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

• (No) Random Access to any frame in a limited amount of time

• Fast Forward/Reverse Search to display only selected frames (also Freeze mode)

• Coding/Decoding Delay 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

Hierarchical structure, with similarity to H.261

- Sequence Layer
- Group of Pictures (GOP)
- Picture Layer: I, P, B, D Types
- Slices
- MB Layer
- Blocks Layer

- Not exist in H.261
- Different from H.261
- Similar 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/Bidirectional 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

<table>
<thead>
<tr>
<th>I-Picture</th>
<th>P-Picture</th>
<th>B-Picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra</td>
<td>Intra</td>
<td>Intra</td>
</tr>
<tr>
<td>Intra-A</td>
<td>Intra-A</td>
<td>Intra-A</td>
</tr>
<tr>
<td>Inter-D</td>
<td>Inter-F</td>
<td></td>
</tr>
<tr>
<td>Inter-DA</td>
<td>Inter-FD</td>
<td></td>
</tr>
<tr>
<td>Inter-F</td>
<td>Inter-FDA</td>
<td></td>
</tr>
<tr>
<td>Inter-FD</td>
<td>Inter-B</td>
<td></td>
</tr>
<tr>
<td>Inter-FDA</td>
<td>Inter-BD</td>
<td></td>
</tr>
<tr>
<td>Skipped</td>
<td>Inter-BDA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inter-I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inter-ID</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inter-IDA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skipped</td>
<td></td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>8</th>
<th>16</th>
<th>19</th>
<th>22</th>
<th>26</th>
<th>27</th>
<th>29</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>16</td>
<td>22</td>
<td>24</td>
<td>27</td>
<td>29</td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td>19</td>
<td>22</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>34</td>
<td>34</td>
<td>38</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>34</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>26</td>
<td>27</td>
<td>29</td>
<td>34</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>29</td>
<td>32</td>
<td>35</td>
<td>40</td>
<td>48</td>
<td>58</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>29</td>
<td>34</td>
<td>38</td>
<td>46</td>
<td>56</td>
<td>69</td>
</tr>
<tr>
<td>27</td>
<td>29</td>
<td>35</td>
<td>38</td>
<td>46</td>
<td>56</td>
<td>69</td>
<td>83</td>
</tr>
</tbody>
</table>
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

• **P-Pictures**: Forward prediction, with reference to previous I or P pictures:

\[ b = \hat{c} \]

MB corresponding to b in the reconstructed previous frame

---

Frame \( k \)

Frame \( k-1 \)
**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

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

Holes or overlap regions are created in the predicted image
B-Picture Prediction Types

- $a_1=0; a_2=1$: Backward prediction
- $a_1=1; a_2=0$: Forward prediction
- $a_1=a_2=0.5$: Bidirectional prediction

For this type, 2 displacement vectors $(d_1, d_2)$ and prediction error $(b-b\sim)$ need to be coded for each MB $b$. 
Half-Pel refinement

• motion estimation from previous reconstructed and interpolated frame

Horizontal Interpolated pixel
\[ h = (A+B)/2 \]

Vertical interpolated pixel
\[ v = (A+C)/2 \]

Central interpolated pixel
\[ c = (A+B+C+D)/4 \]
Half-Pel Refinement Example

Search Window

Example at $-1.5,-2$

$$
\begin{array}{c}
\frac{86+33}{2} & \frac{33+76}{2} \\
\frac{68+54}{2} & \frac{54+62}{2}
\end{array}
$$

$$
\begin{array}{c}
60 & 55 \\
61 & 58
\end{array}
- 
\begin{array}{c}
70 & 58 \\
64 & 53
\end{array}
= 5.25

Example at $2,-0.5$

$$
\begin{array}{c}
\frac{77+82}{2} & \frac{59+82}{2} \\
\frac{82+54}{2} & \frac{82+96}{2}
\end{array}
$$

$$
\begin{array}{c}
80 & 71 \\
68 & 89
\end{array}
- 
\begin{array}{c}
70 & 58 \\
64 & 53
\end{array}
= 15.75

Example at $-1.5,-2$

$$
\begin{array}{c}
\frac{68+71+75+56}{4} & \frac{71+75+56+25}{4} \\
\frac{75+56+61+60}{4} & \frac{56+25+60+63}{4}
\end{array}
$$

$$
\begin{array}{c}
68 & 57 \\
63 & 51
\end{array}
- 
\begin{array}{c}
70 & 58 \\
64 & 53
\end{array}
= 1.50

Current Block

$$
\begin{array}{c}
70 & 58 \\
64 & 53
\end{array}$$
# Half-Pel Refinement Example cont’d

## Search Window

<table>
<thead>
<tr>
<th>Search Window</th>
<th>Mean Absolute Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>11.50</td>
</tr>
<tr>
<td>-1.5</td>
<td>16.00</td>
</tr>
<tr>
<td>-1</td>
<td>24.25</td>
</tr>
<tr>
<td>-0.5</td>
<td>22.75</td>
</tr>
<tr>
<td>0</td>
<td>29.50</td>
</tr>
<tr>
<td>+0.5</td>
<td>20.75</td>
</tr>
<tr>
<td>+1</td>
<td>17.00</td>
</tr>
<tr>
<td>+1.5</td>
<td>12.25</td>
</tr>
<tr>
<td>+2</td>
<td>15.25</td>
</tr>
<tr>
<td>+2.5</td>
<td>18.25</td>
</tr>
</tbody>
</table>

## Current Block

![Current Block](image)

## Motion Vector

Motion Vector: \((-0.5, +1.5)\) + Error:

![Error](image)
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

- Past reference frame buffer
- Motion compensation
- Future reference frame buffer
- Motion compensation
- Motion estimation
- Inverse DCT
- Block difference
- Inverse block difference
- Block vector coding
- DCT
- Quantization
- Inverse quantization
- Inverse DCT

- Entropy coding
- Entropy coding
- Entropy coding
- Entropy coding

- Intra-coded frames
- Predictive-coded frames
- Bidirectional-coded frames
- Coded motion vectors
# H.261 Vs. MPEG-1

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

Objectives

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

Not standardized!

Image Sequence → Encoder → Output Buffer → Compressed Bitstream

Increase/Decrease Quantization factor

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