Introduction to MPEG Video Coding
Topics today…

✓ MPEG -1
✓ MPEG -2
MPEG -1

- Moving Pictures Experts Group, established in 1988
- Approved by ISO/IEC MPEG Group in November’91
  - Coding moving pictures & associated audio up to 1.5Mbit/s
  - Up to 1.2Mbit/s for video & 256kbps for audio
  - Supports only non-interlaced video
MPEG-1 Motion Compensation

- Introduces third frame (in addition to I-Frame, P-Frame) **B-Frame**
- B-Frame uses backward prediction & forward prediction
  - Each MB in the B-Frame will have 2 motion vectors
MPEG-1 Motion Compensation [2]

B-Frame Coding based on bidirectional motion compensation
Multimedia Computing
(CSIT 410)
MPEG-1 Motion Compensation

**Notation**
- $M \leftarrow$ interval between P-Frame and preceding I or P Frame
- $N \leftarrow$ interval between two consecutive I-Frames
- In Fig $M=3$ & $N=9$
MPEG-1  Major differences from H.261

- Source Formats
  - Supports any formats that meet the constrained parameter set

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal size of picture</td>
<td>≤ 768</td>
</tr>
<tr>
<td>Vertical size of picture</td>
<td>≤ 576</td>
</tr>
<tr>
<td>No. of MBs / picture</td>
<td>≤ 396</td>
</tr>
<tr>
<td>No. of MBs / second</td>
<td>≤ 9,900</td>
</tr>
<tr>
<td>Frame rate</td>
<td>≤ 30 fps</td>
</tr>
<tr>
<td>Bit-rate</td>
<td>≤ 1,856 kbps</td>
</tr>
</tbody>
</table>

Constrained Parameter Set
MPEG-1 Major differences from H.261[2]

- Slices
  - Instead of GOBs, MPEG-1 picture is divided into slices
  - Variable number of macroblocks
  - Each slice is coded separately
  - Unique start code

Slices in MPEG-1 Picture
MPEG-1  Major differences from H.261 [3]

- Quantization
  Different quantization tables for *intra* & *inter* coding
  - For intra-coding
    \[
    QDCT[i,j] = \text{round}\left(\frac{8 \times DCT[i,j]}{\text{step_size}[i,j]}\right) = \text{round}\left(\frac{8 \times DCT[i,j]}{Q_1[i,j] \times \text{scale}}\right)
    \]
  - For inter-coding
    \[
    QDCT[i,j] = \left|\frac{8 \times DCT[i,j]}{\text{step_size}[i,j]}\right| = \left|\frac{8 \times DCT[i,j]}{Q_2[i,j] \times \text{scale}}\right|
    \]
### MPEG-1 Video Stream

#### Sequence layer
- Header contains the info about the picture *horizontal size, vertical size, pixel aspect ratio, frame rate, bitrate, buffer size* etc

#### GOP layer
- One of the pictures must be an I-picture
- Header contains time code to indicate hour-min-sec-frame

#### Picture layer
- I, P, B, D pictures

#### Slice layer

#### MB layer
- 6 blocks of 8x8

#### Block layer
- DC coefficient is sent first followed by AC coefficients. (These are coded using variable-length code)
MPEG-1 Video Stream [2]

- **Major Differences from H.261**
  - Source format besides CIF (352x288) and QCIF (76x144)
  - Slices instead of Groups of Blocks, coded independently
  - Different quantization for intra- and inter-coding. Quantizer varies within a macroblock.
  - Bigger range for motion vectors (>>15)
  - Allows random access
MPEG-1 Video Stream [3]

- Compression performance of MPEG-1
  - I-Frame 7:1
  - P-Frame 20:1
  - B-Frame 50:1
  - Average 27:1
MPEG-1 Video Stream [4]
MPEG-2 Introduction

- Focused at higher bitrates, more than 4Mbps
- Development of standard started in 1990
- Approved by ISO/IEC Moving Picture Experts Group in 1994
MPEG-2 Introduction [2]

- Similar to MPEG-1 has parts for:
  - Systems (Packets, synchronization of audio and video, etc.)
  - Video (coding and decoding)
  - Audio (coding and decoding)
  - Conformance (to standards)
  - Software (implementation of coder & decoder)
MPEG-2 Introduction \[3\]

