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

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

Introduction to Video Compression-视频压缩简介

- What is video compression?-什么是视频压缩
  - A video consists of a time-ordered sequence of frames, i.e., images.-视频由时间上有序的图像组成
  - An obvious solution to video compression would be predictive coding based on previous frames.-一种视频压缩方案是基于过往帧的预测编码
    Compression proceeds by subtracting images: subtract in time order and code the residual error.-压缩也可通过前后帧相减，然后对残差部分编码实现
  - It can be done even better by searching for just the right parts of the image to subtract from the previous frame.-通过搜索图像中适当部分并和前一帧相减来实现

Outline of Lecture 12

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

Video Compression with Motion Compensation-基于运动补偿的视频压缩

- Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Consecutive frames in a video are similar-temporal redundancy exists.-连续帧存在时间冗余
  - **Temporal redundancy** is exploited so that not every frame of the video needs to be coded independently as a new image.-由于存在时间冗余，不是所有帧都需要单独编码
  - The difference between the current frame and other frame(s) in the sequence will be coded - small values and low entropy, good for compression.-对帧间误差编码好处是信息量少、低熵值
  - Steps of Video compression based on Motion Compensation(MC):-运动补偿步骤
    1. Motion Estimation (motion vector search).-运动估计
    2. MC-based Prediction.-基于运动补偿的预测
    3. Derivation of the prediction error, i.e., the difference.-生成预测误差

Video Compression with Motion Compensation-基于运动补偿的视频压缩

- Motion Compensation-运动补偿
  - Each image is divided into macroblocks of size N × N.-图像由宏块组成

  a) By default, N = 16 for luminance images. For chrominance images, N = 8 if 4:2:0 chroma subsampling is adopted.-亮度图像N = 16， 色度图像N = 8，采样格式4:2:0

  - Motion compensation is performed at the macroblock level.-宏块补偿

  a) The current image frame is referred to as Target Frame.-当前帧为目标帧

  b) A match is sought between the macroblock in the Target Frame and the most similar macroblock in previous and/or future frame(s)referred to as Reference frame(s)).-寻找目标帧宏块和参考帧中最相似宏块的匹配

  c) The displacement of the reference macroblock to the target macroblock is called a motion vector MV.-参考宏块到目标宏块的位移称作运动向量

  d) Figure 10.1 shows the case of forward prediction in which the Reference frame is taken to be a previous frame.-前向预测情况

Video Compression with Motion Compensation-基于运动补偿的视频压缩

- Motion Compensation-运动补偿

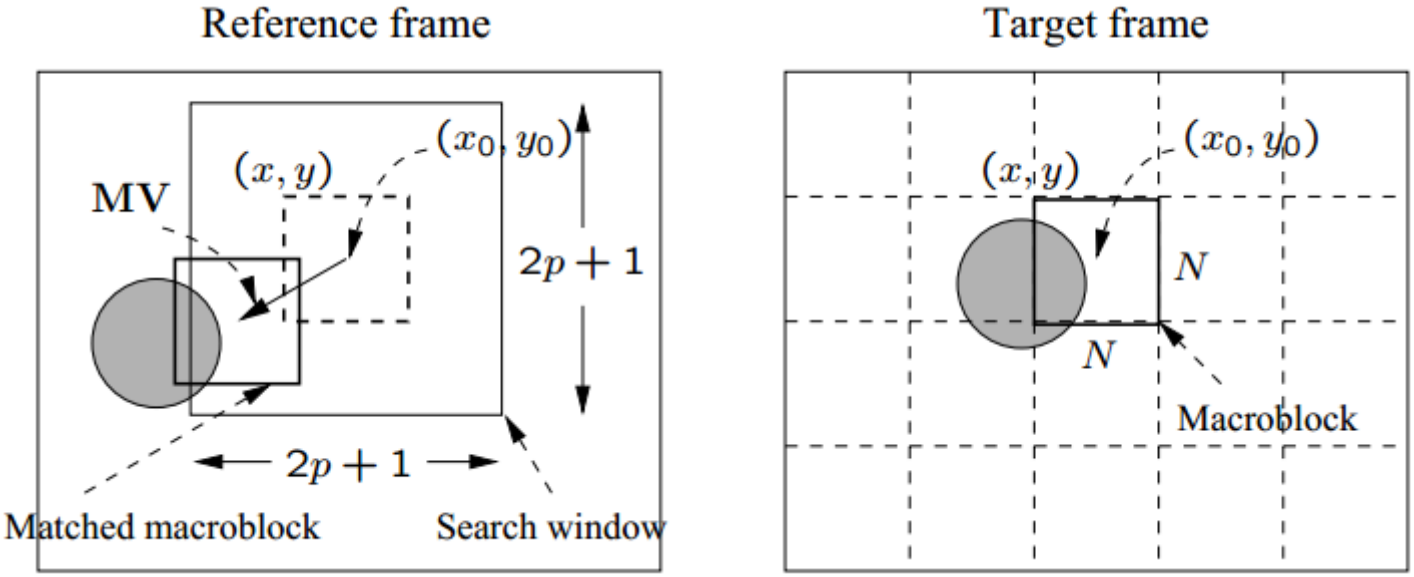


Fig. 10.1: Macroblocks and Motion Vector in Video Compression.

视频压缩中的宏块和运动向量

e) MV search is usually limited to a small immediate neighborhood - both horizontal and vertical displacements in the range [−p; p]: -搜索范围限制在[−p; p]

This makes a search window of size (2p + 1) × (2p + 1).-窗口大小

Outline of Lecture 12

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

Search for Motion Vectors-搜索运动向量

- Objective function for motion vector searching-目标函数
  - The difference between two macroblocks can then be measured by their Mean Absolute Difference (MAD):-采用平均绝对值误差来衡量

$$MAD(i,j) = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}|C(x+k,y+l) - R(x+i+k,y+j+l)| \quad (10.1)$$

N — size of the macroblock,-宏块大小

k and l — indices for pixels in the macroblock,-宏块中像素的索引

i and j — horizontal and vertical displacements,-水平和垂直的位移

$C(x+k,y+l)$ — pixels in macroblock in Target frame,-目标帧宏块的像素

$R(x+i+k,y+j+l)$ — pixels in macroblock in Reference frame.-参考帧宏块的像素

  - The goal of the search is to find a vector (i ; j) as the motion vector MV = (u; v), such that $MAD(i,j)$is minimum:-当MAD值最小，找到最佳MV

$$(u,v) = [(i,j)|MAD(i,j)\ is\ minimum, i \in [-p,p], j \in [-p,p]] \quad (10.2)$$

Search for Motion Vectors-搜索运动向量

- Sequential search-顺序搜索
  - sequentially search the whole (2p + 1) × (2p + 1) window in the Reference frame (also referred to as Full search). –参考帧窗口的全搜索

  a) a macroblock centered at each of the positions within the window is compared to the macroblock in the Target frame pixel by pixel and their respective MAD is then derived using Eq. (10.1).-窗口宏块与目标宏块 比较

  b) The vector (i ; j) that offers the least MAD is designated as the MV (u; v) for the macroblock in the Target frame.-MAD最小的向量为目标帧宏块的运动向量

  c) sequential search method is very costly — assuming each pixel comparison requires three operations (subtraction, absolute value, addition), the cost for obtaining a motion vector for a single macroblock is $(2p+1)\cdot(2p+1)\cdot N^2\cdot 3 \Rightarrow O(p^2N^2)$.-顺序搜索非常耗时

Search for Motion Vectors-搜索运动向量

- Sequential search-顺序搜索

PROCEDURE 10.1 Motion-vector:sequential-search

寻找运动向量的顺序搜索算法

```
Begin
    min_MAD = LARGE _NUMBER;      /* Initialization */
    for i = -p to p
        for j = -p to p
        {
            cur_MAD = MAD(i,j)
            if cur_MAD < min_MAD
            {
                min_MAD = cur_MAD;
                u = i;                    /* Get the coordinates for MV. */
                v = j;
            }
        }
end
```

Search for Motion Vectors-搜索运动向量

- **2D Logarithmic search-2D对数搜索**
  - a cheaper version, that is suboptimal but still usually effective. –不是最优，但非常有效
  - The procedure for 2D Logarithmic Search of motion vectors takes several iterations and is akin to a binary search-多次迭代，类似折半查找

a) As illustrated in Fig.10.2, initially only nine locations in the search window are used as seeds for a MAD-based search; they are marked as '1'.-起始搜索的9个候选位置被标记为1

b) After the one that yields the minimum MAD is located, the center of the new search region is moved to it and the step-size ("offset") is reduced to half.-MAD最小时确定新的搜索区域中心，同时步长减半

c) In the next iteration, the nine new locations are marked as '2', and so on-新一轮迭代的9个候选位置被标记为2，如此类推

