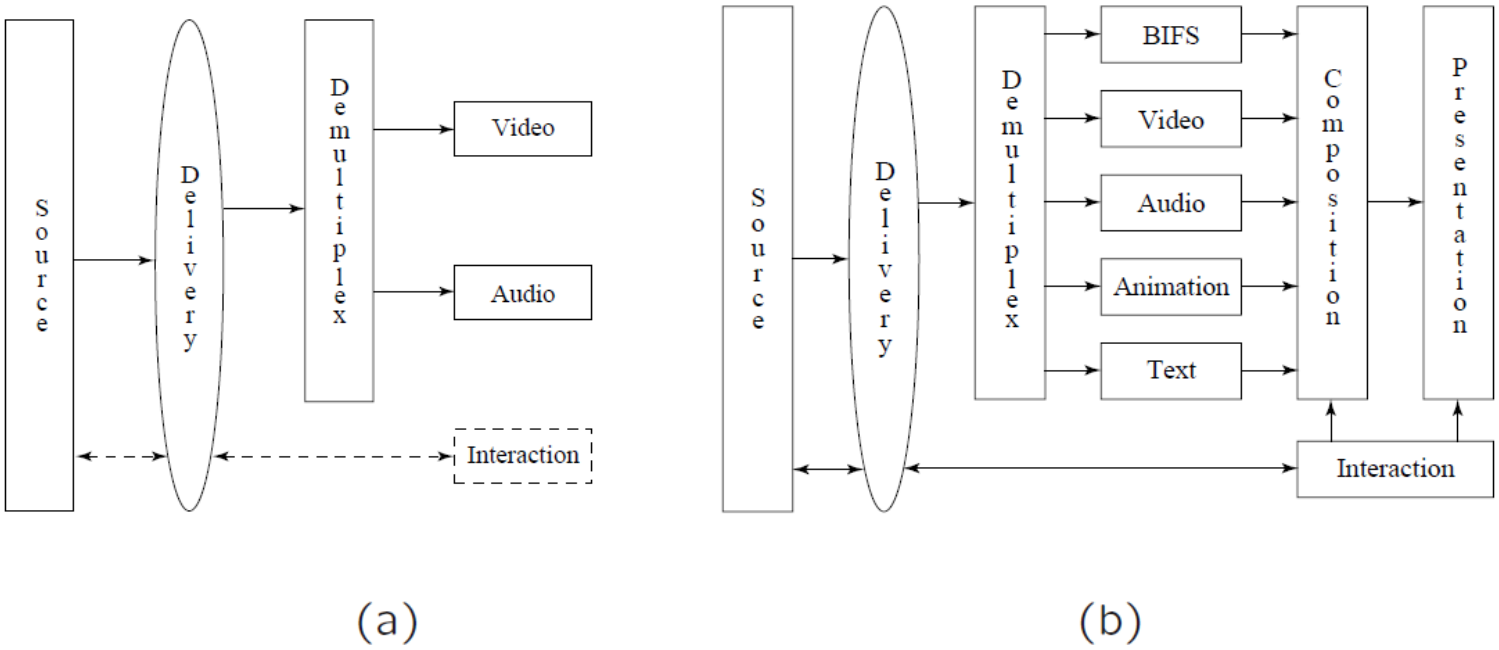


Fig. 13.2: Comparison of interactivities in MPEG standards: (a) reference models in MPEG-1 and 2 (interaction in dashed lines supported only by MPEG-2); (b) MPEG-4 reference model.

MPEG标准中交互性的比较

MPEG-4

- Overview of MPEG-4-MPEG-4概述
  - The hierarchical structure of MPEG-4 visual bitstreams is very different from that of MPEG-1 and -2, it is very much video object-oriented.-MPEG-4视觉位流层次结构与MPEG-1和MPEG-2视觉位流区别很大

Video-object Sequence (VS)
Video Object (VO)
Video Object Layer (VOL)
Group of VOPs (GOV)
Video Object Plane (VOP)

Fig. 13.3: Video Object Oriented Hierarchical Description of a Scene in MPEG-4 Visual Bitstreams.

