MPEG-1:
A Standard for Digital Storage of Audio and Video

Nimrod Peleg

Update: Feb. 2005

MPEG-1 Target (1988-1990)

- MPEG-1 has been developed for storage of CIF format video and its associated audio at about 1.5Mbps on various digital storage media such as *CD-ROM*, DAT, Winchester disks, Optical drives etc.
- Primary application perceived as Multimedia systems (Similar quality to VHS)
- Similar to H.261 + Additional features

Main Features

- Standartizes a Syntax that supports ME, MC, DCT, Quantization, VLC etc.
- Does not define specific algorithms needed to produce valid data
- A number of parameters, contained in the coded bitstream, allowing more flexibility than H.261

Application Specific Features

- <u>(No) Random Access</u> to any frame in a limited amount of time
- <u>Fast Forward/Reverse Search</u> to display only selected frames (also Freeze mode)
- <u>Coding/Decoding Delay</u> of about 1Sec (vs. 150mSec in H.261)

Input Video Format

- Progressive (Non-Interlaced) video only
- Input format usually converted to *Standard Input Format (SIF)*:

352x240, 8b/pixel, Chroma subsampled by 2 in both axis (Similar to H.261), 30fps

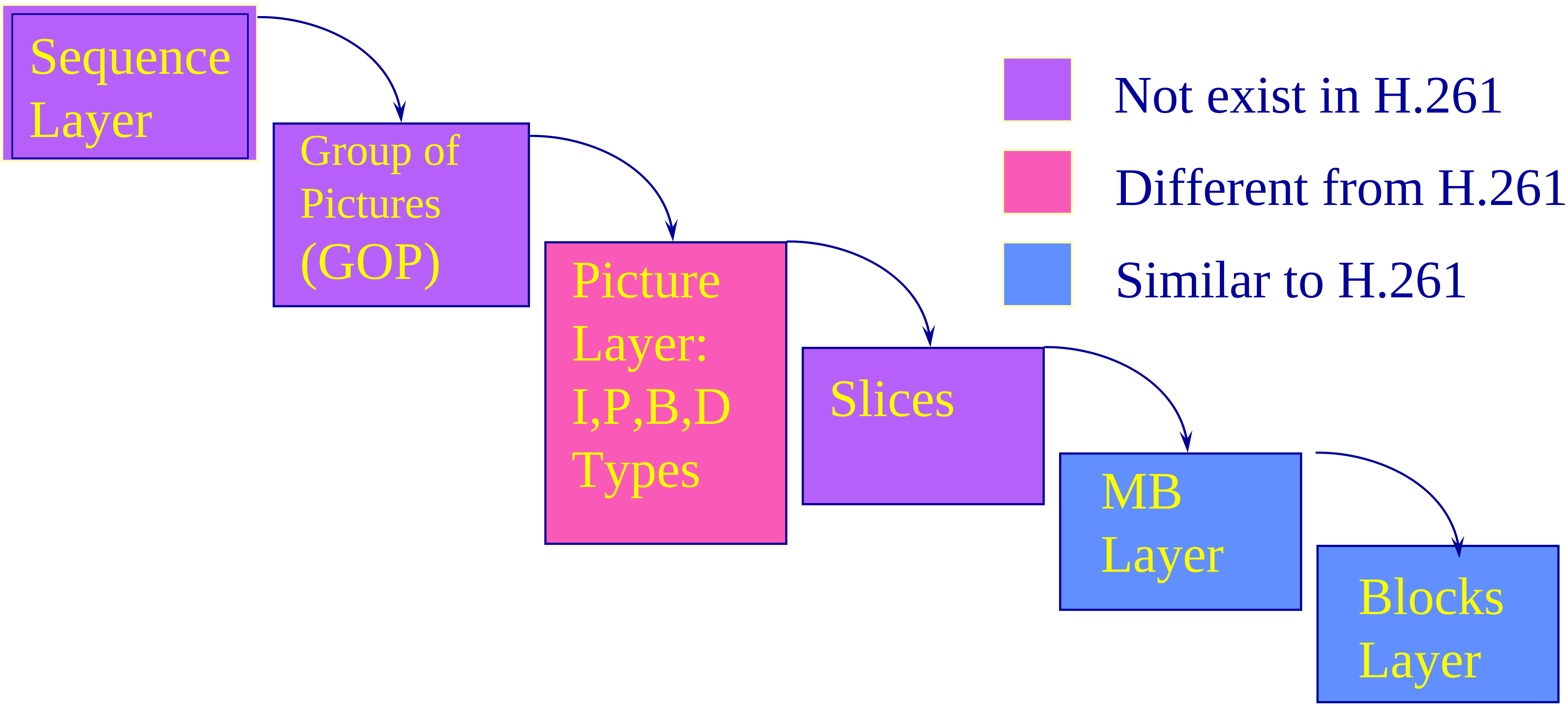
- Color space (YCbCr) adopted from CCIR-601

Constrained Parameters

• Max. Horizon. resolution:	720 pixels/line
• Max. Vertical resolution:	576 lines/pict.
• Max. Temporal rate:	30 frames/sec.
• Max. number of MB/picture:	396
• Max. MB rate:	9900 MB/sec.
• Max. Bitrate:	1.86 Mbps
• Max. Decoder buffer size:	376,832 bits
• Max. MV Range:	-64 to +63.5 pels

Data Structure

Hierarcical structure, with similarity to H.261



Compression Modes

- *Sequences*: Several GOPs,
- *Group of Pictures (GOP)*: Smallest unit that can be independently decoded
- *Pictures:* 4 compression modes defined
  - *Intra (I)*: No reference to any other picture, JPEG-like coding, serve as "random access" points
  - *Predicted (P)*: MC prediction errors are coded. Forward prediction from previous I or P frames