Search for Motion Vectors-搜索运动向量

- 2D Logarithmic search-2D对数搜索

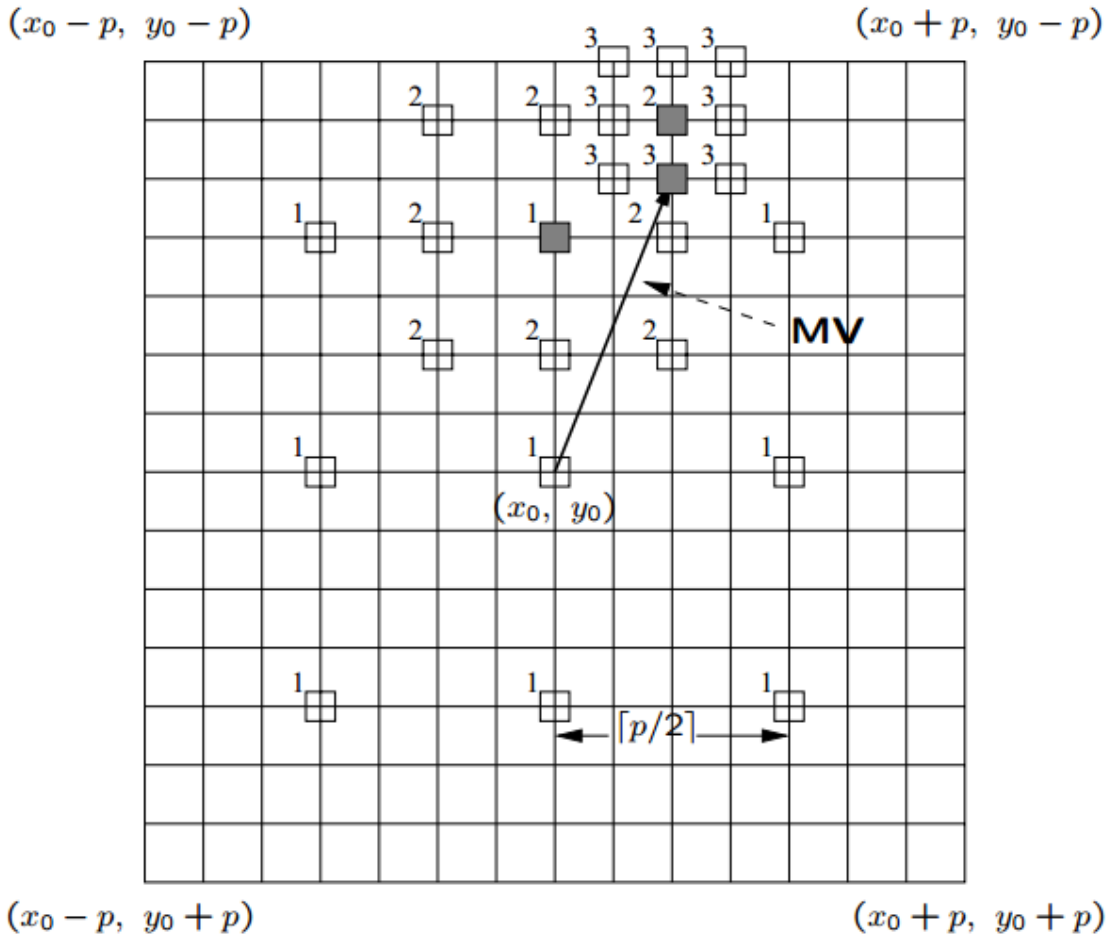


Fig. 10.2: 2D Logarithmic Search for Motion Vectors.

2D对数搜索算法示意图

Search for Motion Vectors-搜索运动向量

- 2D Logarithmic search-2D对数搜索

PROCEDURE 10.2 Motion-vector:2D-logarithmic-search

寻找运动向量的2D对数搜索算法

begin

offset = $\left\lceil \frac{p}{2} \right\rceil$;

Specify nine macroblocks within the search window in the Reference frame, they are centered at $(x_0; y_0)$ and separated by offset horizontally and/or vertically;

while last ≠ TRUE

{

Find one of the nine specified macroblocks that yields minimum *MAD*;

if offset = 1 then last = TRUE;

offset = $\lceil offset/2 \rceil$;

Form a search region with the new offset and new center found;

}

end

Search for Motion Vectors-搜索运动向量

- 2D Logarithmic search-2D对数搜索

d) Using the same example as in the previous subsection, the total operations per second is dropped to:-每秒的计算量相应减小

$$OPS\_per\_second = (8 \cdot (\lceil \log_2 p \rceil + 1) + 1) \cdot N^2 \cdot 3 \cdot \frac{728 \times 480}{N \cdot N} \cdot 30$$

$$= (8 \cdot (\lceil \log_2 15 \rceil + 1) + 1) \times 16^2 \times 3 \times \frac{728 \times 480}{16 \times 16} \times 30$$

$$\approx 1.25 \times 10^9$$

Search for Motion Vectors-搜索运动向量

- Hierarchical Search-分层搜索
  - The search can benefit from a hierarchical (multiresolution) approach in which initial estimation of the motion vector can be obtained from images with a significantly reduced resolution.-通过降低分辨率对初始运动向量估计
  - Figure 10.3: a three-level hierarchical search in which the original image is at Level 0, images at Levels 1 and 2 are obtained by down-sampling from the previous levels by a factor of 2, and the initial search is conducted at Level 2-分层搜索第0层是原始图像，第i层是i-1层的下采样（下采样因子为2）
  - Since the size of the macroblock is smaller and p can also be proportionally reduced, the number of operations required is greatly reduced.-低分辨率下的宏块更小，计算量大大减小

Search for Motion Vectors-搜索运动向量

- Hierarchical Search-分层搜索

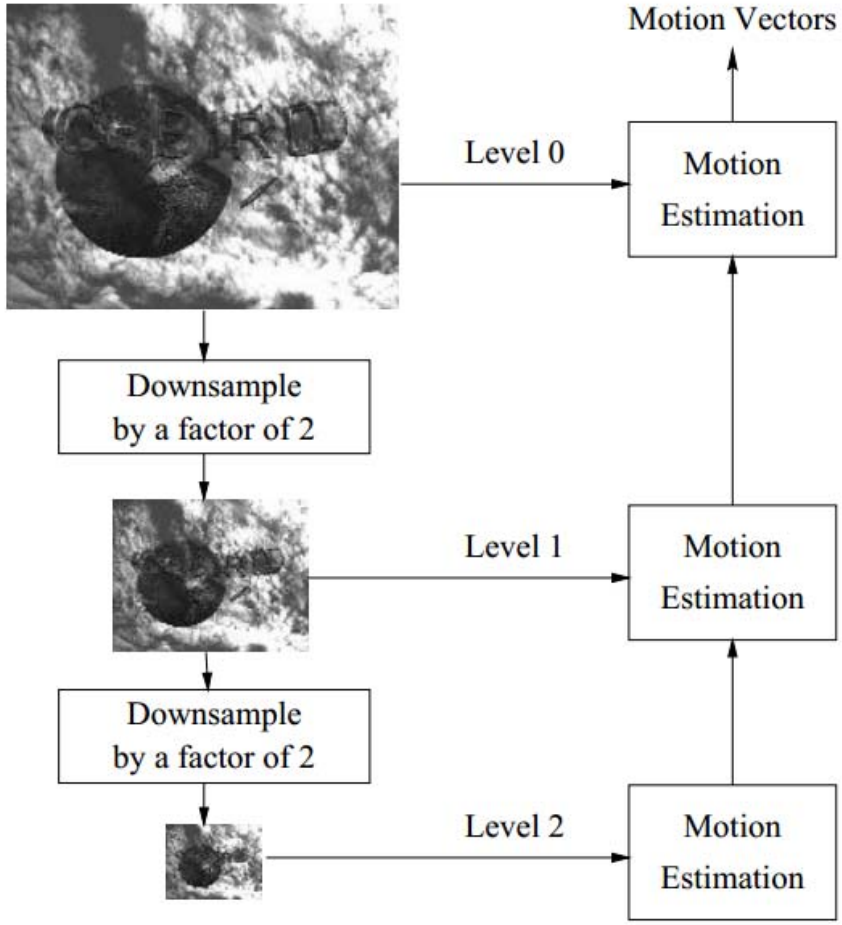


Fig. 10.3: A Three-level Hierarchical Search for Motion Vectors.