MPEG-4视觉位流中面向视频对象的场景分层描述

MPEG-4

- Overview of MPEG-4-MPEG-4概述
  - 1. Video-object Sequence (VS) — delivers the complete MPEG-4 visual scene, which may contain 2-D or 3-D natural or synthetic objects.-视频对象序列
  - 2. Video Object (VO) — a particular object in the scene, which can be of arbitrary (non-rectangular) shape corresponding to an object or background of the scene.-视频对象
  - 3. Video Object Layer (VOL) — facilitates a way to support (multi-layered) scalable coding. A VO can have multiple VOLs under scalable coding, or have a single VOL under non-scalable coding.-视频对象层
  - 4. Group of Video Object Planes (GOV) — groups Video Object Planes together (optional level).-视频对象平面组
  - 5. Video Object Plane (VOP) — a snapshot of a VO at a particular moment.-视频对象屏幕

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based vs. Frame-based Coding-基于对象的视觉编码与基于帧的编码

a) MPEG-1 and -2 do not support the VOP concept, and hence their coding method is referred to as frame-based (also known as Block-based coding).-MPEG-1和MPEG-2的编码方式为基于帧的编码方式

b) Fig. 12.4 (c) illustrates a possible example in which both potential matches yield small prediction errors for block-based coding.-图中举例说明潜在的匹配可能对基于块的编码产生小的预测误差

c) Fig. 12.4 (d) shows that each VOP is of arbitrary shape and ideally will obtain a unique motion vector consistent with the actual object motion.-基于对象的视觉编码可有任意形状，并在理性情况可以获取唯一的与对象运动一致的运动向量。

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码

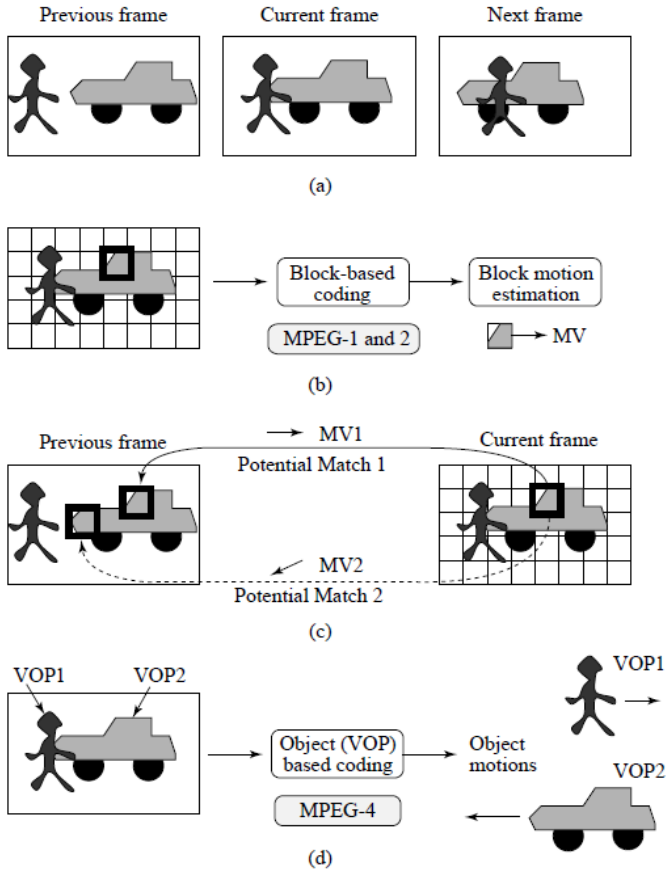


Fig. 13.4: Comparison between Block-based Coding and Object-based Coding.

