10. MPEG-1

Prof. Chia-Wen Lin
Department of Electrical Engineering
National Tsing Hua University
03-5731152
cwlin@ee.nthu.edu.tw

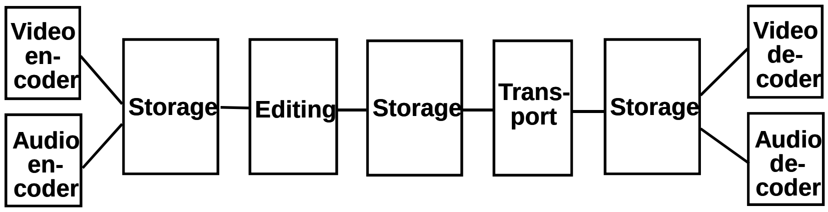
History of MPEG-1 Standard

- Applications
  - Video on digital storage media
  - Computer and telecommunication networks
- History
  - 1988: started
  - Sep. 1989: Proposal registration
  - Oct. 1989: Subjective tests
  - Mar. 1990: Definition of video algorithm (Simulation Model 1)
  - Sep. 1990: Committee Draft for video

Competition (1988 – Sep. 1989)
Convergence (Oct. 1989 – Sep. 1990)

MPEG-1 Video Coding System

For basic digital storage media application



MPEG-1 Coding Standard

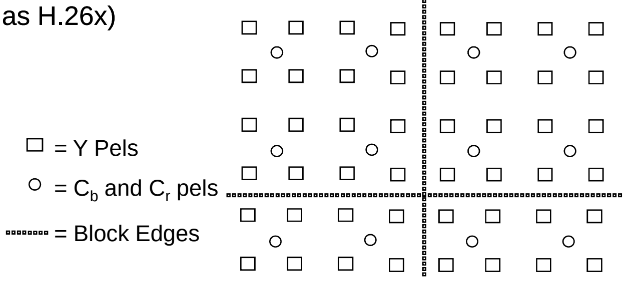
- **For multimedia applications at about 1.5 Mb/s**
- **Used the results of JPEG, H.261, RM8**
- **CD11172: Systems, Video, Audio, Conformance Testing, Simulation Software**
- **ISO-IEC JTC1/SC2/WG11**
  - **- Semiconductor** **- Consumer electronics**
  - **- Computer** **- Telecommunications**
  - **- Broadcasting**
- **Sub working groups**
  - **- Video** **- Audio** **- Systems**
  - **- VLSI** **- Test** **- Digital storage media**
- **Method of working**
  - **- Requirement / Competition / Convergence**
  - **- Simulation Models**
  - **- Draft Standards**

Important Features for MPEG-1 Applications

- Normal playback
- Random access
- Reverse playback
- Fast forward / reverse searches
- Audio-visual synchronization
- Robustness to errors
- Editability
- Format flexibility
- Cost tradeoffs

Parameters

- Picture size up to 4096 × 4096 supported
  - Normally at 360 × 240
- Pel aspect ratio: 14 choices
- Picture rates: 23.976, 24 (movies), 25 (PAL), 29.97, 30, 50, 59.94, 60
- 4:2:0 format
  - (same as H.26x)



Constrained Parameters

- To allow interoperability of MPEG-1 equipment, a subset of parameter space is defined
- All MPEG-1 conformant decoders shall decode constrained parameter bitstreams

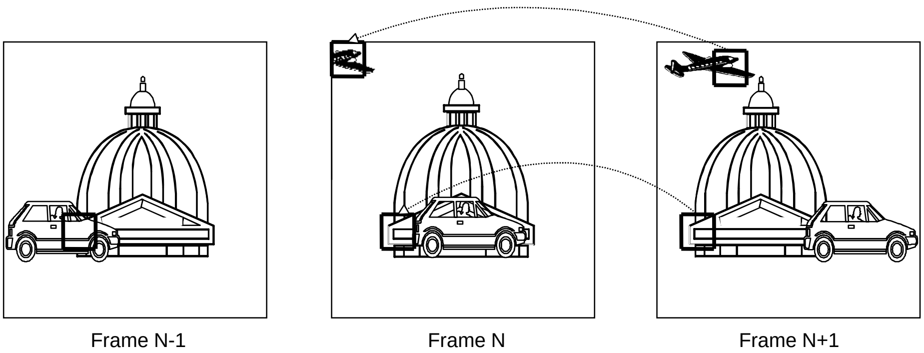
Horizontal picture size	≤ 768 pels
Vertical picture size	≤ 576 lines
Picture area	≤ 396 macroblocks
Pel rate	≤ 396´ 25 macroblocks per second
Picture rate	≤ 30 Hz
Motion vector range	< −64 to +63.5 pels (using half-pel vectors), etc.
Input buffer size (in VBV mode)	≤ to 327 680 bits
Bitrate	≤ 1 856 000 bits/second (constant bitrate)

Typical MPEG-1 Video Source Format

Format	**SIF (525/625)**	
Signal component	**Lines/Frame**	**Pixels/Line**
Luminance (Y)	**240/288**	**352**
Chrominance (Cb)	**120/144**	**176**
Chrominance (Cr)	**120/144**	**176**

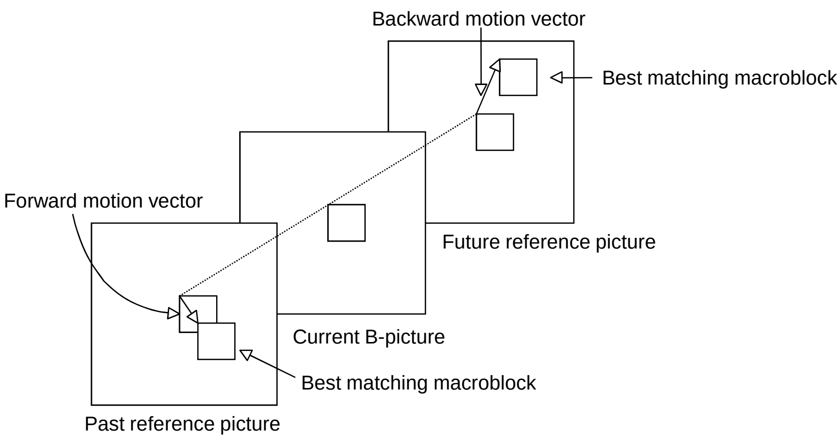
- Uncompressed bit-rate for transmitting SIF at 30 fps is 30.4 Mb/s