一个三层的运动向量搜索示意图

Search for Motion Vectors-搜索运动向量

- Hierarchical Search-分层搜索

  - Given the estimated motion vector (u^k, v^k) at Level k, a 3×3 neighborhood centered at $(2 \cdot u^k; 2 \cdot v^k)$ at Level k – 1 is searched for the refined motion vector. – 给定第k层的估计运动向量，返回k-1层修正运动向量的估计值

  - the refinement is such that at Level k – 1 the motion vector (u^{k-1}, v^{k-1}) satisfies:-第k-1层运动向量满足以下条件

    $$\left(2u^k - 1 \le u^{k-1} \le 2u^k + 1, 2v^k - 1 \le v^{k-1} \le 2v^k + 1\right)$$

  - Let $\left(x_0^k, y_0^k\right)$ denote the center of the macroblock at Level k in the Target frame. The procedure for hierarchical motion vector search for the macroblock centered at (x_0^0, y_0^0)in the Target frame can be outlined as follows:-给定第k层宏块的位置，搜索目标帧宏块位置的分层算法如下

Search for Motion Vectors-搜索运动向量

- Hierarchical Search-分层搜索

PROCEDURE 10.3 Motion-vector:hierarchical-search

分层搜索算法

begin

// Get macroblock center position at the lowest resolution Level k

$x_0^k = x_0^0/2^k; y_0^k = y_0^0/2^k;$

Use Sequential (or 2D Logarithmic) search method to get initial estimated MV(u^k, v^k) at Level k;

while last ≠ TRUE

{

Find one of the nine macroblocks that yields minimum MAD at Level $k-1$ centered at

$\left(2\left(x_0^k + u^k\right) - 1 \le x \le 2\left(x_0^k + u^k\right) + 1, 2\left(y_0^k + v^k\right) - 1 \le y \le 2\left(y_0^k + v^k\right) + 1\right);$

if $k = 1$ then last = TRUE;

$k = k - 1;$

Assign (x_0^k, y_0^k) and (u^k, v^k) with the new center location and MV;

}

end

Search for Motion Vectors-搜索运动向量

- 运动向量搜索算法效率对比

Table 10.1 Comparison of Computational Cost of Motion Vector Search based on examples

运动向量搜索方法计算代价对比

Search Method	OPS_per_second for 728×480 at 30 fps	
	$p = 15$	$p = 7$
Sequential search	29.89×10^9	7.00×10^9
2D Logarithmic search	1.25×10^9	0.78×10^9
3-level Hierarchical search	0.51×10^9	0.40×10^9

Outline of Lecture 12

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

H.261

- Introduction to H.261-H.261简介
  - An earlier digital video compression standard, its principle of MC-based compression is retained in all later video compression standards.-H.261基于运动补偿压缩原理，并被后来所有标准采用

a) The standard was designed for videophone, video conferencing and other audiovisual services over ISDN.-服务可视电话、视频会议、视听服务

b) The video codec supports bit-rates of p × 64 kbps, where p ranges from 1 to 30 (Hence also known as p ∗ 64).-支持p × 64kbps的码率

c) Require that the delay of the video encoder be less than 150 msec so that the video can be used for real-time bidirectional video conferencing.-编码器延迟必须低于150ms，用于实时双向会议

H.261

- Introduction to H.261-H.261简介
  - ITU Recommendations & H.261 Video Formats-H.261由国际电信联盟标准组织ITU推荐
  - H.261 belongs to the following set of ITU recommendations for visual telephony systems:-H.261属于下列标准

1. H.221 — Frame structure for an audiovisual channel supporting 64 to 1,920 kbps.-H.221支持64~1920kbps视听通道格式
2. H.230 — Frame control signals for audiovisual systems.-视听系统中的帧控制信号
3. H.242 — Audiovisual communication protocols.-视听通信协议
4. H.261 — Video encoder/decoder for audiovisual services at $p \times 64$ kbps.-用于视听服务的视频编码/解码器
5. H.320 — Narrow-band audiovisual terminal equipment for $p \times 64$ kbps transmission.-H.320为窄带视听终端标准

H.261

- Introduction to H.261-H.261简介

Table 10.2 Video Formats Supported by H.261

H.261支持的视频格式

Table 10.2 Video formats supported by H.261

Video format	Luminance image resolution	Chrominance image resolution	Bitrate (Mbps) (if 30 fps and uncompressed)	H.261 support
QCIF	176 × 144	88 × 72	9.1	Required
CIF	352 × 288	176 × 144	36.5	Optional

H.261

- Introduction to H.261-H.261简介

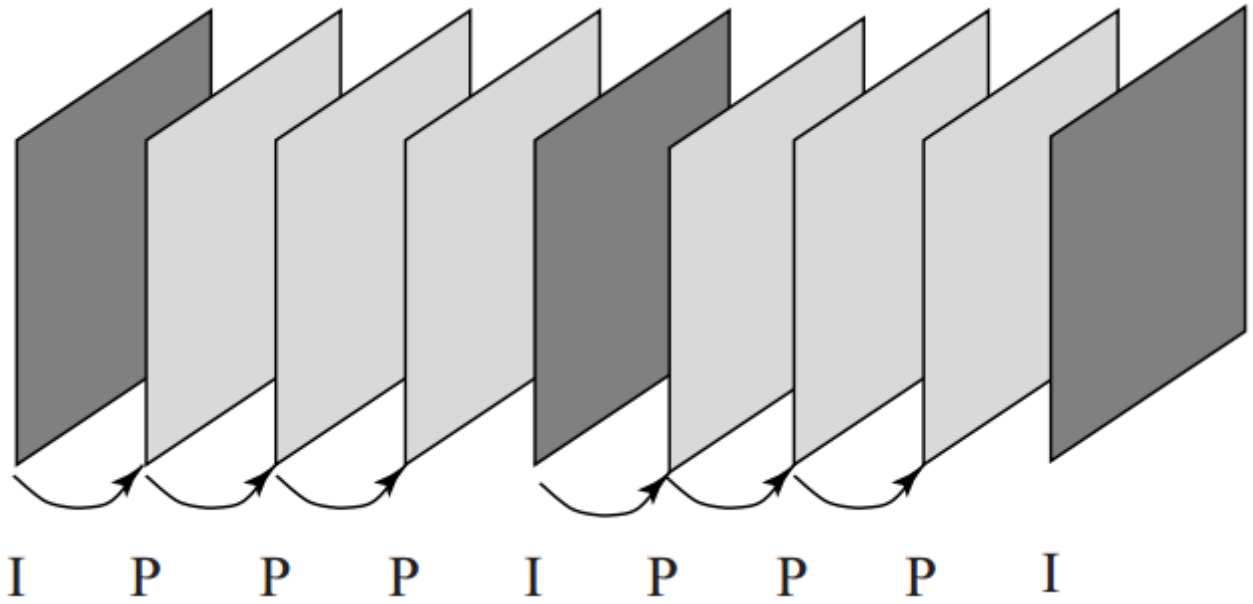


Fig. 10.4: H.261 Frame Sequence.