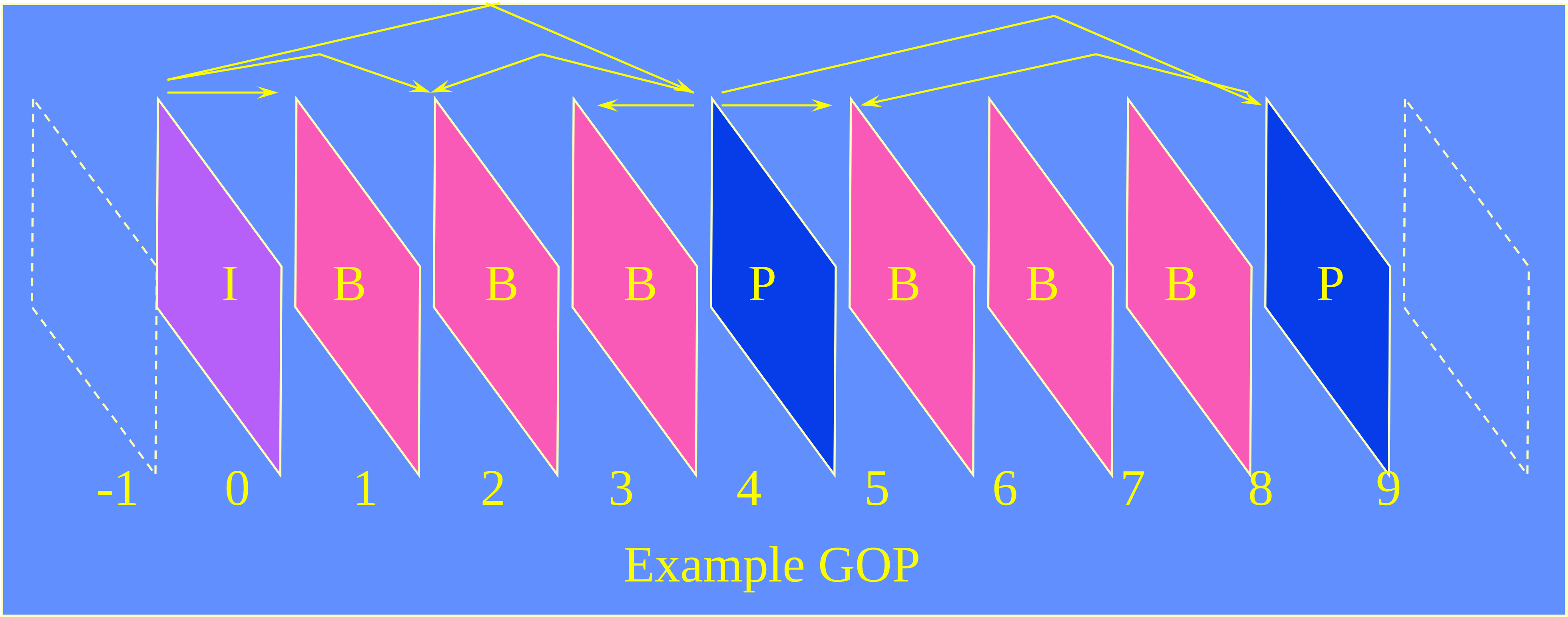
Compression Modes (Cont'd)

– *Bi-directional (B)*: MC prediction errors are coded. Forward/Backward/Bidirctional prediction from both previous and future I or P

– *DC (D)*: Contains only DC component of each block, serve fast forward search mode, for very low bitrates

The number of I,P,B frames in a GOP are application-dependent (see next slide)

GOP Arrangement



- Encoding order options: 0,4,1,2,3,8,5,6,7

  or 0,1,4,2,3,8,5,6,7

Compression Modes (Cont'd)

- *Slices*: Made up of *MacroBlocks*. One or more in a picture, provide some header information for a fast error recovery
- *MacroBlocks (MB)*: Basic unit for MC and Quantizer table. Composition similar to H.261 (see next slides for MB types)
- *Block:* Smallest DCT unit, 8x8 pixel array

MB Types

I-Picture	P-Picture	B-Picture
Intra	Intra	Intra
Intra-A	Intra-A	Intra-A
	Inter-D	Inter-F
	Inter-DA	Inter-FD
	Inter-F	Inter-FDA
	Inter-FD	Inter-B
	Inter-FDA	Inter-BD
	Skipped	Inter-BDA
		Inter-I
		Inter-ID
		Inter-IDA
		Skipped

Intraframe Compression Modes

- For 8-bit input image, DCT Coeff. is 11-bit, [0,2040] for DC and [-1024,1023] for AC
- Quantized coeff. obtained by dividing DCT coeff. value by quantized step size + rounding
- Default Intra Q-Table:

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

Intraframe Types

- *Intra*: MBs are coded with current quantization table
- *Intra-A*: Quantization table is scaled by MQuant (transmitted in the header), that can be varied on a MB basis