基于块的编码和基于对象的编码的比较

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) MPEG-4 VOP-based coding also employs the Motion Compensation technique:-基于对象的视觉编码也采用运动补偿技术

  - An Intra-frame coded VOP is called an I-VOP.-VOP中的帧内编码I-VOP
  - The Inter-frame coded VOPs are called P-VOPs if only forward prediction is employed, or B-VOPs if bi-directional predictions are employed.-VOP中的帧间编码与基于双向预测的帧间编码

b) The new difficulty for VOPs: may have arbitrary shapes, shape information must be coded in addition to the texture of the VOP.-基于对象的视觉编码的难点

Note: texture here actually refers to the visual content, that is the gray-level (or chroma) values of the pixels in the VOP.-纹理指像素灰度值

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) MC-based VOP coding in MPEG-4 again involves three steps:-基于运动补偿的VOP编码涉及三个步骤

- Motion Estimation.-运动估计
- MC-based Prediction.-基于运动补偿的预测
- Coding of the prediction error.-预测误差编码

b) Only pixels within the VOP of the current (Target) VOP are considered for matching in MC.-只有当前VOP内部的像素才需要进行运动补偿的匹配

c) To facilitate MC, each VOP is divided into many macroblocks (MBs). MBs are by default 16×16 in luminance images and 8 × 8 in chrominance images.-每一个VOP都被分为许多宏块

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) MPEG-4 defines a rectangular bounding box for each VOP (see Fig. 13.5 for details).-MPEG-4为每一个VOP定义矩形边界框

b) The macroblocks that are entirely within the VOP are referred to as Interior Macroblocks. –完全处于VOP内部的宏块称为内部宏块

c) The macroblocks that straddle the boundary of the VOP are called Boundary Macroblocks.-跨越了VOP边界的宏块，称为边界宏块

d) To help matching every pixel in the target VOP and meet the mandatory requirement of rectangular blocks in transform coding (e.g., DCT), a pre-processing step of padding is applied to the Reference VOPs prior to motion estimation.-匹配前要对参考VOP进行填充处理