H.261的帧序列

H.261

◆ Introduction to H.261-H.261简介

- Two types of image frames are defined: Intra-frames (I-frames) and Inter-frames (P-frames):-I帧编码和P帧编码的定义

a) I-frames are treated as independent images. Transform coding method similar to JPEG is applied within each I-frame, hence “Intra”.-I帧是独立图像

b) P-frames are not independent: coded by a forward predictive coding method (prediction from a previous P-frame is allowed | not just from a previous I-frame).-P帧不独立

c) **Temporal redundancy removal** is included in P-frame coding, whereas I-frame coding performs only **spatial redundancy removal**.-P帧可进行时间冗余消除，I帧只能进行空间冗余消除

d) To avoid propagation of coding errors, an I-frame is usually sent a couple of times in each second of the video.-为避免编码误差传播，数字视频每秒发出多个I帧

- Motion vectors in H.261 are always measured in units of full pixel and they have a limited range of ±15 pixels, i.e., p = 15.-运动向量以全部像素为单位进行测量

H.261

- Intra-frame Coding-帧内编码

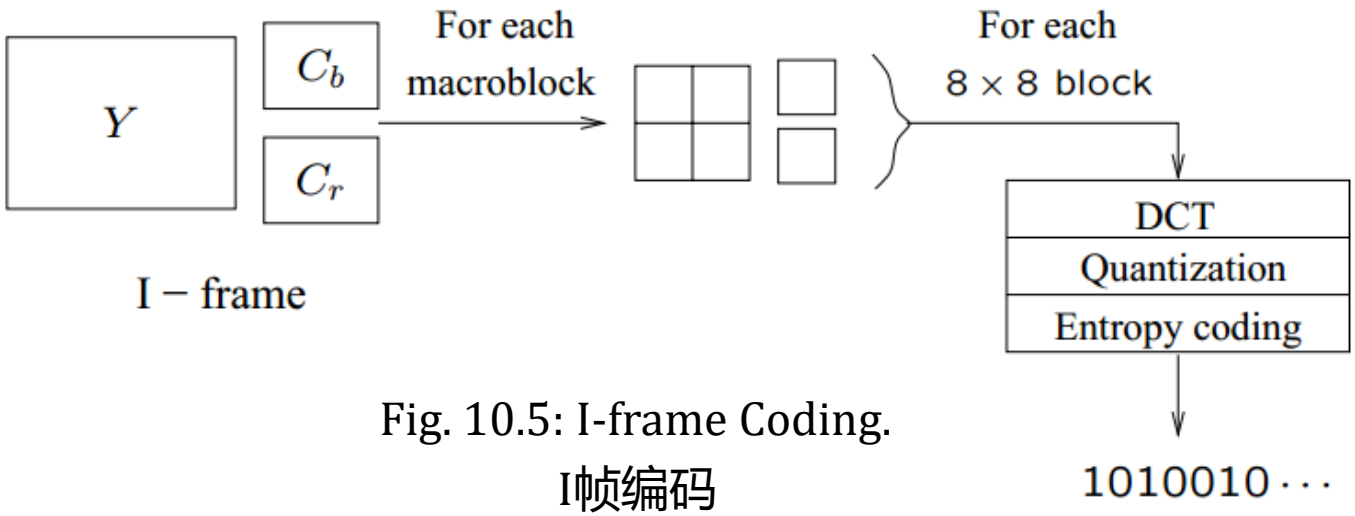


Fig. 10.5: I-frame Coding.

I帧编码

- **Macroblocks** are of size 16 × 16 pixels for the Y frame, and 8 × 8 for Cb and Cr frames, since 4:2:0 chroma subsampling is employed. A mac roblock consists of four Y, one Cb, and one Cr 8 × 8 blocks.-宏块由4个Y块、1个Cb块、1个Cr块组成
- For each 8 × 8 block a DCT transform is applied, the DCT coefficients then go through quantization zigzag scan and entropy coding.-对每个8 × 8 块都要进行DCT变换、量化、熵编码

H.261

- Inter-frame Predictive Coding-帧间预测编码

  - Figure 10.6 shows the H.261 P-frame coding scheme based on motion compensation:-图10.6展示H.261的基于运动补偿的P帧编码

a) For each macroblock in the Target frame, a motion vector is allocated by one of the search methods discussed earlier.-在目标帧宏块内，采用任一运动向量搜索算法定位运动向量

b) After the prediction, a difference macroblock is derived to measure the prediction error.-用差值宏块测量预测误差

c) Each of these 8×8 blocks go through DCT, quantization, zigzag scan and entropy coding procedures.-每一个8×8块都需经历DCT变换、量化、Z字扫描和熵编码

H.261

- Inter-frame Predictive Coding-帧间预测编码

  - The P-frame coding encodes the difference macroblock (not the Target macroblock itself).-对差值宏块进行P帧编码

  - Sometimes, a good match cannot be found, i.e., the prediction error exceeds a certain acceptable level.-可能出现预测误差不在接受范围内

  a) The MB itself is then encoded (treated as an Intra MB) and in this case it is termed a non-motion compensated MB.-需要对预测误差大的宏块本身进行编码，即未进行运动补偿

  - For motion vector, the difference MVD is sent for entropy coding:-运动向量是对差值MVD进行熵编码

$$MVD = MV_{Preceding} - MV_{Current} \qquad (10.3)$$

H.261

- Inter-frame Predictive Coding-帧间预测编码

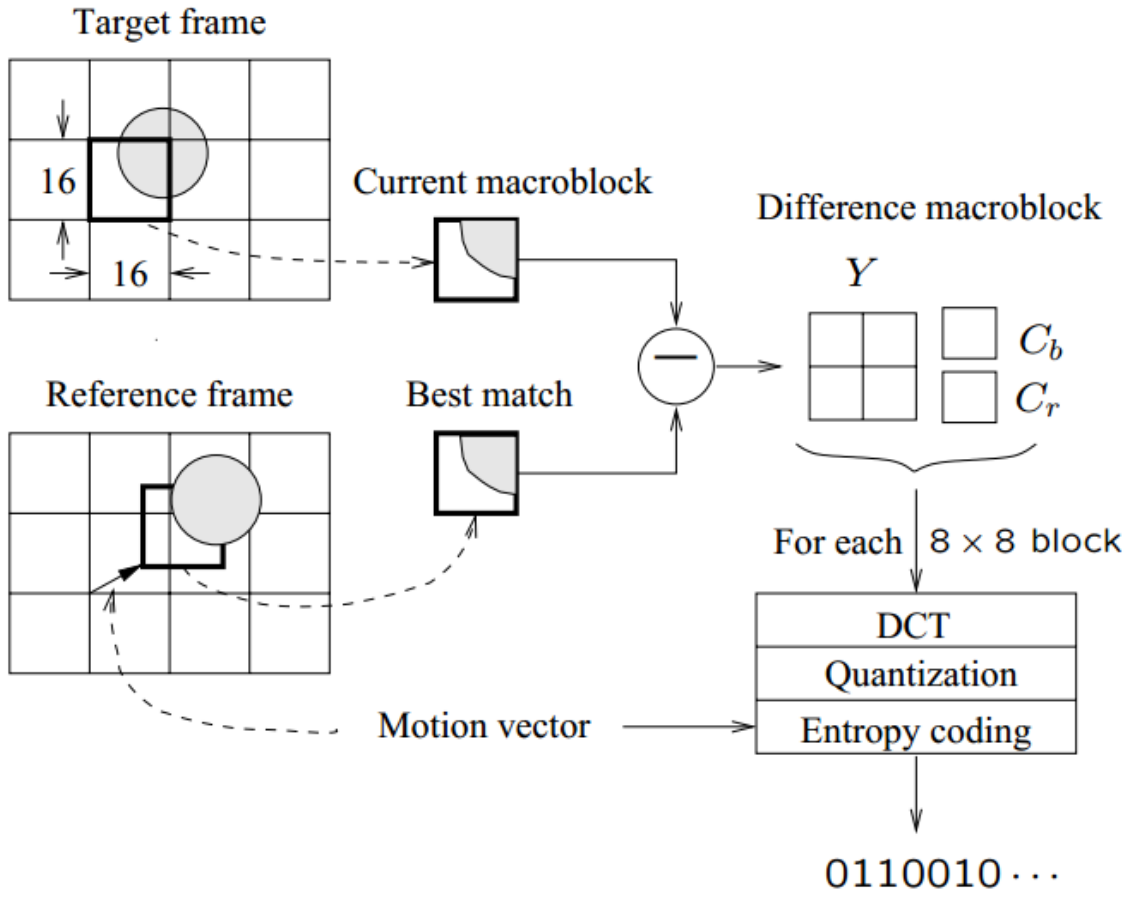


Fig. 10.6: H.261 P-frame Coding Based on Motion Compensation.