HVS suggests that "busy" MBs can be quantized relatively coarsely

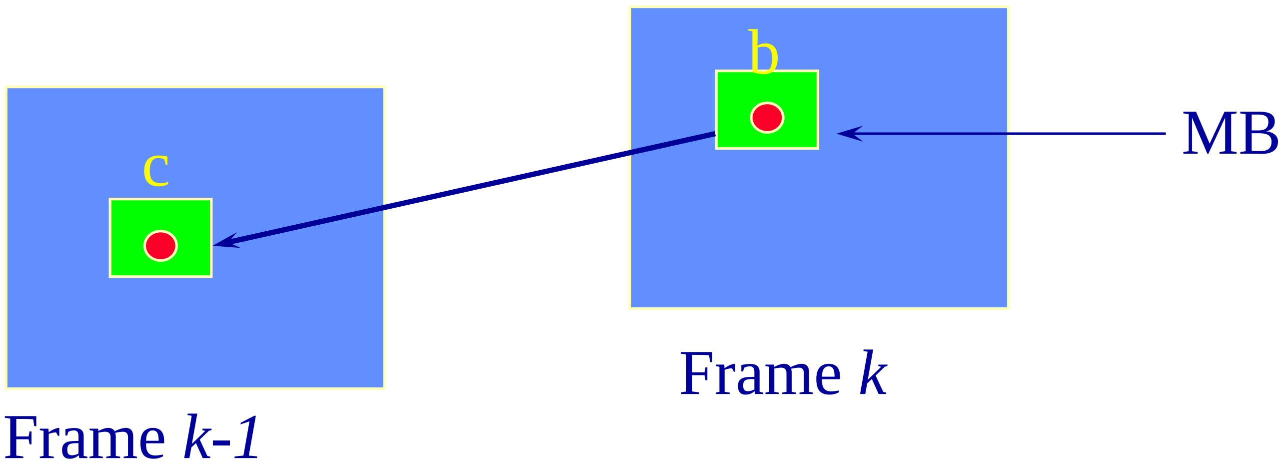
Due to adaptive quantization, MPEG (Intra) provides 30% better compression compared with JPEG ! (No adaptation in JPEG)

Intraframe Coding

- DC coeff. are DPCM coded with a fixed Huffman table (with logarithmic amplitude - similar to JPEG)
- AC coeff. are Zig-Zag scanned and converted into run-length pairs (similar to JPEG)
- A single code table is used for all blocks. Only highly-probable pairs are VLC coded, and the rest with a fixed length code, to avoid extremely long codewords
- Codebook is superset of H.261 (Not JPEG !)

Interframe Compression Modes

- <u>P-Pictures</u>: Forward prediction, with reference to previous I or P pictures:



$$\hat{b} = \tilde{c}$$

MB corresponding to b in the reconstructed previous frame

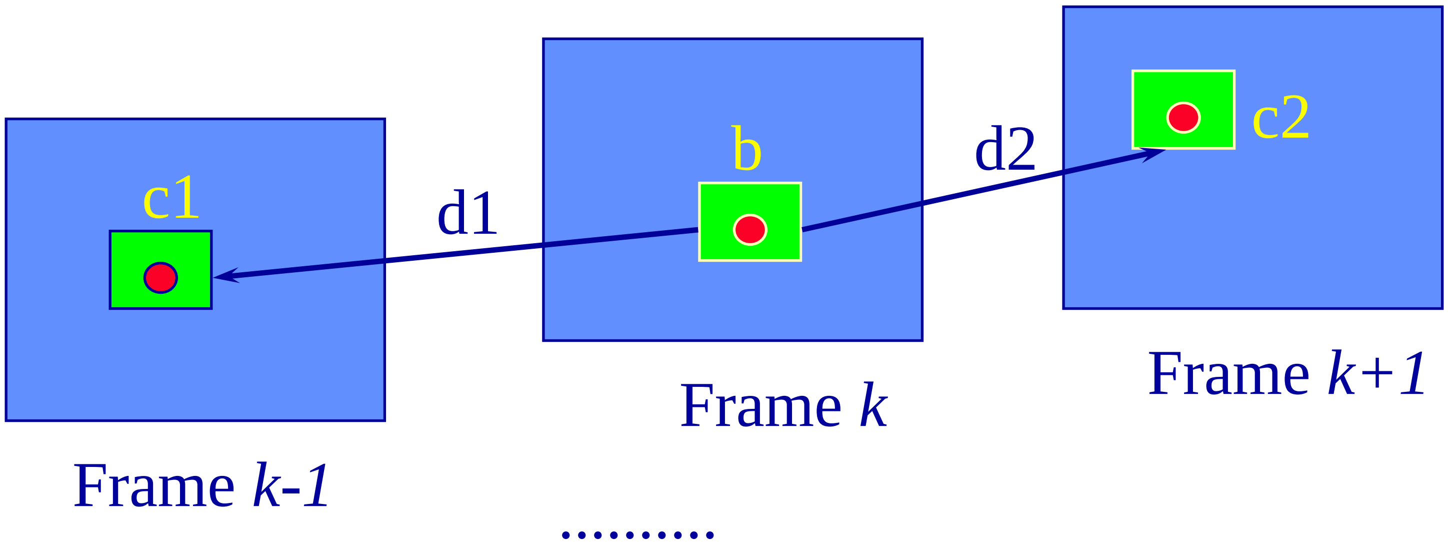
P-Interframe Types

- *Intra / Intra-A* : same as in *Intraframe*
- *Inter-D*: DCT of prediction error will be coded
- *Inter-F*: Forward MC active
- *Inter-A*: Adaptive quantization (new MQuant)
- *Skipped*: If the MB at the same position in the previous frame (without MC) is good enough (stationary area)

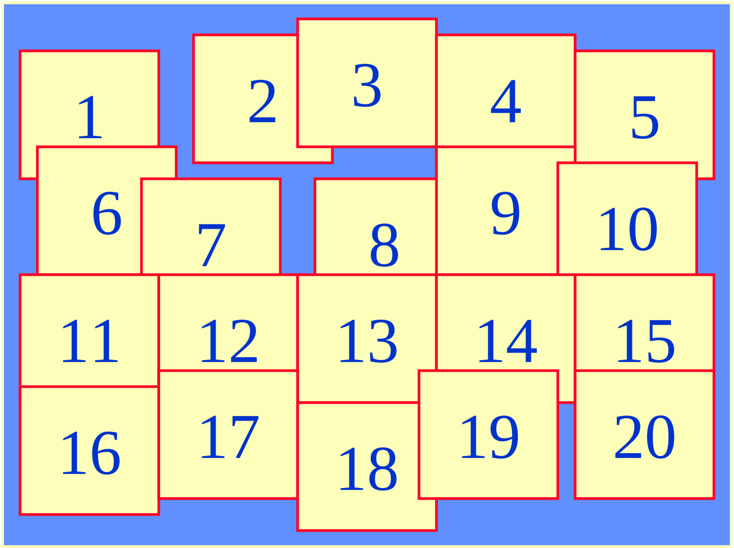
Note: D/F/A combinations also possible

Interframe Modes (Cont'd)