Note: Padding only takes place in the Reference VOPs.-填充只在参考VOP发生

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

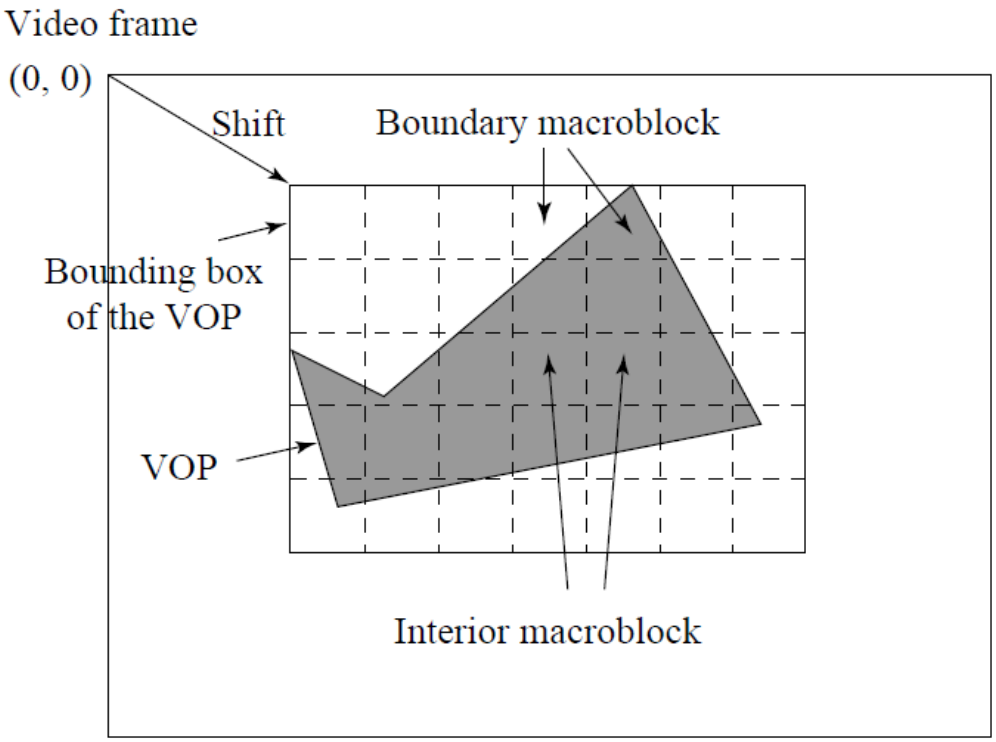


Fig. 13.5: Bounding Box and Boundary Macroblocks of VOP.

VOP的边界框和边界宏块

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) Padding-填充

- For all Boundary MBs in the Reference VOP, Horizontal Repetitive Padding is invoked first, followed by Vertical Repetitive Padding.-先水平重复填充，再垂直重复填充

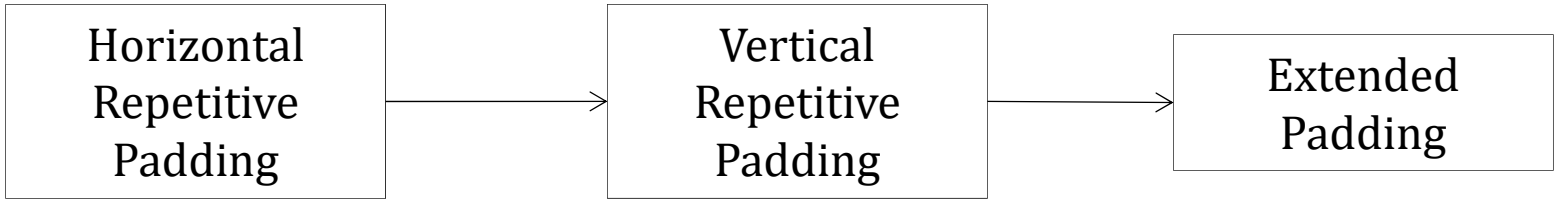


Fig. 13.6: A Sequence of Paddings for Reference VOPs in MPEG-4.
MPEG-4中参考VOP的一系列填充

- Afterwards, for all Exterior Macroblocks that are outside of the VOP but adjacent to one or more Boundary MBs, extended padding will be applied.-对在VOP之外但与边界宏块相邻的外部宏块，进行扩展填充

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) Padding-填充

Algorithm 13.1 Horizontal Repetitive Padding: 水平重复填充

```
begin
  for all rows in Boundary MBs in the Reference VOP
    if ∃ (boundary pixel) in the row
      for all interval outside of VOP
        if interval is bounded by only one boundary pixel b
          assign the value of b to all pixels in interval
        else    // interval is bounded by two boundary pixels b1 and b2
          assign the value of (b1 +b2)/2 to all pixels in interval
end
```

➢ The subsequent Vertical Repetitive Padding algorithm works in a similar manner.-垂直重复填充算法与之类似