Bi-directional Motion Estimation



Bi-directional Motion Estimation

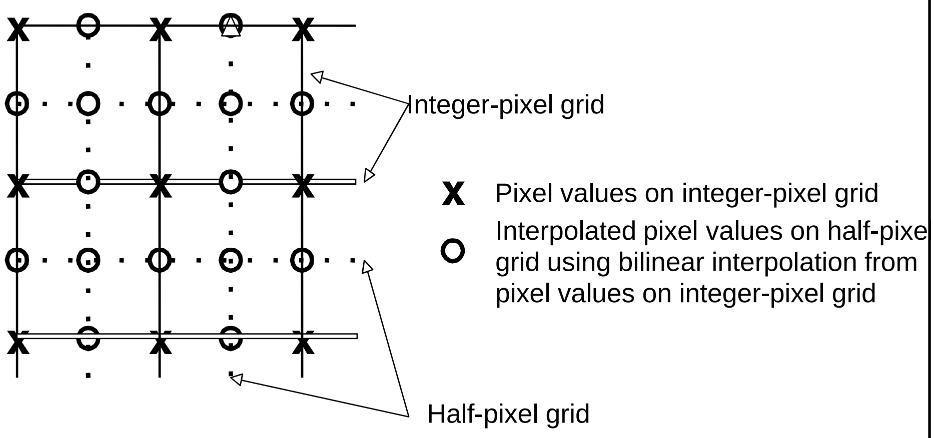
- Forward, backward, or average prediction:
  one or two motion vectors per 16x16 block



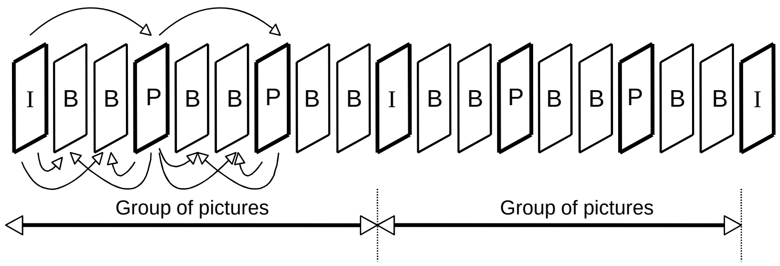
Bi-directional Motion Estimation (Cont.)

- Advantages
  - Higher coding efficiency
    - Increased frame rate with few extra bits
  - No uncovered background problem
  - No error propagation
- Disadvantages
  - More frame storage is needed
  - More delay

Motion Estimation with Half-Pixel Accuracy



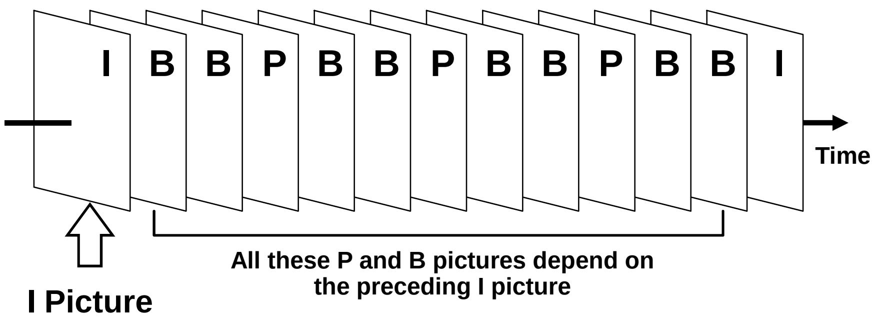
Coding Structure of GOP



I = **Intra-Picture** Coding, allow random access, for reference
P = **Predictive** coding, causal prediction only, can be referenced
B = **Bi-directional** coding, noncausal prediction, never referenced

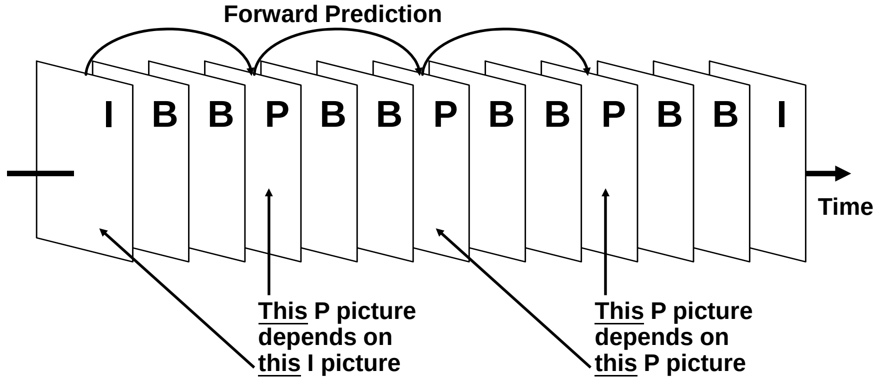
I-Pictures

- DCT coded without reference to any other pictures
- Stored in a frame buffer in encoder and decoder
- Used as basis of prediction for entire GOP



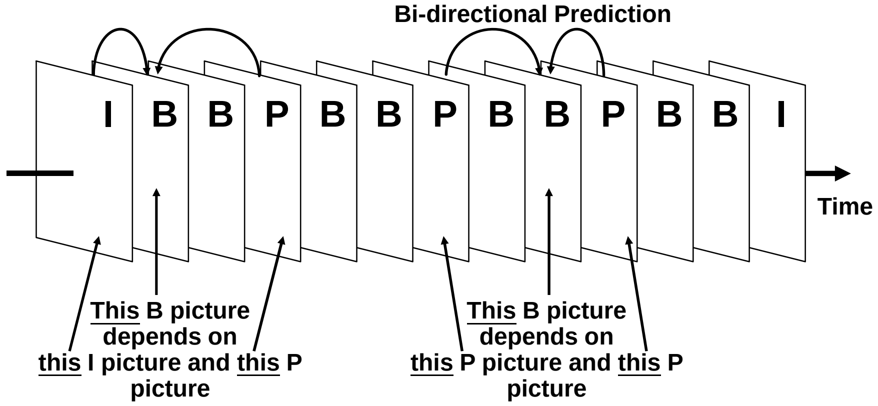
P-Pictures

- DCT coded with reference to the preceding anchor picture
- Stored in a frame buffer in encoder and decoder
- Use forward prediction only