H.261中基于运动补偿的P帧编码

H.261

- Quantization in H.261-H.261中的量化
  - The quantization in H.261 uses a constant step size, for all DCT coefficients within a macroblock.-H.261中采用固定步长对宏块所有DCT系数进行量化
  - If we use DCT and QDCT to denote the DCT coefficients before and after the quantization, then for DC coefficients in Intra mode:-帧内编码的DC系数量化如下

$$QDCT = round\left(\frac{DCT}{step\_size}\right) = round(\frac{DCT}{8}) \qquad (10.4)$$

for all other coefficients:-对于其他情况的系数的量化如下

$$QDCT = \left\lfloor\frac{DCT}{step\_size}\right\rfloor = \left\lfloor\frac{DCT}{2*scale}\right\rfloor \qquad (10.5)$$

scale — an integer in the range of [1, 31]-整形尺度因子

H.261

- H.261 Encoder and Decoder-H.261编码器和解码器
  - Fig. 10.7 shows a relatively complete picture of how the H.261 encoder and decoder work.-图10.7展示H.261的编码和解码过程
  - A scenario is used where frames I, P_1, and P_2 are encoded and then decoded.-假设有一个I帧，两个P帧（P1, P2）需要进行编解码
  - Note: decoded frames (not the original frames) are used as reference frames in motion estimation.-解码的帧在运动补偿中将作为参考帧
  - The data that goes through the observation points indicated by the circled numbers are summarized in Tables 10.3 and 10.4.-表中展示数据经过观察点的变化情况

H.261

- H.261 Encoder and Decoder-H.261编码器和解码器

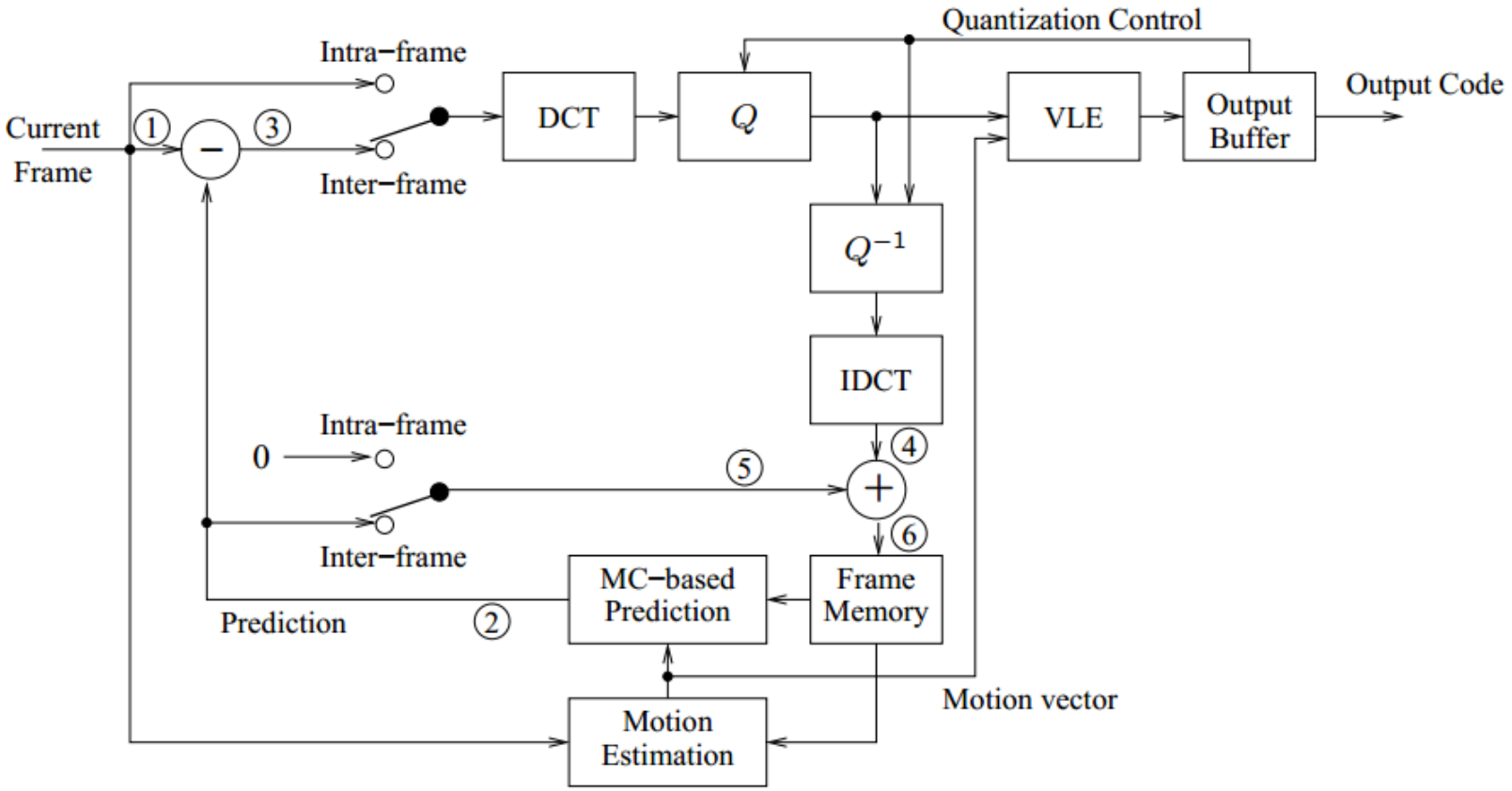


Fig. 10.7: H.261 Encoder and Decoder.

H.261的编码和解码

H.261

- H.261 Encoder and Decoder-H.261编码器和解码器

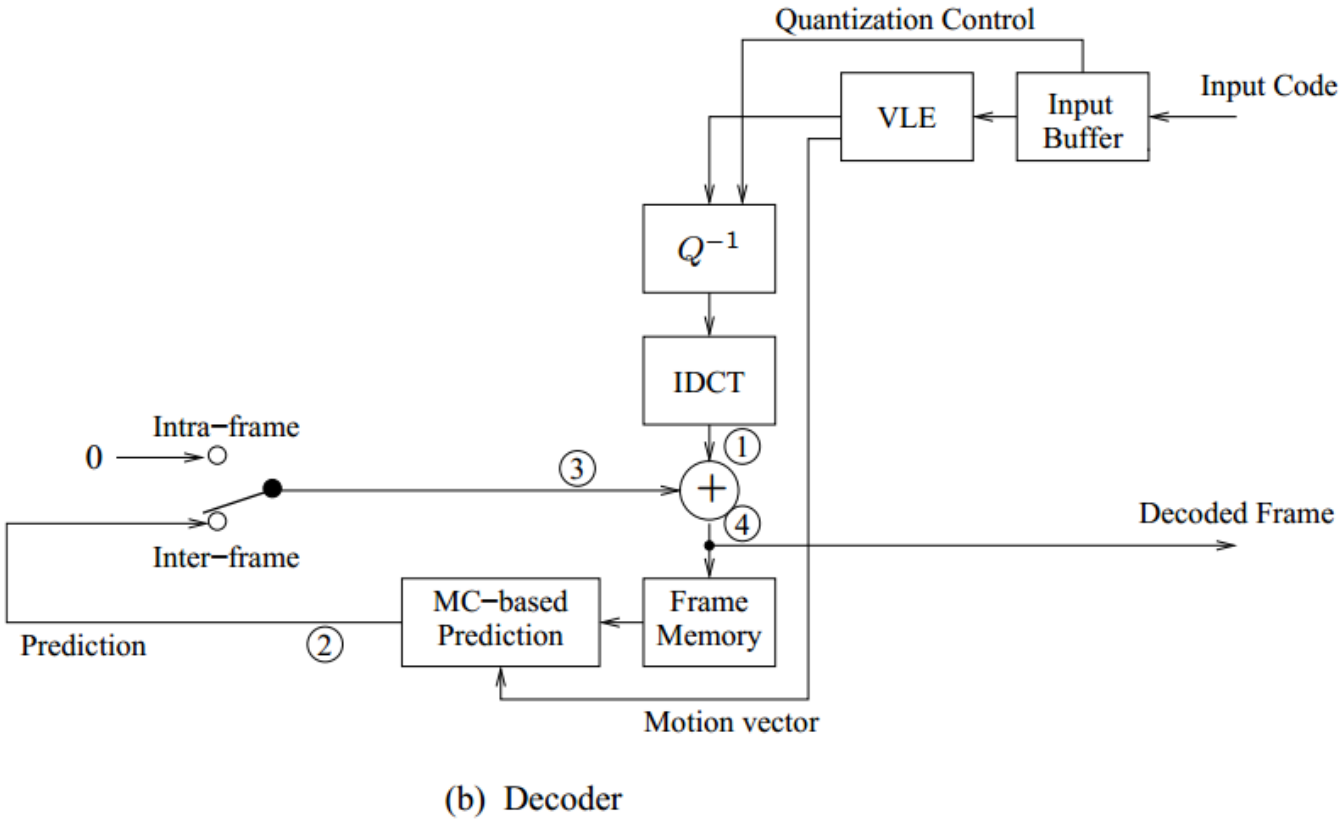


Fig. 10.7 (Cont'd): H.261 Encoder and Decoder.