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) Padding-填充

Example 13.1: Repetitive Paddings

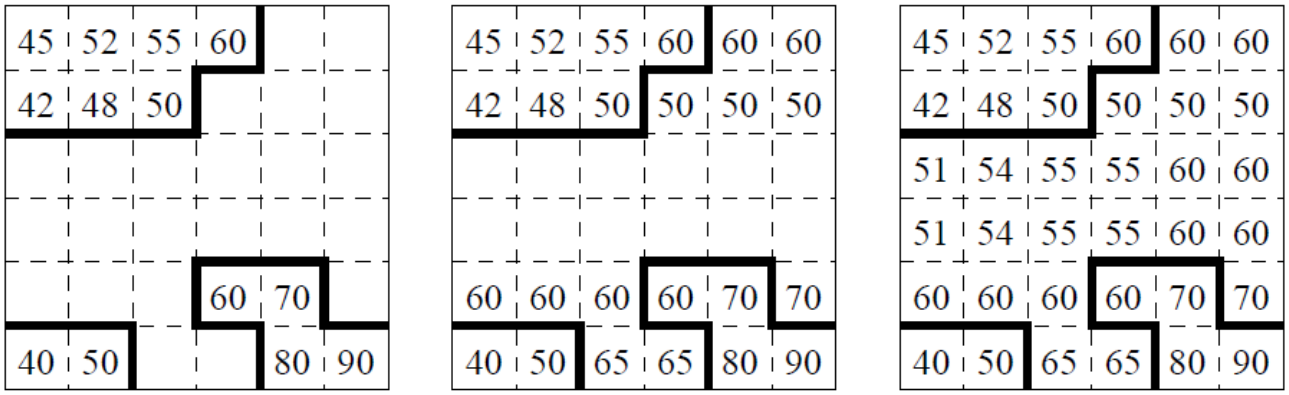


Fig. 13.7: An example of Repetitive Padding in a boundary macroblock of a Reference VOP: (a) Original pixels within the VOP, (b) After Horizontal Repetitive Padding, (c) Followed by Vertical Repetitive Padding.

参考VOP的边界宏块中重复填充的示例

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - VOP-based Coding-基于对象的视觉编码

a) Motion Vector Coding-运动向量编码

- Let C(x+k;y+l) be pixels of the MB in Target VOP, and R(x+i+k;y+j+l)be pixels of the MB in Reference VOP.-定义目标VOP和参考VOP的宏块像素
- A Sum of Absolute Difference (SAD) for measuring the difference between the two MBs can be defined as:-采用SAD作为误差度量

$$SAD(i,j) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} |\, C(x+k, y+l) - R(x+i+k, y+i+l)| \,.\, Map(x+k, y+l)$$

N—the size of the MB. $Map(p,q) = 1$ when $C(p,q)$ is a pixel within the target VOP, otherwise $Map(p,q) = 0$

- The vector (i:j)that yields the minimum SAD is adopted as the motion vector MV(u;v):-SAD最小对应为运动向量

$$(u,v) = [(i,j) \mid SAD(i,j) \text{ is minimum, i} \in [-p,p], j \in [-p,p]]$$

p — the maximal allowable magnitude for u and v

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Texture Coding-纹理编码

a) Texture coding in MPEG-4 can be based on:-MPEG-4纹理编码基础

- DCT or-离散余弦变换
- Shape Adaptive DCT (SA-DCT).-基于形状自适应DCT

b) Texture coding based on DCT:-基于DCT的纹理编码

- In I-VOP, the gray values of the pixels in each MB of the VOP are directly coded using the DCT followed by VLC, similar to what is done in JPEG.-使用VLC接着使用DCT直接编码
- In P-VOP or B-VOP, MC-based coding is employed — it is the prediction error that is sent to DCT and VLC.-在P-VOP和B-VOP中采用基于运动补偿的编码

MPEG-4

- Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Texture Coding-纹理编码

  b) Texture coding based on DCT:-基于DCT的纹理编码

  ➢ Coding for the Interior MBs:-对内部宏块编码

  Each MB is 16 ×16 in the luminance VOP and 8×8 in the chrominance VOP.-宏块的大小

  Prediction errors from the six 8×8 blocks of each MB are obtained after the conventional motion estimation step.-获得预测误差

  ➢ Coding for Boundary MBs:-对边界宏块编码

  For portions of the Boundary MBs in the Target VOP outside of the VOP, zeros are padded to the block sent to DCT since ideally prediction errors would be near zero inside the VOP.-在目标VOP边界宏块中VOP外的部分用0填充

  After MC, texture prediction errors within the Target VOP are obtained.-运动补偿后，获得目标VOP的纹理预测误差