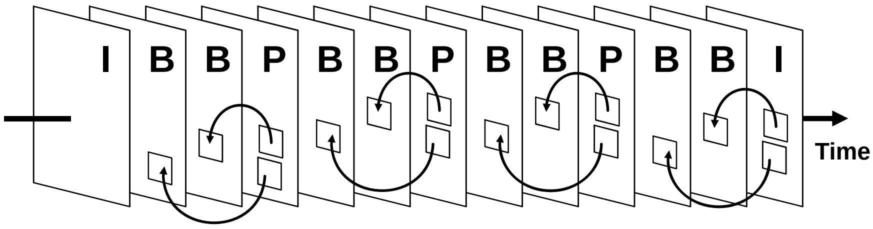
B-Pictures

- DCT coded with reference to either the preceding anchor picture, the following anchor picture, or both
- Use forward, backward or bi-directional prediction



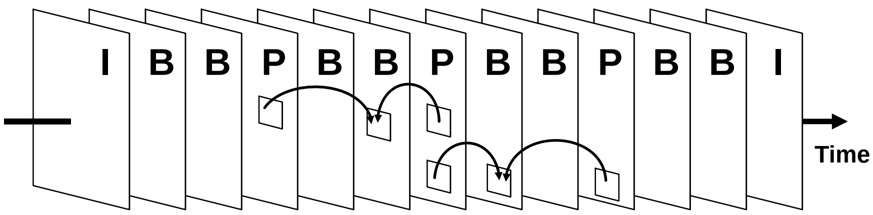
Backward Prediction

- A backward-predicted macroblock depends on decoded pixels from the immediately following anchor picture
- Can only be used to code macroblocks in B pictures



Bidirectional (Interpolated) Prediction

- A bi-directionally-predicted macroblock depends on decoded pixels from the anchor pictures immediately following and immediately preceding
- Can only be used to code macroblocks in B pictures



MPEG Picture Types

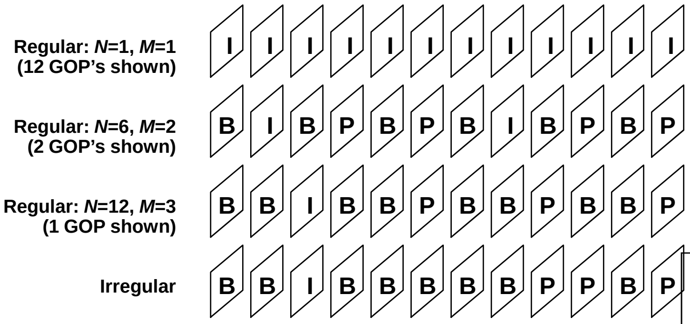
- Generated number of bits: I > P > B. For example, I~300 kbits, P~100-65 kbits (fast/slow motion), B~18-7 kbits (fast/slow motion), per frame.
- **B Pictures:**
  - Pros: Best prediction and compression, object occlusion and entrance into scene, noise averaging.
  - Cons: Encoder delay, high complexity, large encoder buffers required.

GOP Rules

- A GOP must contain at least one I picture
- This I picture may be followed by any number of I and P pictures
- Any number of B pictures may occur between anchor pictures, and B pictures may precede the first I picture
- A GOP, in coding order, must start with an I picture
- A GOP, in display, order must start with an I or B picture and must end with an I or P picture

Regular and Irregular GOPs

- Regular GOP's are defined by *N* and *M**:
  - *N* is the I picture interval
  - *M* is the anchor picture interval. There are *M*-1 B pictures between anchor pictures
- Irregular GOP's are not defined by *N* and *M*, but are still allowed as long as they follow the GOP Rules.

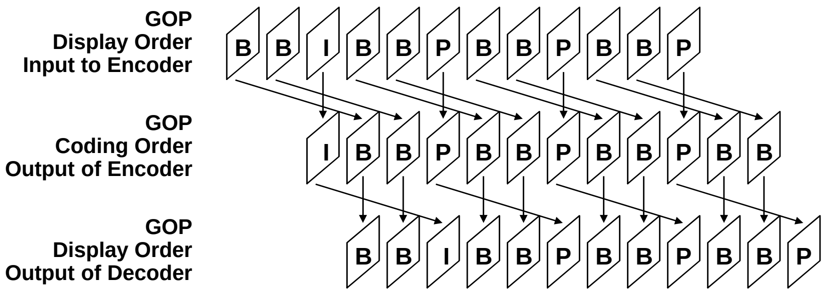


**N and M are not MPEG syntax elements and are not used in any way by the specification.*

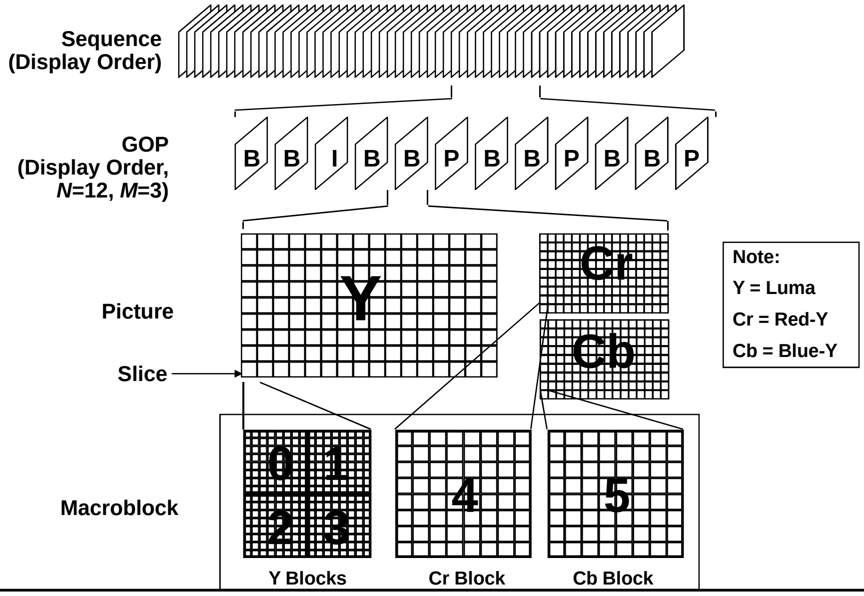
All GOP's in Display Order

GOP Picture Orderings

- **Two Distinct Picture Orderings**
  - **<u>Display Order</u> (input to encoder, output of decoder)**
  - **<u>Coding Order</u> (output of encoder, input to decoder)**
  - **These are different if <u>B frames</u> are present**
  - **B frames must be reordered so that "future" anchor pictures are available for prediction. Note that reordering causes DELAY!**