H.261的编码和解码

H.261

- H.261 Encoder and Decoder-H.261编码器和解码器

Table 10.3: Data Flow at the Observation Points in H.261 Encoder
H.261编码过程中观测点的数据流变化情况

Current Frame	Observation Point					
	1	2	3	4	5	6
I	I			$\tilde{I}$	0	$\tilde{I}$
P_1	P_1	P_1'	D_1	$\tilde{D}_1$	P_1'	$\tilde{P}_1$
P_2	P_2	P_2'	D_2	$\tilde{D}_2$	P_2'	$\tilde{P}_2$

Table 10.4: Data Flow at the Observation Points in H.261 Decoder
H.261解码过程中观测点的数据流变化情况

Current Frame	Observation Point			
	1	2	3	4
I	$\tilde{I}$		0	$\tilde{I}$
P_1	$\tilde{D}_1$	P_1'	P_1'	$\tilde{P}_1$
P_2	$\tilde{D}_2$	P_2'	P_2'	$\tilde{P}_2$

H.261

- A Glance at Syntax of H.261 Video Bitstream-H.261视频位流语法概述

  - Fig. 10.8 shows the syntax of H.261 video bitstream: a hierarchy of four layers: Picture, Group of Blocks (GOB), Macroblock, and Block.-H.261视频位流的语法

a) **The Picture layer:** PSC (Picture Start Code) delineates boundaries between pictures. TR (Temporal Reference) provides a time-stamp for the picture.-图像层

b) **The GOB layer:** H.261 pictures are divided into regions of 11 × 3 macroblocks, each of which is called a Group of Blocks (GOB).-块组层

- Fig. 10.9 depicts the arrangement of GOBs in a CIF or QCIF luminance image.-CIF和QCIF亮度图像中块组的排列
- For instance, the CIF image has 2 × 6 GOBs, corresponding to its image resolution of 352×288 pixels. Each GOB has its Start Code (GBSC) and Group number (GN).-CIF的块组层包含起始码和组号

H.261

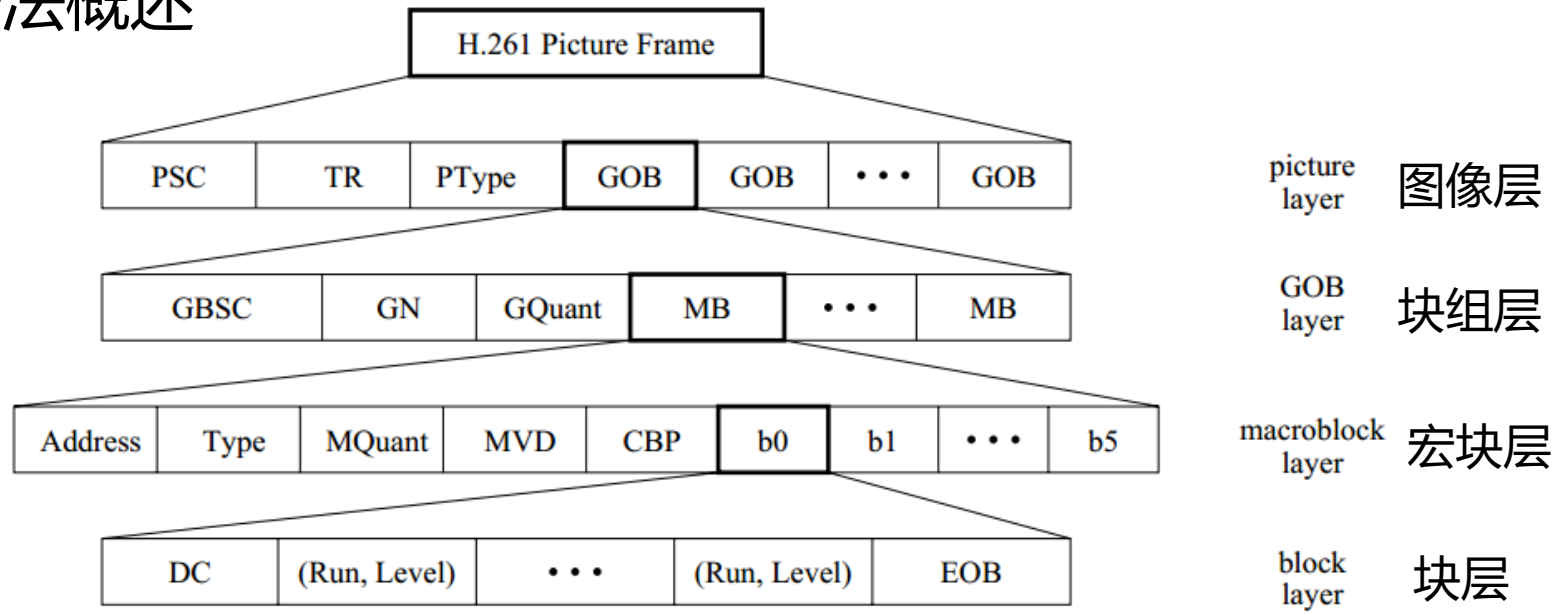
- A Glance at Syntax of H.261 Video Bitstream-H.261视频位流语法概述
  - In case a network error causes a bit error or the loss of some bits, H.261 video can be recovered and resynchronized at the next identifiable GOB.-数据丢失情况下，起始码可以帮助数据恢复和同步
  - GQuant indicates the Quantizer to be used in the GOB unless it is overridden by any subsequent MQuant (Quantizer for Macroblock). GQuant and MQuant are referred to as scale in Eq. (10.5).-Gquant和Mquant是量化器

c) **The Macroblock layer:** Each Macroblock (MB) has its own Address indicating its position within the GOB, Quantizer (MQuant), and six 8 × 8 image blocks (4 Y, 1 Cb, 1 Cr).-宏块有自己地址

d) **The Block layer:** For each 8 × 8 block, the bitstream starts with DC value, followed by pairs of length of zero un (Run) and the subsequent non-zero value (Level) for ACs, and finally the End of Block (EOB) code. The range of Run is [0; 63]. Level reflects quantized values — its range is [−127,127] and Level 6≠ 0.-块层的位流从DC值开始

H.261

- A Glance at Syntax of H.261 Video Bitstream-H.261视频位流语法概述



PSC:	Picture Start Code	TR:	Temporal Reference
PType:	Picture Type	GOB:	Group of Blocks
GBSC:	GOB Start Code	GN:	Group Number
GQuant:	GOB Quantizer	MB:	Macro Block
MQuant:	MB Quantizer	MVD:	Motion Vector Data
CBP:	Coded Block Pattern	EOB:	End of Block

Fig. 10.8: Syntax of H.261 Video Bitstream.

H.261视频位流语法

H.261

- A Glance at Syntax of H.261 Video Bitstream-H.261视频位流语法概述

GOB 0	GOB 1
GOB 2	GOB 3
GOB 4	GOB 5
GOB 6	GOB 7
GOB 8	GOB 9
GOB 10	GOB 11

CIF

GOB 0
GOB 1
GOB 2

QCIF

Fig. 10.9: Arrangement of GOBs in H.261 Luminance Images.

H.261亮度图像块组设置

Outline of Lecture 12

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

H.263

- Introduction to H.263-H.263简介

  - H.263 is an improved video coding standard for video conferencing and other audiovisual services transmitted on Public Switched Telephone Networks (PSTN). –H.263服务于视频会议和其他公共电话网络视听服务

  a) Aims at low bit-rate communications at bit-rates of less than 64 kbps.-码率尽可能在64kbps以下

  b) Uses predictive coding for inter-frames to reduce temporal redundancy and transform coding for the remaining signal to reduce spatial redundancy (for both Intra-frames and inter-frame prediction).-相邻帧采用预测编码消除时间冗余，其他情况采用信号变换消除空间冗余

H.263

- Introduction to H.263-H.263简介

Table 10.5 Video Formats Supported by H.263

H.263支持的视频格式

Video format	Luminance image resolution	Chrominance image resolution	Bit-rate (Mbps) (if 30 fps and uncompressed)	Bit-rate (kbps) BPPmaxKb (compressed)
sub-QCIF	128 × 96	64 × 48	4.4	64
QCIF	176 × 144	88 × 72	9.1	64
CIF	352 × 288	176 × 144	36.5	256
4CIF	704 × 576	352 × 288	146.0	512
16CIF	1,408 × 1,152	704 × 576	583.9	1024