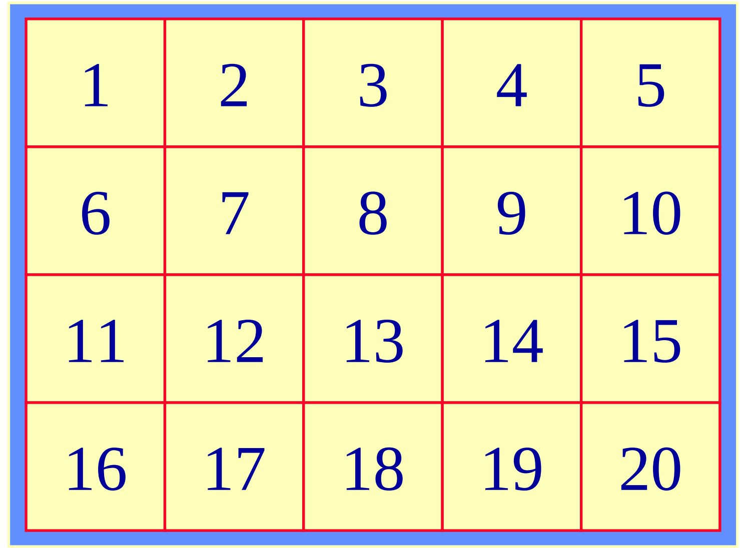
- B-Picture: Allows interpolative coding (Bi-directional prediction)



Backward Prediction



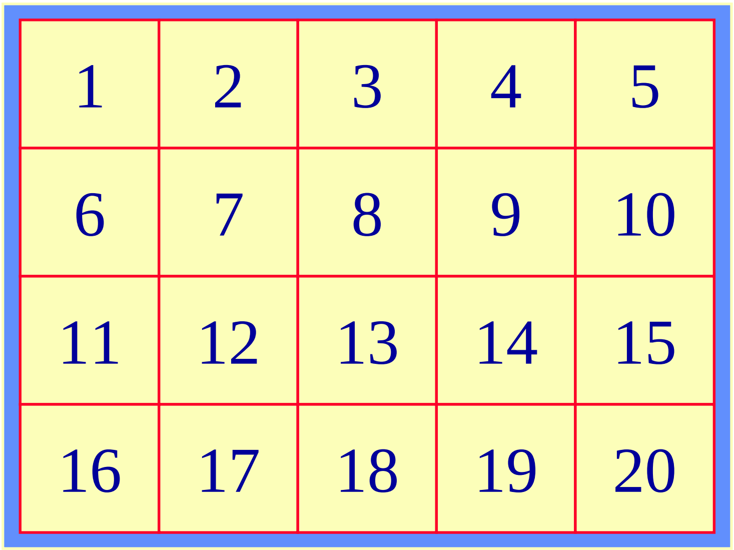
Previous Frame



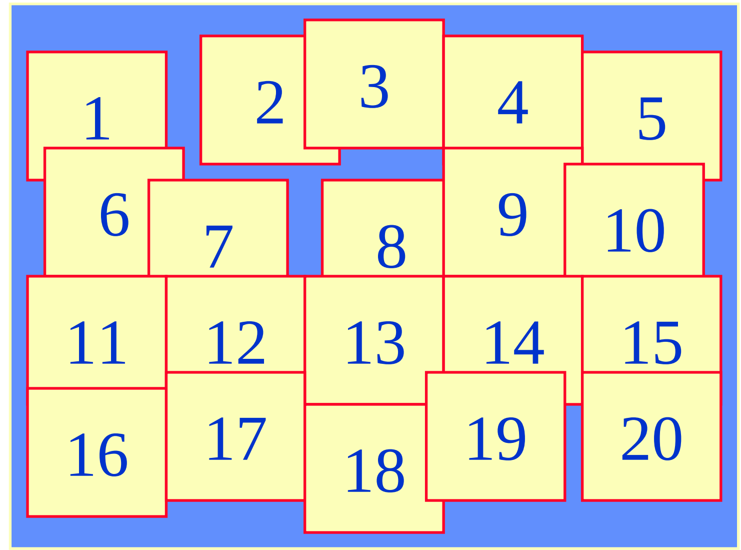
Predicted (current) Frame

- No holes or overlap is created in the predicted image
- All standards use this approach
- Introduces less delays