Six Hierarchical Layers of MPEG

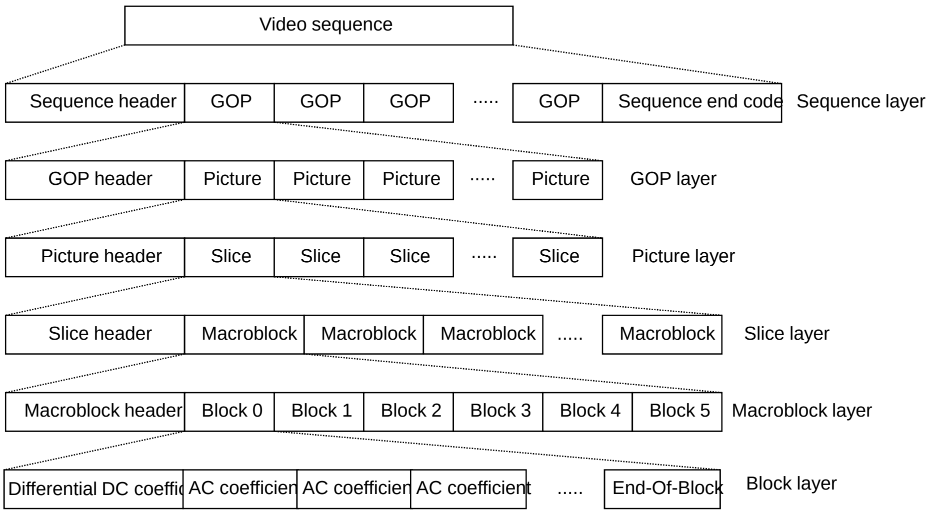


Six Hierarchical Layers of MPEG (Cont.)

- Important syntax elements in each layer:

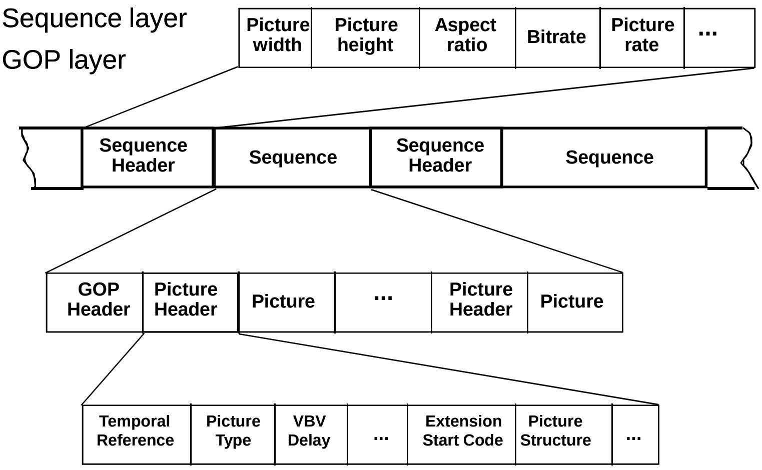
Sequence	Picture Size; Frame Rate Bit Rate; Buffering Requirements Programmable Coding Parameters
GOP	Random Access Unit SMPTE Time-Code
Picture	Timing information (buffer fullness, temporal reference), Coding type (I, P, or B)
Slice	Intra-frame addressing information Coding re-initialization (error resilience)
Macroblock	Basic coding structure, Coding method, Motion Vectors, Quantization
Block	DCT coefficients

MPEG-1 Video Bit-stream



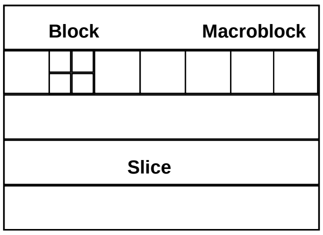
MPEG Bit-Stream Structure

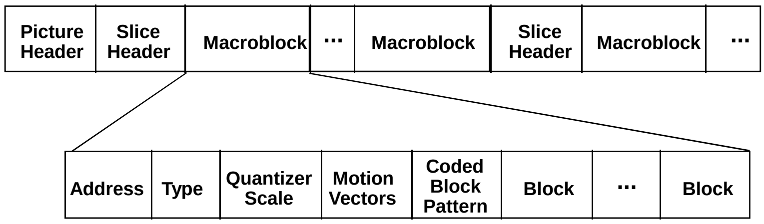
- Sequence layer
- GOP layer



MPEG Bit-Stream Structure (Cont.)

- Picture layer
- Slice layer
- Macroblock layer
- Block Layer





MPEG-1 Video Bitstream Hierarchy

Sequence - picture and aspect ratio, picture rate, *bit rate*,
minimum *decoder buffer size (vbv_size)*,
constrained parameters flag,
(intra quantization table),
(non-intra quantization table), (user data)

Group Of Pictures - time code, closed GOP flag,
broken link flag, (user data)

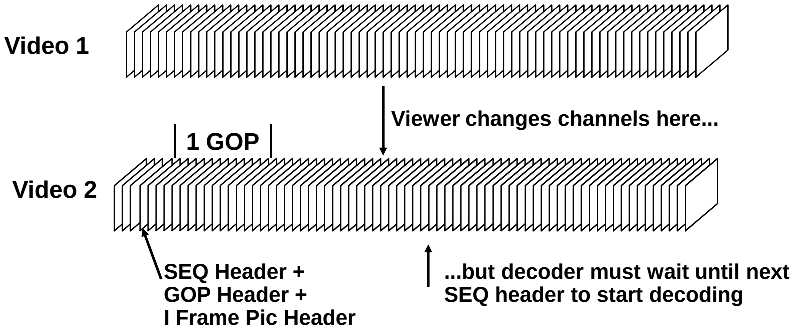
Picture - temporal reference, picture type (I/P/B/D),
decoder buffer initial occupancy (vbv_delay),
(forward motion vector resolution and range),
(backward motion vector resolution and range),
(user data)