H.263

- Group of Blocks (GOB) of H.263-H.263的组块
  - As in H.261, H.263 standard also supports the notion of Group of Blocks (GOB).-H.263同样支持组块
  - The difference is that GOBs in H.263 do not have a fixed size, and they always start and end at the left and right borders of the picture.-H.263组块没有固定大小，起止于图像左右边界
  - As shown in Fig. 10.10, each QCIF luminance image consists of 9 GOBs and each GOB has 11 × 1 MBs (176 × 16 pixels), whereas each 4CIF luminance image consists of 18 GOBs and each GOB has 44 × 2 MBs (704 × 32 pixels).-图10.10展示QCIF和4CIF的组块情况

H.263

- Group of Blocks (GOB) of H.263-H.263的组块

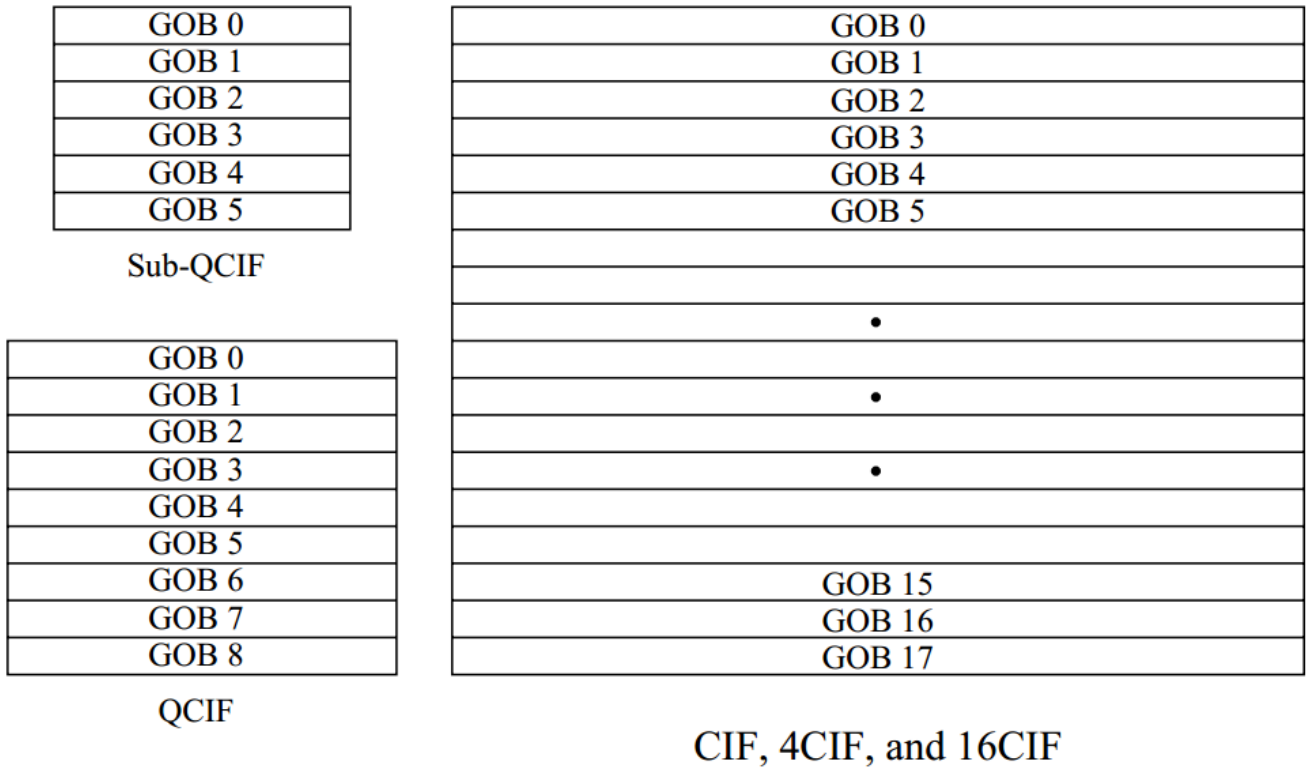


Fig. 10.10 Arrangement of GOBs in H.263 Luminance Images.

H.263中亮度图像的组块设置

H.263

- Motion Compensation in H.263-H.263的运动补偿
  - The horizontal and vertical components of the MV are predicted from the median values of the horizontal and vertical components, respectively, of MV1, MV2, MV3 from the “previous”, “above” and “above and right” MBs (see Fig.10.11 (a)).-H.263运动向量估计由当前宏块左边、上方、右上方宏块的运动分量共同决定
  - For the Macroblock with MV(u; v):-对于宏块MV(u; v)运动向量计算如下

$$u_p = median(u_1, u_2, u_3)$$

$$v_p = median(v_1, v_2, v_3) \qquad (10.6)$$

  - Instead of coding the MV(u, v) itself, the error vector (δu, δv) is coded, where δu = u – u_p and δv = v – v_p.-对误差向量进行编码

H.263

- Motion Compensation in H.263-H.263的运动补偿

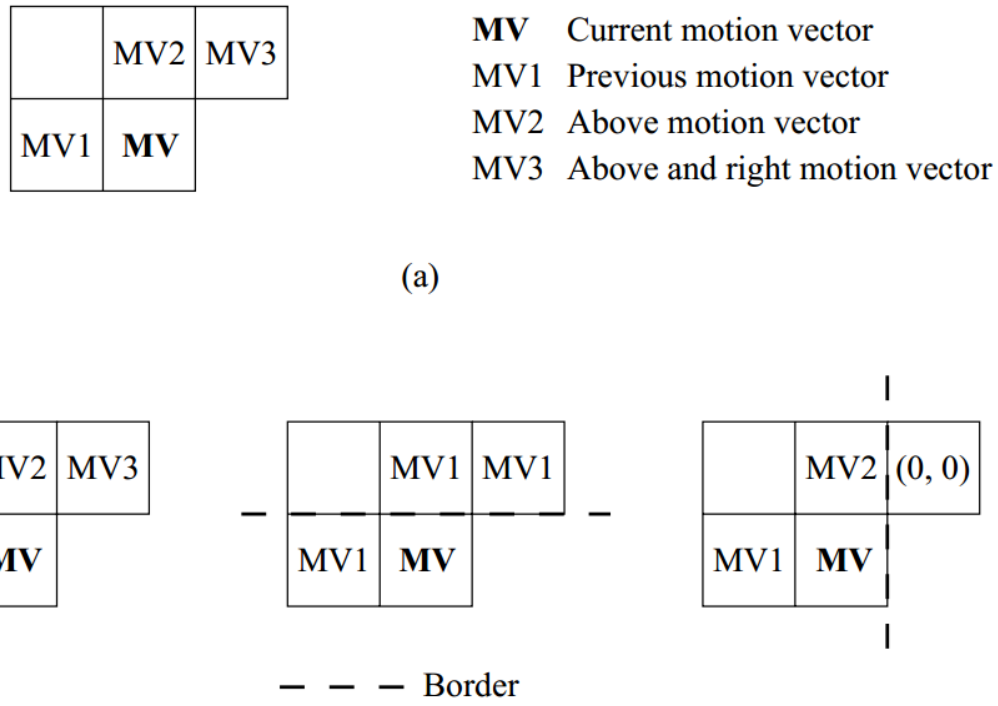


Fig. 10.11 Prediction of Motion Vector in H.263.