Forward Prediction



Current Frame



Predicted (future) Frame

Holes or overlap regions are created in the predicted image

B-Picture Prediction Types

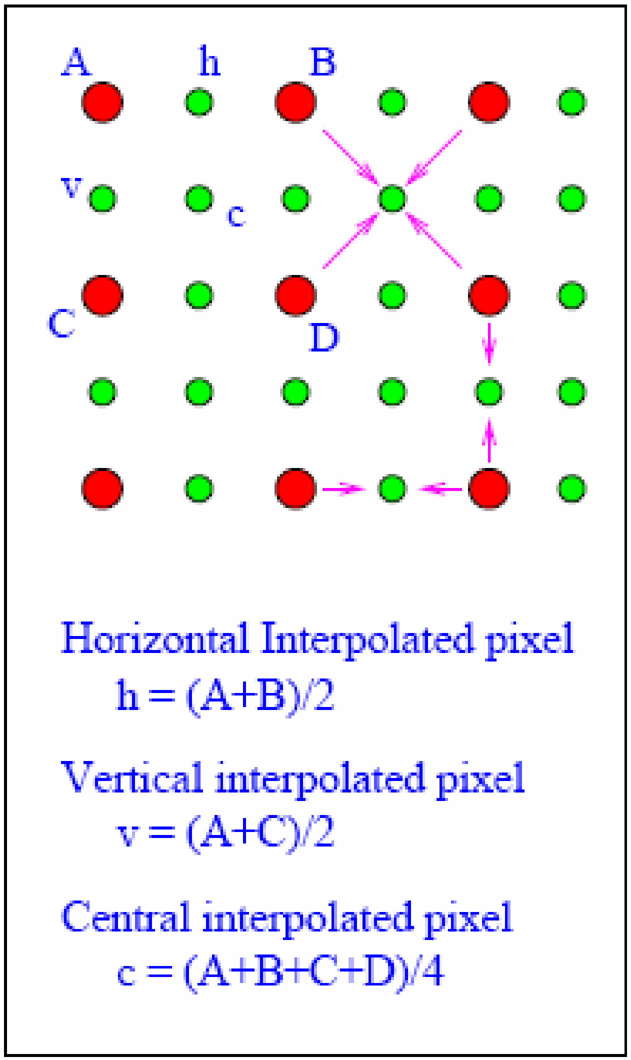
- $a1=0$; $a2=1$: Backward prediction
- $a1=1$; $a2=0$: Forward prediction

- $a1=a2=0.5$: Bidirectional prediction

For this type, 2 displacement vectors (d1, d2) and prediction error (b-b~) need to be coded for each MB b.

Half-Pel refinement

- motion estimation from previous reconstructed and interpolated frame



Half-Pel Refinement Example

Search Window

−2	86	33	76	94	55	11
−1	68	54	62	34	77	59
0	10	16	28	73	82	82
+1	65	68	71	75	54	96
+2	95	75	56	25	99	39
	53	61	60	63	34	55

−2 −1 0 +1 +2

Example at −1.5,−2

(86+33)/2	(33+76)/2
(68+54)/2	(54+62)/2

→

60	55
61	58

−

70	58
64	53

= 5.25

Example at 2,−0.5

(77+82)/2	(59+82)/2
(82+54)/2	(82+96)/2

→

80	71
68	89

−

70	58
64	53

= 15.75

Example at −1.5,−2

(68+71+75+56)/4	(71+75+56+25)/4
(75+56+61+60)/4	(56+25+60+63)/4

→

68	57
63	51

−

70	58
64	53

= 1.50

Current Block

70	58
64	53

Half-Pel Refinement Example cont'd

Search Window

	−2	−1	0	+1	+2	
−2	86	33	76	94	55	11
−1	68	54	62	34	77	59
0	10	16	28	73	82	82
+1	65	68	71	75	54	96
+2	95	75	56	25	99	39
	53	61	60	63	34	55

Mean Absolute Differences

	−2	−1.5	−1	−0.5	0	+0.5	+1	+1.5	+2	+2.5
−2	11.50	5.25	18.50	13.25	15.75	12.75	20.25	13.25	20.25	19.25
−1.5	16.00	13.00	18.50	12.75	6.75	9.75	12.75	9.25	15.25	13.00
−1	24.25	22.75	23.25	16.50	22.00	15.50	23.25	16.75	13.75	16.00
−0.5	22.75	20.50	18.25	16.50	16.00	12.50	15.75	13.25	15.75	14.00
0	29.50	28.25	26.50	20.25	21.50	15.00	9.75	13.75	22.25	20.50
+0.5	20.75	17.25	13.75	9.00	9.75	9.50	13.00	10.50	15.25	15.50
+1	17.00	12.25	7.25	7.25	13.50	10.50	23.50	10.00	25.75	13.75
+1.5	12.25	8.25	4.25	1.50	7.25	8.25	18.25	8.50	6.50	9.50
+2	15.25	9.50	4.25	8.25	15.25	9.75	26.50	10.50	20.00	14.25
+2.5	18.25	16.00	14.50	17.50	20.25	21.00	25.75	21.50	21.50	28.50

Current Block

70	58
64	53

$\rightarrow$

Motion Vector: (−0.5, +1.5) + Error:

2	1
1	2

Mode Decision

- Macroblock MSE < tsh_1:
  - transmit motion only
- tsh_1 < MB MSE < tsh_2:
  - transmit motion + DCT on DFD
  - Displaced Frame Difference: motion compensated error image (predicted-original)
  - Adapted DCT quantization (around 0)
- Macroblock MSE > tsh_2:
  - INTRA MB

Not Standardized !