MPEG-4

- MPEG-4 Object types, Profiles and Levels -MPEG-4部分、规范和层次
  - The standardization of Profiles and Levels in MPEG-4 serve two main purposes: –MPEG-4 规范和层次标准化两个原因
    - ensuring interoperability between implementations-确保具体实现部分的协调性
    - allowing testing of conformance to the standard-标准一致性
  - MPEG-4 not only specified Visual profiles and Audio profiles, but it also specified Graphics profiles, Scene description pro- files, and one Object descriptor profile in its Systems part.-MPEG-4定义了视觉规格和音频规格
  - **Object type** is introduced to dene the tools needed to create video objects and how they can be combined in a scene.-引入部件并如何结合到场景中
  - For “Main Profile”, for example, only Object Types “Simple”, “Core”, “Main”, and “Scalable Still Texture” are supported. MPEG-4 Levels in Simple, Core, and Main Visual Profiles.-MPEG-4定义了视频规范，简单、高阶简单、核心、主要、简单扩展等

Outline of Lecture 12&13

- Overview-概述
- MPEG-1
  - Motion Compensation in MPEG-1 –MPEG-1的运动补偿
  - Other Major Differences from H.261-与H.261的其他主要区别
- MPEG-2
  - Supporting Interlaced Video –支持隔行扫描视频
  - MPEG-2 Scalabilities –MPEG-2的可伸缩性
  - Other Major Differences from MPEG-1-与MPEG-1的其他主要区别
- MPEG-4
  - Overview of MPEG-4-MPEG 4概述
  - Object-based Visual Coding in MPEG-4-MPEG-4的基于对象的视觉编码
  - Synthetic Object Coding in MPEG-4-MPEG-4的合成对象编码
  - MPEG-4 Object types, Profile and Levels-MPEG-4部分、规范和层次
- MPEG-7
- Experiments-实验

63

MPEG-7

- Introduction of MPEG-7 -MPEG-7介绍
  - The main objective of MPEG-7 is to serve the need of audiovisual content-based retrieval (or audiovisual object retrieval) in applications such as digital libraries.-MPEG-7主要服务于视听内容检索服务
  - Nevertheless, it is also applicable to any multimedia applications involving the generation (content creation) and usage (content consumption) of multimedia data.-MPEG-7可应用于任何多媒体
  - MPEG-7 became an International Standard in September 2001 — with the formal name Multimedia Content Description Interface.-MPEG-7于2001成为国际标准-多媒体内容描述接口

MPEG-7

- Introduction of MPEG-7 -MPEG-7介绍
  - The main objective of MPEG-7 is to serve the need of audiovisual content-based retrieval (or audiovisual object retrieval) in applications such as digital libraries.-MPEG-7主要服务于视听内容检索服务
  - Nevertheless, it is also applicable to any multimedia applications involving the generation (content creation) and usage (content consumption) of multimedia data.-MPEG-7可应用于任何多媒体
  - MPEG-7 became an International Standard in September 2001 — with the formal name Multimedia Content Description Interface.-MPEG-7于2001成为国际标准-多媒体内容描述接口

MPEG-7

- Applications Supported by MPEG-7-MPEG-7支持的应用
  - MPEG-7 supports a variety of multimedia applications. Its data may include still pictures, graphics, 3D models, audio, speech, video, and composition information (how to combine these elements).- MPEG-7支持多种多样多媒体应用
  - These MPEG-7 data elements can be represented in textual format, or binary format, or both.-MPEG-7可用文本或二进制进行表示
  - Fig. 13.17 illustrates some possible applications that will benefit from the MPEG-7 standard.-下图展示采用MPEG-7的一些应用