H.263中运动向量的预测

H.263

- H.263 supports Half-Pixel Precision-支持半像素精度
  - In order to reduce the prediction error, half-pixel precision is supported in H.263 vs. full-pixel precision only in H.261.-H.263 支持半像素精度

a) The default range for both the horizontal and vertical components u and v of MV(u; v) are now [−16; 15:5].-运动向量水平和垂直分量有固定范围

b) The pixel values needed at half-pixel positions are generated by a simple bilinear interpolation method, as shown in Fig. 10.12.-通过双线性插值获得半像素位置像素值

H.263

- H.263 supports Half-Pixel Precision-支持半像素精度

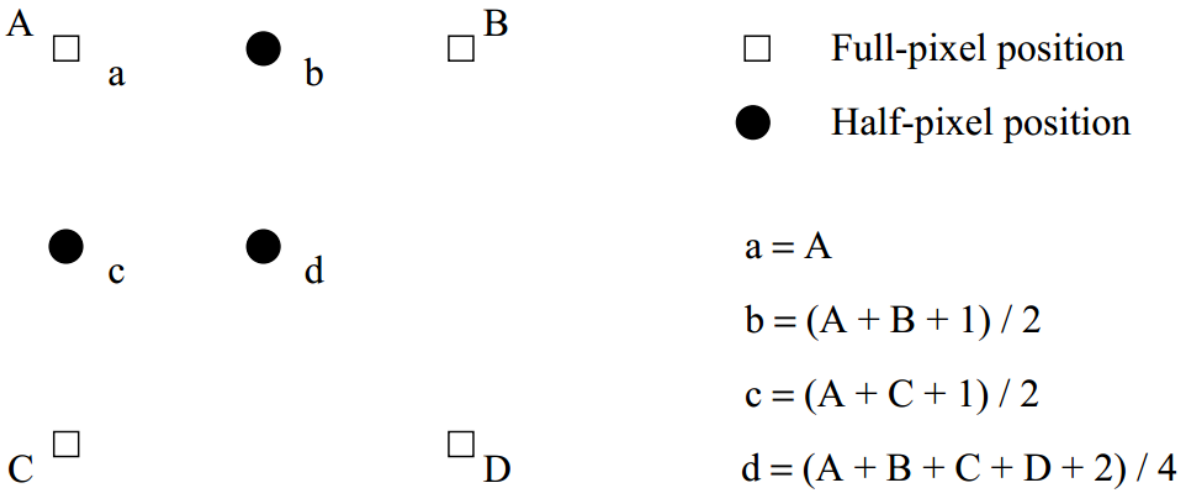


Fig. 10.12: Half-pixel Prediction by Bilinear Interpolation in H.263.

H.263中通过双线性插值进行半像素精度的预测

H.263

- Optional H.263 Coding Modes-H.263可选的编码模式
  - H.263 specifies many negotiable coding options in its various Annexes. Four of the common options are as follows:-H.263有4种可选的编码模式

a) Unrestricted motion vector mode:-无限制的运动向量模式

- The pixels referenced are no longer restricted to be within the boundary of the image.-参考像素不再限制在图像边界内
- When the motion vector points outside the image boundary, the value of the boundary pixel that is geometrically closest to the referenced pixel is used.-利用最靠近参考像素的像素点的值计算图像外的参考像素值
- The maximum range of motion vectors is [-31.5, 31.5].-运动向量有固定范围

H.263

- Optional H.263 Coding Modes-H.263可选的编码模式

b) Syntax-based arithmetic coding mode:-基于语法的算术编码模式

- As in H.261, variable length coding (VLC) is used in H.263 as a default coding method for the DCT coefficients.-H.263采用变长编码方法进行编码
- Similar to H.261, the syntax of H.263 is also structured as a hierarchy of four layers. Each layer is coded using a combination of fixed length code and variable length code.-H.263语法也有4层，结合定长或变长编码算法进行编码

c) Advanced prediction mode:-高级预测模式

- In this mode, the macroblock size for MC is reduced from 16 to 8.-运动补偿中的宏块大小变小
- Four motion vectors (from each of the 8 × 8 blocks) are generated for each macroblock in the luminance image.-每个宏块产生4个运动向量

H.263

- Optional H.263 Coding Modes-H.263可选的编码模式

d) PB-frames mode:-PB帧模式

- In H.263, a PB-frame consists of two pictures being coded as one unit, as shown Fig. 10.13.-一个 PB帧由两个编码为一个单元的图片组成
- The use of the PB-frames mode is indicated in PTYPE.-PTYPE说明PB帧模式的使用
- The PB-frames mode yields satisfactory results for videos with moderate motions.-PB帧模式在视频图像变化不太大时效果较好
- Under large motions, PB-frames do not compress as well as B-frames and an improved new mode has been developed in Version 2 of H.263.-PB帧模式在视频图像变化较大情况效果不如B帧模式压缩效果好

H.263

- Optional H.263 Coding Modes-H.263可选的编码模式

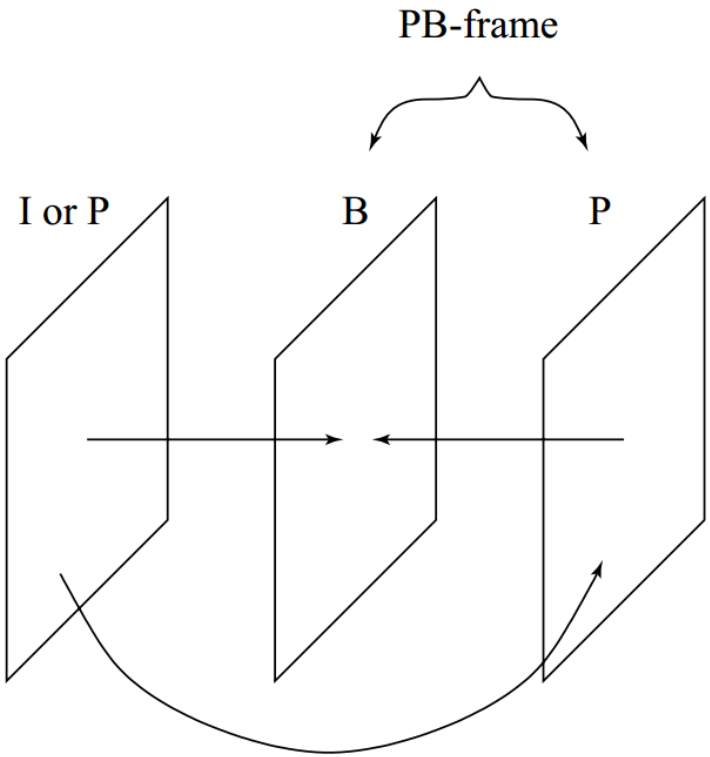


Fig. 10.13: A PB-frame in H.263.
H.263中的PB帧

H.263

◆ H.263+ and H.263++

- The aim of H.263+: broaden the potential applications and offer additional flexibility in terms of custom source formats, different pixel aspect ratio and clock frequencies.-H.263+应用性更强，更加灵活

- H.263+ provides 12 new negotiable modes in addition to the four optional modes in H.263.-提供12种全新的可选模式

a) It uses Reversible Variable Length Coding (RVLC) to encode the difference motion vectors.-采用可逆变长编码方法对运动向量差编码

b) A slice structure is used to replace GOB to offer additional flexibility.-用宏块片结构替代块组

H.263

- **H.263+ and H.263++**

c) H.263+ implements *Temporal, SNR,* and *Spatial scalabilities.*-实现时间、空间、信噪比的可伸缩性

d) Support of Improved PB-frames mode in which the two motion vectors of the B-frame do not have to be derived from the forward motion vector of the P-frame as in Version 1.-H.263+支持改善的PB帧模式

e) H.263+ includes deblocking filters in the coding loop to reduce blocking effects.-包含解块过滤器，减少编码阻塞

H.263

◆ H.263+ and H.263++

- H.263++ includes the baseline coding methods of H.263 and additional recommendations for Enhanced Reference Picture Selection (ERPS), Data Partition Slice (DPS), and Additional Supplemental Enhancement Information.-H.263++是更为先进的版本

a) The ERPS mode operates by managing a multi-frame buffer for stored frames - enhances coding efficiency and error resilience capabilities. -利用多帧缓冲区提高编码效率和错误恢复能力

b) The DPS mode provides additional enhancement to error resilience by separating header and motion vector data from DCT coefficient data in the bitstream and protects the motion vector data by using a reversible code.-DPS提供了更强的错误恢复能力

Outline of Lecture 12

- Introduction to Video Compression-视频压缩简介
- Video Compression with Motion Compensation-基于运动补偿的视频压缩
  - Three Main Steps of Video Compression-视频压缩三个主要步骤
  - Motion Compensation-运动补偿
- Search for Motion Vectors-搜索运动向量
  - Sequential Search-顺序搜索
  - 2D Logarithmic Search-2D对数搜索
  - Hierarchical Search-分层搜索
- H.261
  - Introduction to H.261-H.261简介
  - Intra-frame Coding-帧内编码
  - Inter-frame Predictive Coding-帧间预测编码
  - H.261 Encoder and Decoder-H.261编码器和解码器
- H.263
- Experiments-实验

56

Experiments & Class Assignments

- Experiments

  - 第10章程序10.2 运动向量：2D对数搜索

- Class Assignments

  1、比较H.261和H.263的相同以及不同之处。

  2、编程实现运动向量的分层搜索（语言不限）。

期末考试题型

- 单选题
- 简答题
- 计算分析题
- 综合论述题
- 附加题