Slice - slice vertical position, quantizer scale

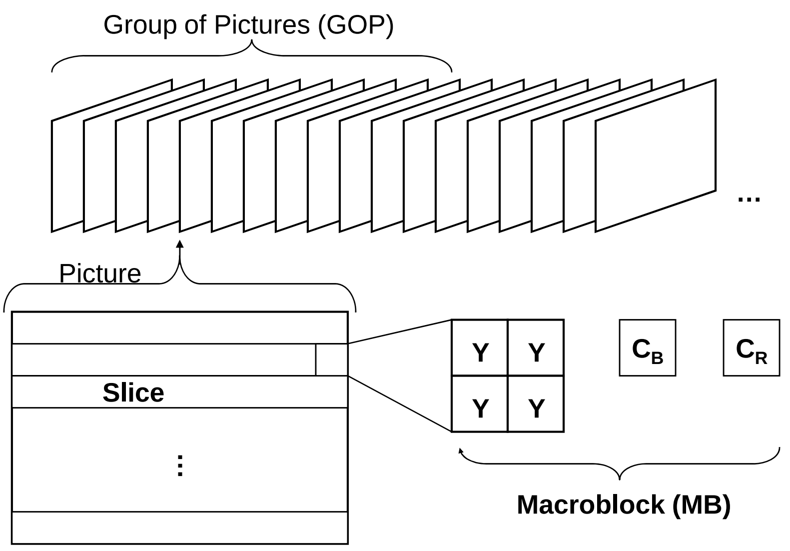
Macroblock - (stuffing), address increment, macroblock type,
quantizer scale, motion vectors,
coded block pattern, DCT coefficients

Sequence

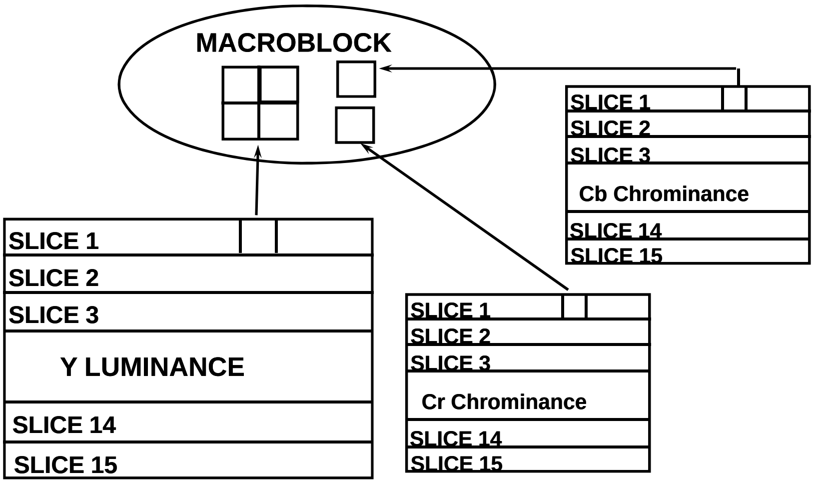
- For CD-ROM applications, sequences can be used to indicate relatively long clips (e.g. shots, scenes or entire movies)
- For broadcast applications, sequence headers are usually sent frequently (e.g., every GOP) so that key bitstream information is obtained at channel changes



Data Structure

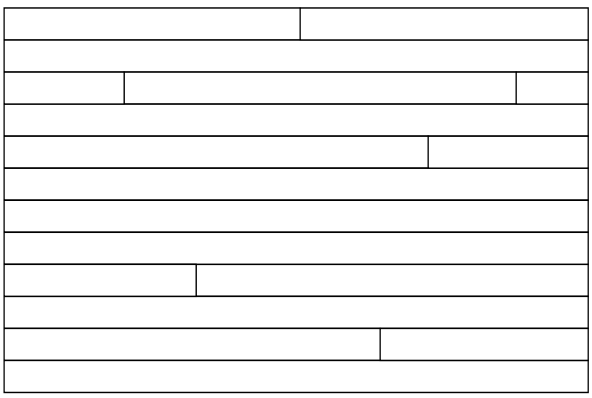


MPEG-1 Picture

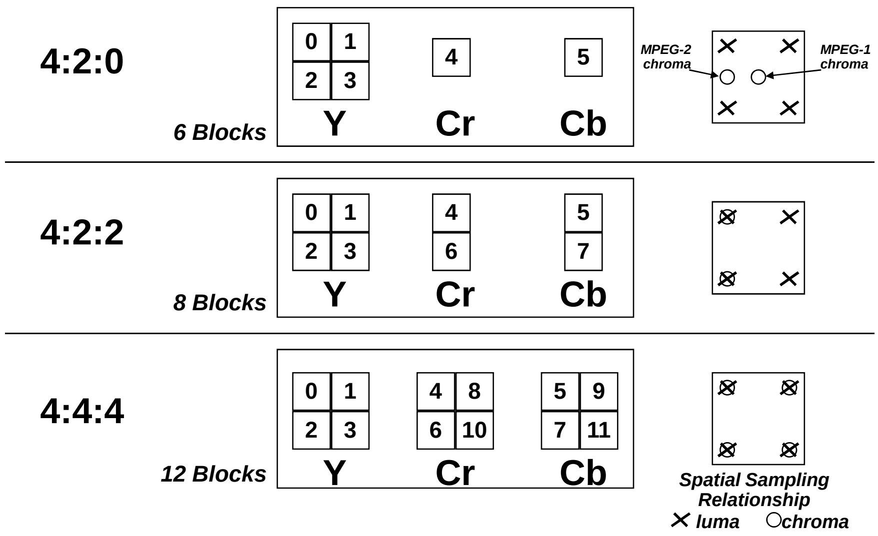


Slice Structure

- A slice is a collection of macroblocks in raster scan order.
- A slice in MPEG-1 video frame can be single MB or entire picture.