MPEG-7

- Applications Supported by MPEG-7-MPEG-7支持的应用

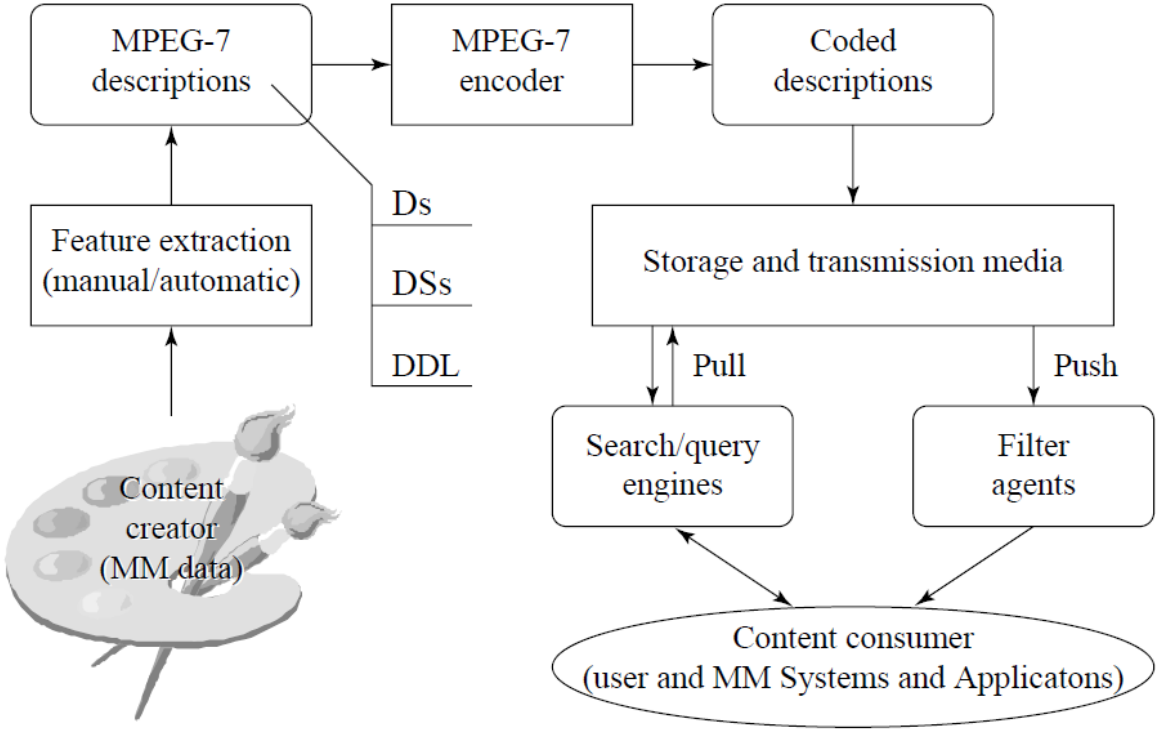


Fig. 13.17: Possible Applications using MPEG-7.

可能使用MPEG-7的应用

MPEG-7

- MPEG-7 and Multimedia Content Description-MPEG-7和多媒体内容描述
  - MPEG-7 has developed Descriptors (D), Description Schemes (DS) and Description Definition Language (DDL). The following are some of the important terms:-MPEG-7开发了多媒体内容描述的描述子、描述方案等
    - Feature — characteristic of the data.-特征
    - Descriptors(D) — definition (syntax and semantics) of the feature.-描述子
    - Description Schemes(DS) — specification of the structure and relationship between Ds and between DSs.-描述方案
    - Description — a set of instantiated Ds and DSs that describes the structural and conceptual information of the content, the storage and usage of the content, etc.-描述
    - Description Definition Language(DDL) — syntactic rules to express and combine DSs and Ds.-描述定义语言
  - The scope of MPEG-7 is to standardize the Ds, DSs and DDL for descriptions. The mechanism and process of producing and consuming the descriptions are beyond the scope of MPEG-7.-MPEG-7的范畴

Experiments & Class Assignments

- Experiments

  - 编程实现第11章程序11.1 水平重复填充程序（以例11.1为例子）

- Class Assignments

  1、为什么B帧不能在运动补偿中作为参考帧？假设任何类型的帧都可作为参考帧。讨论在视频序列中用参考B帧代替P帧的利弊（即完全消除P帧）。

  2、编程实现运动向量的分层搜索（语言不限）。