B-frame Encoding Process

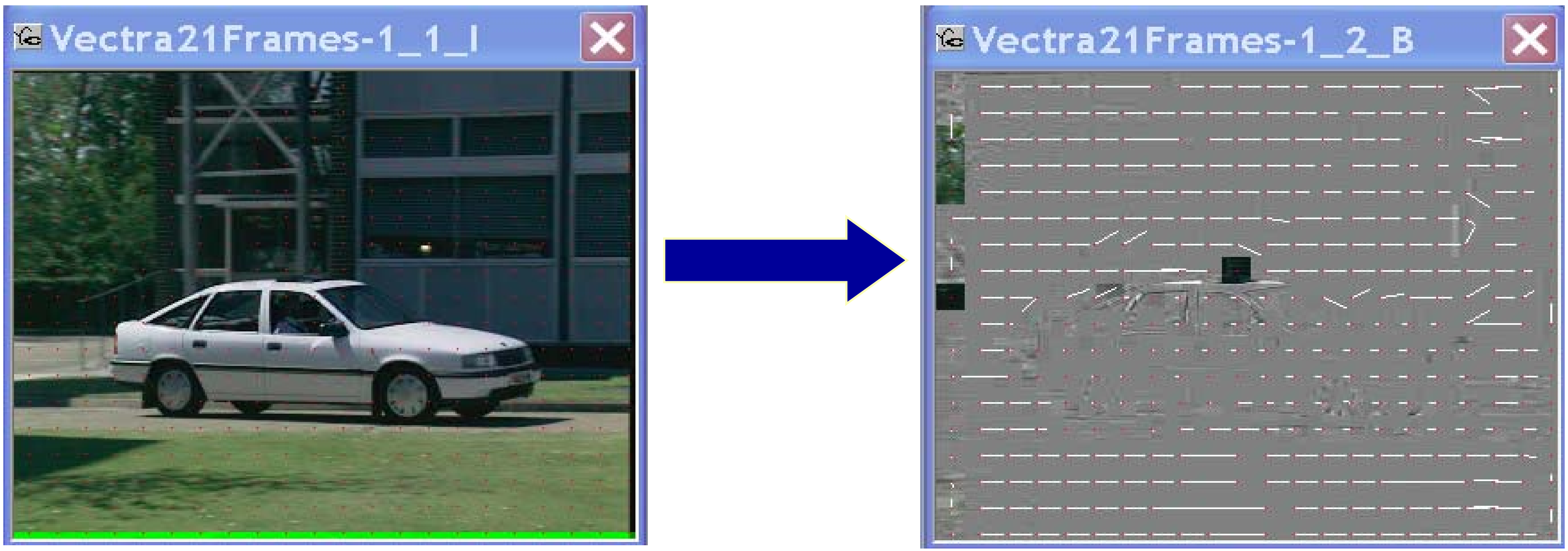
- For each GOP, we first encode all I and P frames (typically 1/3 of all frames)
- The remaining (B frames) can be interpolated from the reconstructed I and P frames
- The resulting interpolation error is DCT encoded

B-frame Pro's and Con's

- ☺ Allow effective handling of covered / uncovered problems
- ☺ Better MC provides a better SNR
- ☺ Since not used to predict other frames, they can be encoded with fewer bits
- ☹ Frame buffers needed (both decoder and encoder)
- ☹ If too many: more bits needed to encode reference frames, and coding delay increase

B-Interframe Types

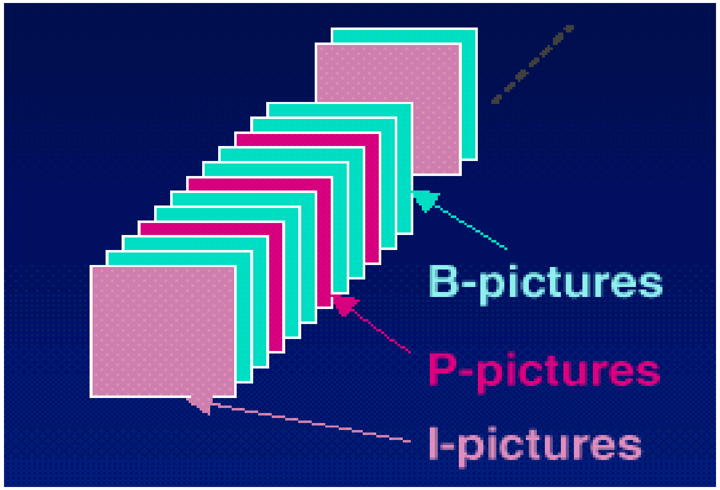
- *Inter-B*: Backward MC active
- The rest types are same as for P-Interframe (see slide 17)