Macroblock Structure



VLC for MV Difference

f-code = 1,2,3,4,5,6,7

$f = 2^{f\text{-}code-1}$=1,2,4,8,16,32,64

dMD=MD – PMD

motion_code=[dMD+Sign(dMD)×(*f*-1)]/*f*

r = | motion_code × *f* | – | dMD |

The residual *r* is sent with log(*f*) bits

dMD = motion_code×*f*

– Sign(motion_code)×*r*

|motion_code|	VLC
0	1
1	01s
2	001s
3	0001 s
4	0000 11s
5	0000 101s
6	0000 100s
7	0000 011s
8	0000 0101 1s
9	0000 0101 0s
10	0000 0100 1s
11	0000 0100 01s
12	0000 0100 00s
13	0000 0011 11s
14	0000 0011 10s
15	0000 0011 01s
16	0000 0011 00s

– Range [–512, +511.5] for half-pel motion vectors
– Range [–1024, +1023] for full-pel motion vectors

MPEG MV Coding

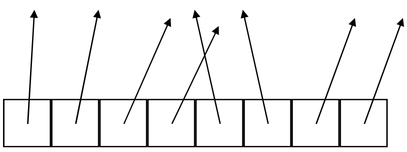
f-code	modulus	MV Range	
		Full-pel=0	Full-pel=1
1	32	-8~7.5	-16~15
2	64	-16~15.5	-32~31
3	128	-32~31.5	-64~63
4	256	-64~63.5	-128~127
5	512	-128~127.5	-256~255
6	1024	-256~255.5	-512~511
7	2048	-512~511.5	-1024~1023

dMD = MV-PMV

motion code = (dMD%modulus)/$2^{f\text{-}code\text{-}1}$

Residual = (dMD%modulus)%$2^{f\text{-}code\text{-}1}$

MV Coding Example

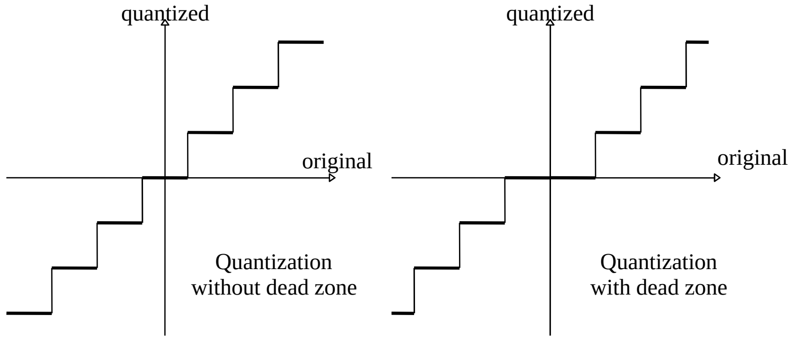


MV Field										**Motion Vectors (MV's) shown for 8 successive macroblocks.**
MV	**x** **y**	**3** **-10**	**10** **-10**	**30** **-9**	**30** **-9**	**-14** **-11**	**-16** **-11**	**27** **-10**	**24** **-10**	**Assume all [x, y] for picture in RANGE [-32, 31] => f_code = 2, MODULUS= 64.**
ΔMV	**x** **y**	**3** **-10**	**7** **0**	**20** **1**	**0** **0**	**-44** **-2**	**-2** **-0**	**43** **1**	**-3** **0**	**ΔMV = Differential MV. [0,0] used as predictor for first MV.**
ΔMV'	**x** **y**	**3** **-10**	**7** **0**	**20** **1**	**0** **0**	**20** **-2**	**-2** **-0**	**-21** **1**	**-3** **0**	**Add or subtract MODULUS if out of RANGE. Keeps all values in RANGE.**
ΔMV" **VLC**	**x** **y**	**0101,000101,00000100100,10,00000100100,0110,00000100110,0110** **000010110,10,11,10,0111,10,11,10**								**Convert to VLC's using table Table 2-B.4 in the MPEG-1 Video spec. VLC's used in this example are for illustration only.**

- **Note that the vertical components of the MV's are much more correlated than the horizontal components.**
- **Therefore, the MV differentials for the vertical components code with fewer bits.**

Quantization

- Same as H.263



Intra

(nearest integer)

Nonintra

(rounded down)

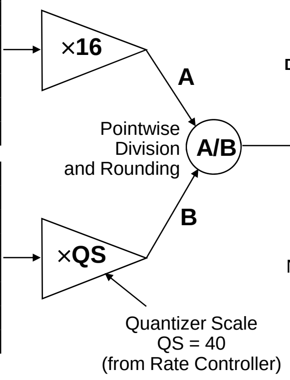
Quantization Example

DCT Frequency Coefficients T[u][v]

DC	276	59	89	39	7	-13	-12	-7
	137	-94	-35	4	17	16	7	2
	51	25	-42	-20	-14	1	5	7
	-12	40	-8	-16	-4	-4	-5	-5
	-8	3	17	-13	-4	0	2	-1
	2	14	14	5	-7	0	-1	0
	-1	-3	-2	12	0	-4	-2	1
	-6	2	-6	6	8	-5	-1	0

DC	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

Default Intra Quantization Matrix QM[u][v]



Quantized DCT Coefficients T'[u][v]

DC	35	1	2	1	0	0	0	0
	3	2	-1	0	0	0	0	0
	1	0	-1	0	0	0	0	0
	0	1	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0

Note: Quantization of DC term is fixed and does not depend on QM or QS.

Default Quantization Matrices

DC								
	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

Intra Matrix: $QM_I[u][v]$

Note: AC coefficients (all coefficients except DC) are first multiplied by 16, then divided by $QS*QM_I[u][v]$.

DC term is treated specially.

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

Non-Intra Matrix: $QM_N[u][v]$

Note: All coefficients are first multiplied by 16, then divided by $QS*QM_N[u][v]$.