- Standard adopted for DVDs
- Defines 7 profiles aimed at different applications
  - Simple, Main, SNR Scalability, Spatially Scalable, High, 4:2:2 and Multiview
  - Each profile could define up to 4 levels
MPEG-2 Introduction [3]

Profiles & Levels in MPEG - 2

<table>
<thead>
<tr>
<th>Level</th>
<th>Simple Profile</th>
<th>Main Profile</th>
<th>SNR Scalable Profile</th>
<th>Spatially Scalable Profile</th>
<th>High Profile</th>
<th>4:2:2 Profile</th>
<th>Multiview Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>High 1440</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Main</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Low</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Four Levels in the main profile

<table>
<thead>
<tr>
<th>Level</th>
<th>Max Resolution</th>
<th>Max fps</th>
<th>Max Pixels/sec</th>
<th>Max coded Data Rate (Mbps)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>1,920 x 1,152</td>
<td>60</td>
<td>62.7 x 10^6</td>
<td>80</td>
<td>film production</td>
</tr>
<tr>
<td>High 1440</td>
<td>1,440 x 1,152</td>
<td>60</td>
<td>47.0 x 10^6</td>
<td>60</td>
<td>consumer HDTV</td>
</tr>
<tr>
<td>Main</td>
<td>720 x 576</td>
<td>30</td>
<td>10.4 x 10^6</td>
<td>15</td>
<td>studio TV</td>
</tr>
<tr>
<td>Low</td>
<td>352 x 288</td>
<td>30</td>
<td>3.0 x 10^6</td>
<td>4</td>
<td>consumer tape equiv.</td>
</tr>
</tbody>
</table>
MPEG-2 Support for Interlaced Video

- MPEG-2 is adopted by broadcast TV’s, so it needs to support interlaced mode.
- Each frame consists of 2 fields:
  - Top-field
  - Bottom-field
- All scanlines are interleaved to form the frame picture, then macroblock is formed & coding proceeds.
- If each field is treated as separate picture, we talk about field-picture.
MPEG-2 Prediction Modes

1. Frame prediction for frame pictures
2. Field prediction for field pictures
   - Uses macroblock 16 x 16 from field pictures
3. Field prediction for frame pictures
4. 16 x 8 MC for field-pictures
5. Dual-prime for P-pictures

Field Prediction for Field pictures

Multimedia Computing
(CSIT 410)
MPEG-2 Alternate Scan & Field DCT

- Techniques used to improve the effectiveness of DCT in prediction errors applied to frame-pictures with interlaced videos. In fact, in interlaced video:
  - The consecutive rows are from different fields, so less correlation
  - Vertically higher spatial frequency components may have larger magnitudes, so scan them earlier.
- In MPEG-2 Field_DCT could be used for the same purpose: Before DCT, rows are reordered so that the first 8 are from the top-field and the last 8 from the bottom-field.
MPEG-2 Alternate Scan of DCT coef.

(a) Zigzag (progressive) scan and (b) alternate (interlaced) scan
MPEG-2 Scalabilities

- MPEG-2 is designed for various applications, digital TV, HDTV & the video will be transmitted over networks with very different characteristics.

- Single coded MPEG-2 bitstream should be scalable for various bitrates.
MPEG-2 Scalabilities\textsuperscript{[2]}

- Coded as base layer and enhancement layers (EL), they allow the scalability.

- Following scalabilities are supported

1. SNR Scalability – EL provides better SNR
2. Spatial Scalability- EL provides better resolution
3. Temporal Scalability- EL provides more frames/time
4. Hybrid scalability- EL provides any two of the three
5. Data Partitioning- Quantized DCT coefficients are split into partitions, one for lower frequency and another for high-frequency DCT coefficients.
MPEG-2 Major differences from MPEG-1

- Better resilience to bit errors
- Support for 4:2:2 & 4:4:4 chroma subsampling
- Nonlinear Quantization
- More restricted slice
  - Slices must start at left edge of the picture
- More flexible video formats
Reference:

Chapter 11