B-Interframe Quantization

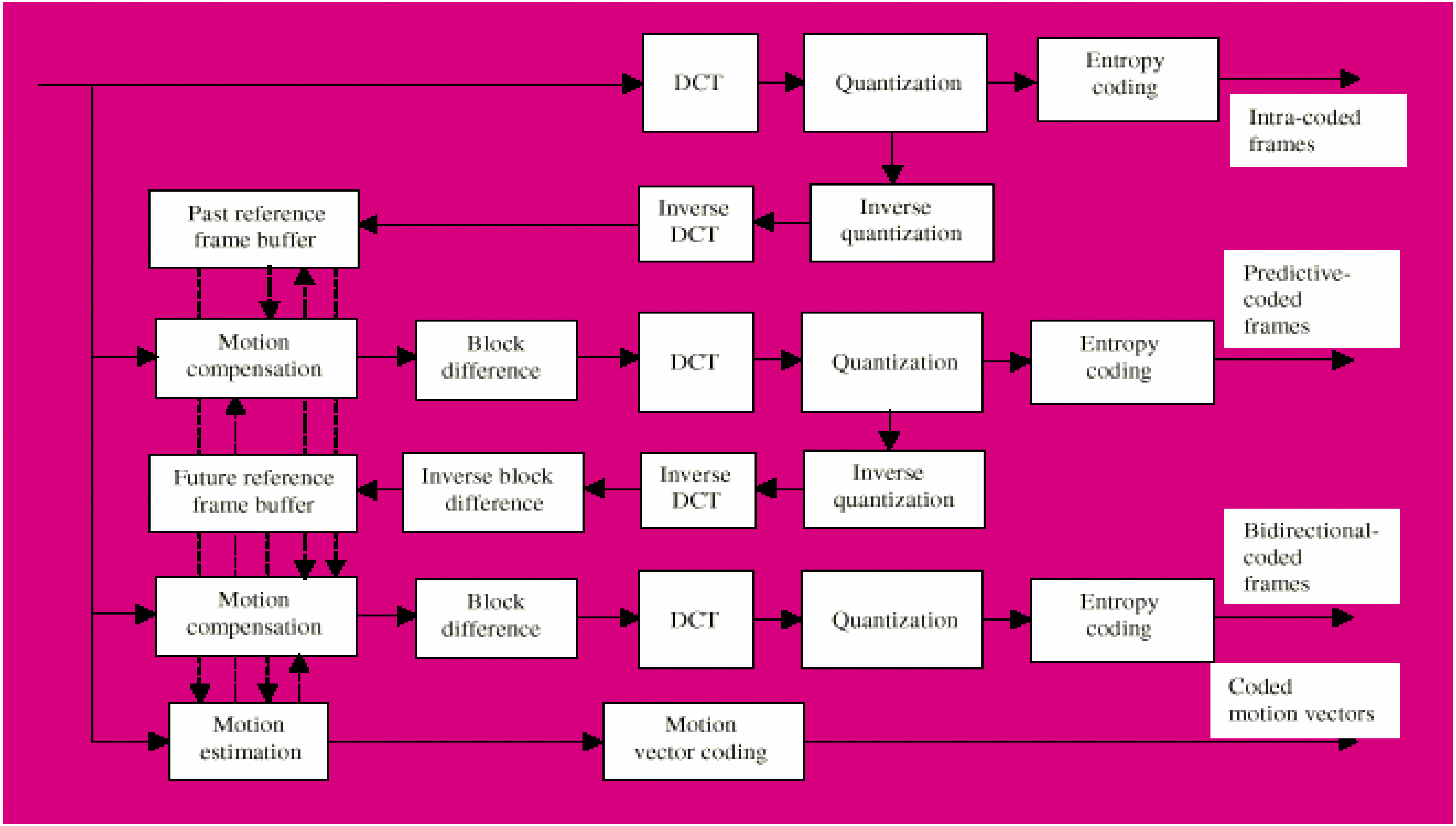
- All DCT coeff. are 11-bit [-2048,2047]
- Q-Table relatively coarser than for I-frame
- All coeff. (including DC) are Zig-Zag scanned for [*run*,*Level*] pairs, and then VLC coded
- Displacement vectors are DPCM coded
- Huffman tables different than P-frames tables

I-P-B Summary



- Intra I- frames:
  - random access
  - error robustness
- Predicted P- frames
  - backward predicted from previous anchor picture
- Bi-directionally predicted B- frames
  - forward/ backward predicted from previous anchor picture (I or P)

MPEG-1 Block Diagram



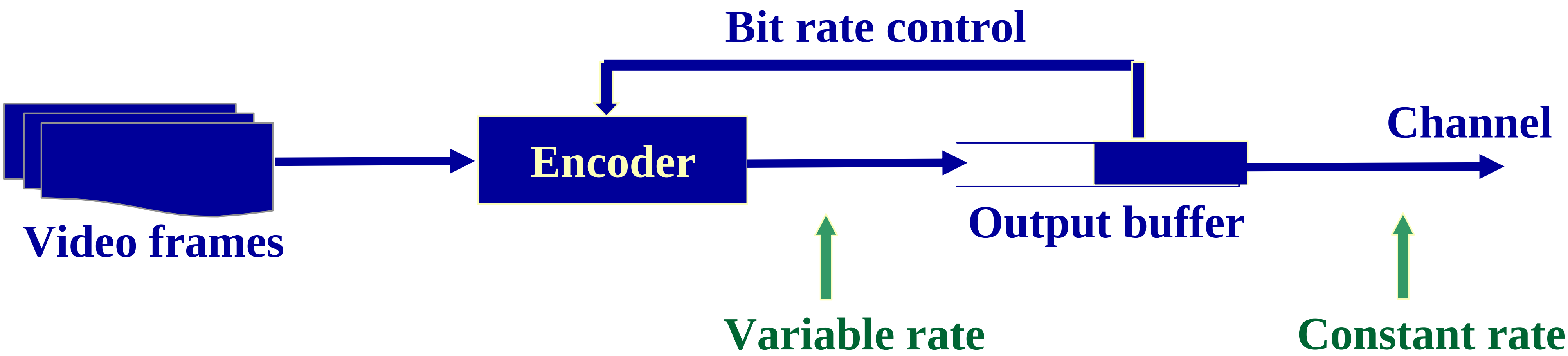
H.261 Vs. MPEG-1

H.261	MPEG-1
Sequential Access	Random Access
1 basic frame rate	Flexible frame rates
CIF/QCIF only	Flexible image size
I and P frames	I,P and B frames
MC over 1 frame	MC over 1 or more frames
1 pixel MV accuracy	1/2 pel MV accuracy
Optional filter in loop	No filter
Variable Th.+Uniform Q	Quantization matrix
BOB structure (no GOF)	GOF and Slice structures

Bit Rate Control (BRC)