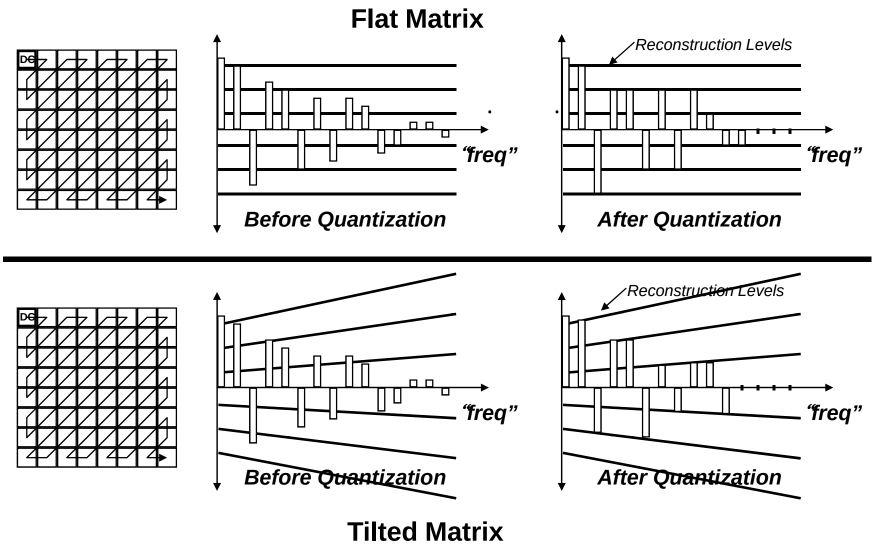
Downloadable Quantization Matrices

- **For improved quality in certain coding situations, quantization matrices for Intra and Non-Intra macroblocks can be downloaded.**
- **The decoder uses these instead of the defaults (which are not sent in the bitstream)**
- **The example at right shows an improved Non-Intra Quant Matrix used by the MPEG-2 Test Model 5 (TM5)**

16	17	18	19	20	21	22	23
17	18	19	20	21	22	23	24
18	19	20	21	22	23	24	25
19	20	21	22	23	24	26	27
20	21	22	23	25	26	27	28
21	22	23	24	26	27	28	30
22	23	24	26	27	28	30	31
23	24	25	27	28	30	31	33

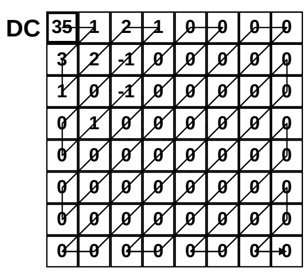
Example of Downloadable Matrix (TM5 Non-Intra Matrix)

Quant Matrix Efeect



Run-Length Coding

- Zeros of the 8x8 block are run length coded
- To optimize the runs, the block is zig-zag scanned



Zig-zag scan (MPEG-1 pattern) through quantized DCT coefficients

35
0, 1
0, 3
0, 1
0, 2
0, 2
0, 1
0, -1
3, 1
0, -1
End of Block

DC Coefficients are differenced from block to block and VLC'd

Common Run/Level Pairs are VLCed

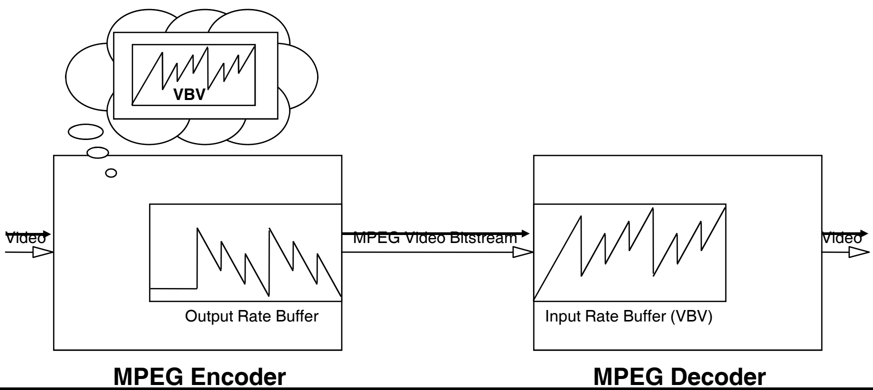
Corresponding Run/Level Pairs

2-D VLC

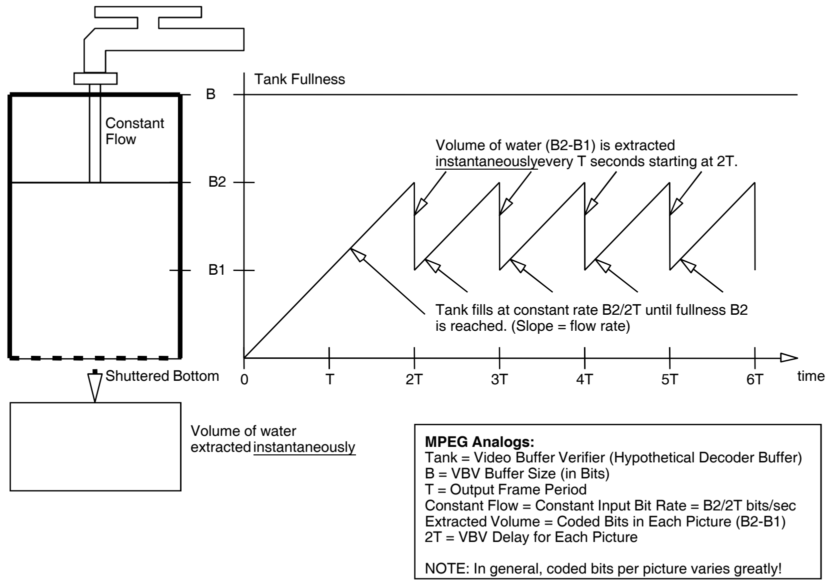
- Similar to H.261
  - Most codes same, some codes shorter
  - Levels up to 255

Video Buffer Verifier (VBV)

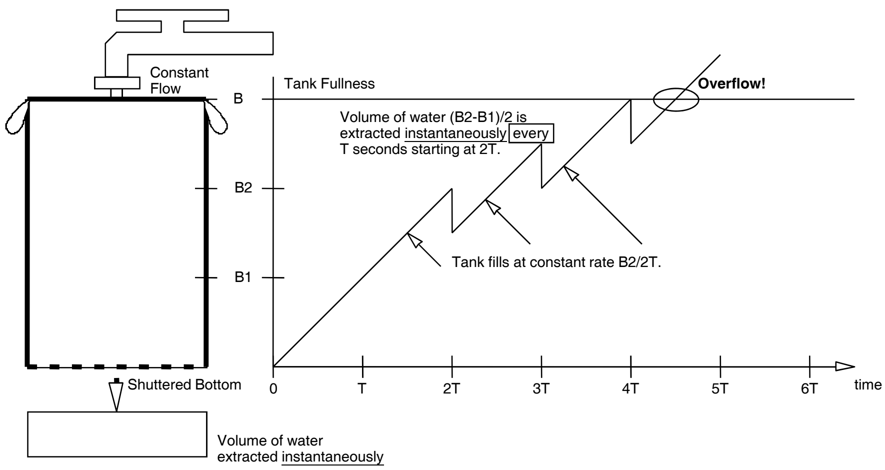
- The VBV is a hypothetical input rate buffer for the video decoder, which is connected to the output of an encoder.
- The encoder keeps track of the VBV fullness, and must ensure that it does not overflow or underflow.
- Assuming constant end-to-end delay, the encoder buffer is the mirror image of the VBV.



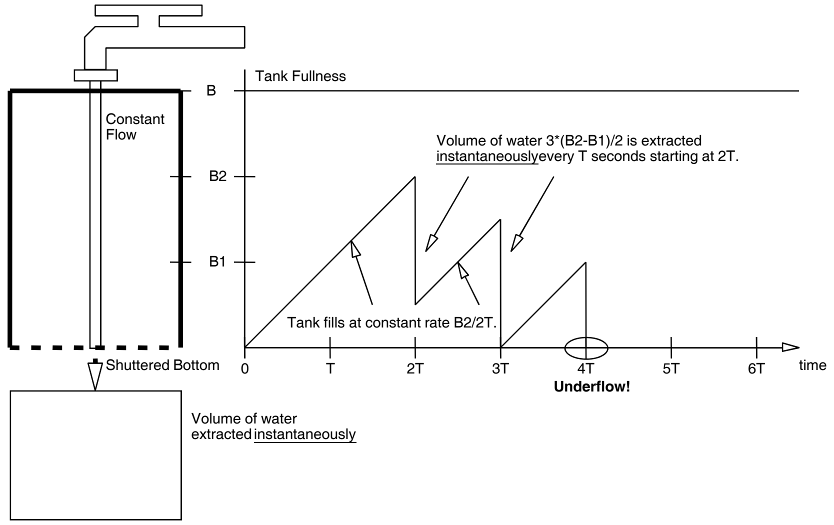
MPEG's VBV Water Tank Analogy (Normal Operation)



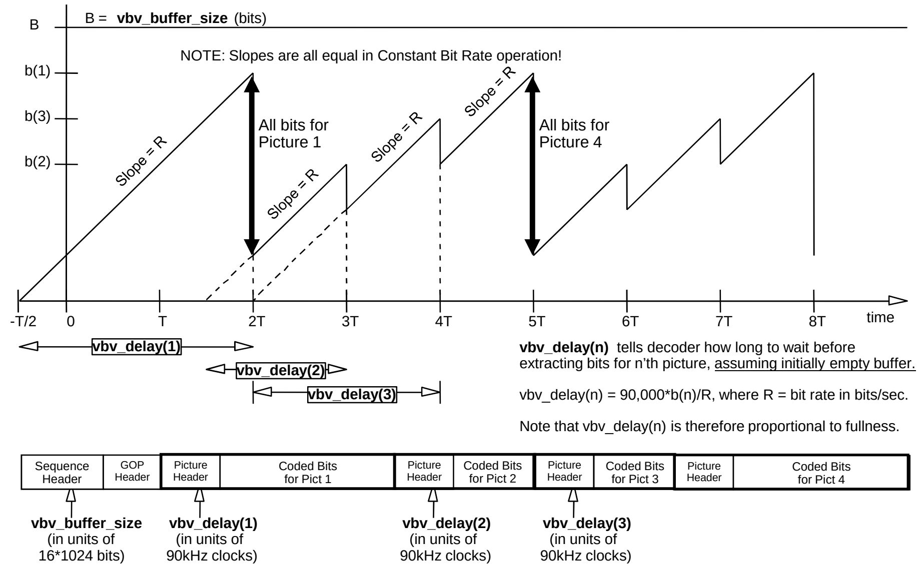
MPEG's VBV Water Tank Analogy (Overflow)



MPEG's VBV Water Tank Analogy (Underflow)



VBV Buffer Size & VBV Delay



MPEG-1 Video Coding Standard

- Similar to H.261, MPEG-1 specifies
  - Bit stream syntax and semantics
  - Decoding to raster representation
  - VBV (Video Buffer Verifier)

  doesn't specify
  - Pre-processing to raster representation
  - Encoding
  - Post-processing from raster representation
- Left flexible (parameters in the bit stream)
  - Coded bit-rate (constant or variable)
  - Lines per picture (< 4096)
  - Pels per line (< 4096)
  - Picture rate (24, 25, or 30 per second)
  - Pel aspect ratio (14 choices)

MPEG-1 Compared to H.261

- **Bi-directional motion compensation with half-pixel accuracy**
- Visually weighted quantization
- In Intra-mode, the DC-coefficient is encoded similar to that in JPEG
- I, P, and B picture types organized as a flexible Group of Pictures (GOP)
- Slice structure instead of Group of Blocks (GOB)
- D-Pictures
- Support maximum motion vector range of -512 to +511.5 pixels, for half-pixel motion vectors: -1024 to + 1023 for full-pixel
- Flexible format: picture sizes up to 4k x 4k, 360 x 240 (SIF) normally used. Variety of picture rates: 23.98, 24, 25, 29.97, 30, 50, 59.94, 60
- VBV

Simulation Model 3 (SM3)

- A specific reference implementation of MPEG-1 encoder including details which were not specified in the standard
- Motion estimation: one forward and/or one backward vector per MB with half-pixel resolution; 2-step search: (i) full search in the range of +/- 7 pixels (2) search 8 neighboring half-pel positions
- Methods for MC / No MC and Intra / Inter decision
- Quantizer, rate control

Summary

- MPEG-1 is mainly for storage media and broadcasting applications
- Due to the use of B-pictures, it may result in long end-to-end delay
- MPEG encoder is much more expensive than the decoder due to the motion estimation which has large search range and may have half-pel accuracy
- MPEG-1 syntax can support a variety of rates and formats for storage media applications
- Pre-processing, encoding, and post-processing are open to improvement
- Extensions to include added features are possible