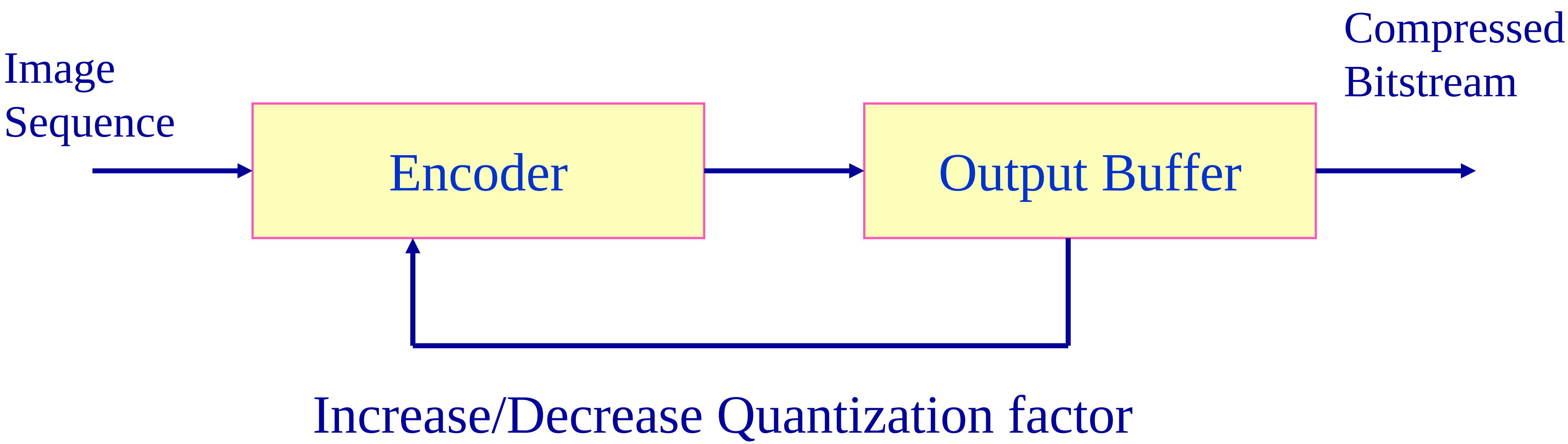
Objectives

- Smart bit allocation:
  - picture level & MB level
- Prevent buffer overflow



Rate Control

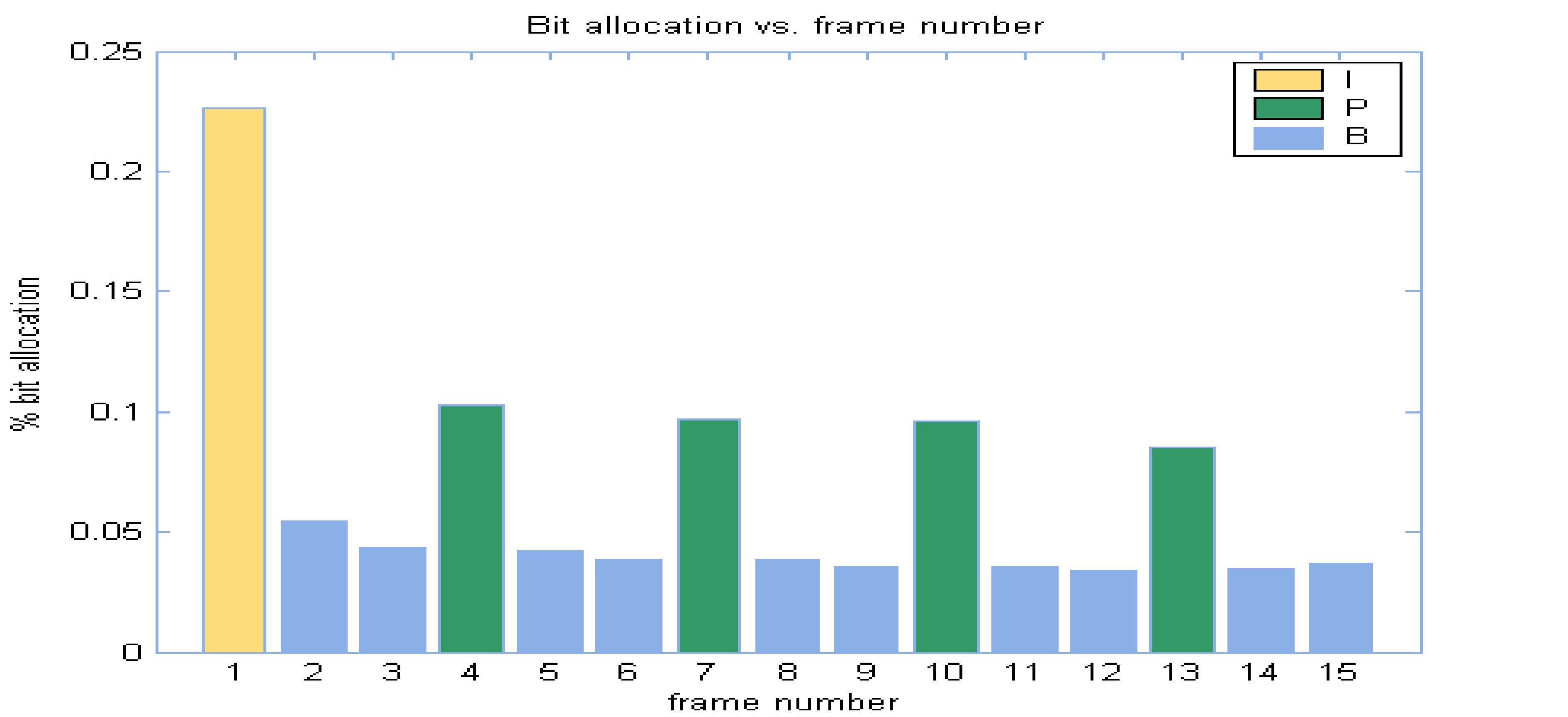
Not standardized !



DCT quantizers *mode decision* scheme can also be changed !

BRC – Test Model 5

- Frame level target bit allocation
  Frame type, remaining bits in the GOP, previous picture complexity



BRC – Test Model 5 (cont'd)

- MB level buffer monitoring - Choose quantizer step size to meet the target frame rate

- MB level adaptive quantization

Current frame

MB activity map (logarithmic scale)