Overview: Video Coding Standards

- Video coding standards: applications and common structure
- Relevant standards organizations
- ITU-T Rec. H.261
- ITU-T Rec. H.263
- ISO/IEC MPEG-1
- ISO/IEC MPEG-2
- ISO/IEC MPEG-4
- Recent progress: H.264/AVC
## Major Applications of Video Compression

<table>
<thead>
<tr>
<th>Application</th>
<th>Bit Rate</th>
<th>Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital television broadcasting</td>
<td>2...5 Mbps (10...20 Mbps for HD)</td>
<td>MPEG-2 H.264/AVC</td>
</tr>
<tr>
<td>DVD video</td>
<td>4...8 Mbps (10...20 Mbps for HD)</td>
<td>MPEG-2 H.264/AVC, VC-1</td>
</tr>
<tr>
<td>HD-DVD, Blu-ray Disk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet video streaming</td>
<td>20...600 kbps</td>
<td>Proprietary, similar to H.263, MPEG-4, or H.264/AVC, VC-1</td>
</tr>
<tr>
<td>Videoconferencing, videotelephony</td>
<td>20...320 kbps</td>
<td>H.261, H.263, H.264/AVC</td>
</tr>
<tr>
<td>Video over 3G wireless</td>
<td>20...200 kbps</td>
<td>H.263, MPEG-4, H.264/AVC, VC-1</td>
</tr>
</tbody>
</table>
Motion-compensated Hybrid Coding
H.261, MPEG-1, MPEG-2, H.263, MPEG-4, H.264/AVC, VC-1

Diagram:
- Coder Control
- Transform/Quantizer
- Decoder
- Deq./Inv. Transform
- Motion-Compensated Predictor
- Motion Estimator
- Entropy Coding
- Control Data
- Quant. Transf. coeffs
- Intra/Inter Coder
- Motion Data
Video Standards: Hierarchical Syntax I
# Video Standards: Hierarchical Syntax II

## TABLE 13.6 SYNTAX HIERARCHY AS USED IN DIFFERENT VIDEO CODING STANDARDS*

<table>
<thead>
<tr>
<th>Syntax layer</th>
<th>Functionality</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence (SC)</td>
<td>Definition of entire video sequence</td>
<td>H.261/3, MPEG-1/2</td>
</tr>
<tr>
<td>VOL (SC)</td>
<td>Definition of entire video object</td>
<td>MPEG-4</td>
</tr>
<tr>
<td>GOP (SC)</td>
<td>Enables random access in video stream</td>
<td>MPEG-1/2</td>
</tr>
<tr>
<td>GVOP (SC)</td>
<td>Enables random access in video stream</td>
<td>MPEG-4</td>
</tr>
<tr>
<td>Picture (SC)</td>
<td>Primary coding unit</td>
<td>H.261/3, MPEG-1/2</td>
</tr>
<tr>
<td>VOP (SC)</td>
<td>Primary coding unit</td>
<td>MPEG-4</td>
</tr>
<tr>
<td>GOB (SC)</td>
<td>Resynchronization, refresh, and error recovery in a picture</td>
<td>H.261/3</td>
</tr>
<tr>
<td>Slice (SC)</td>
<td>Resynchronization, refresh, and error recovery in a picture</td>
<td>MPEG-1/2</td>
</tr>
<tr>
<td>Video Packet (SC)</td>
<td>Resynchronization and error recovery in a picture</td>
<td>MPEG-4</td>
</tr>
<tr>
<td>MB</td>
<td>Motion compensation and shape coding unit</td>
<td>H.261/3, MPEG-1/2/4</td>
</tr>
<tr>
<td>Block</td>
<td>Transform and compensation unit</td>
<td>H.261/3, MPEG-1/2/4</td>
</tr>
</tbody>
</table>

*Each layer starts with a header. An SC in a syntax layer indicates that the header of that layer starts with a start code.
International Telecommunication Union (ITU)

- Formed in 1934 by combination of the International Telegraph Convention of 1865 and the International Radiotelegraph Convention of 1906

- Several “committees,” among them
  - CCITT (International Telephone and Telegraph Consultative Committee) 1956-1992
  - CCIR (International Radio Consultative Committee) 1927-1992

- Reform in 1992
  - CCITT → ITU-T
  - CCIR → ITU-R

- Any recommendation must be agreed upon unanimously by all its member states
ITU Organization with Subgroups Relevant for Video

ITU

ITU-R

ITU-T

ITU-D

Study Group 16 - Multimedia

SG1

SG2

WP1 – Modems and interface
V.34, V.25ter

WP2 – Systems
H.320 – ISDN
H.323 – LAN
H.324 – PSTN
T.120 - DATA

WP3 – Coding
G.7xx – Audio
H.261 – Video
H.263 - Video
IEC and ISO

- **IEC – International Electrotechnical Commission**
  - founded in 1906 to establish international standards for all electrical technologies
  - private, non-profit company under Swiss law

- **ISO – International Organization for Standardization**
  - Established in 1947 “to facilitate the international coordination and unification of industrial standards”
  - Private, non-profit company under Swiss law
  - Agency of the United Nations

- **Joint ISO/IEC Technical Committee 1 (JTC 1)**
  - Jointly addresses all computer-related activities
  - About 30% of total ISO and IEC standards
ISO/IEC organization with subgroups relevant for video

- **IEC**
  - **ISO**
    - **JTC1**
      - **SC29**
        - **AG**
          - **AGM**
          - **RA**
        - **WG**
          - **WG1**
          - **WG12**
          - **WG11**
            - **SG**
              - **JBIG**
              - **JPEG**
            - **SG**
              - **MHEG-5**
              - **MHEG-6**

- **Requirements**
  - **Systems**
  - **Description**
  - **Video**
  - **Audio**
  - **SNHC**
  - **Tests**
  - **Implementation Studies**
  - **Liaisons**
Requirements for a successful video coding standard

- **Interoperability**: should assure that encoders and decoders from different manufacturers work together seamlessly.
- **Innovation**: should perform significantly better than previous standard.
- **Competition**: should be flexible enough to allow competition between manufacturers based on technical merit. Only standardize bit-stream syntax and reference decoder.
- **Independence from transmission and storage media**: should be flexible enough to be used for a range of applications.
- **Forward compatibility**: should decode bit-streams from prior standard
- **Backward compatibility**: prior generation decoders should be able to partially decode new bit-streams
## Standard Development Process

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n x 384 (n= 1-5)</td>
<td>competition</td>
<td>convergence</td>
<td>verification</td>
<td>optimization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p x 64 (p= 1-30)</td>
<td>convergence</td>
<td>verification</td>
<td>optimization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m x 64 (m= 1,2)</td>
<td>competition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Overview of H.261 Standardization Process

- **RM 1 2 3 4**
- **RM 6 7 8**
- **REC. H.261**
ITU-T Rec. H.261

- International standard for ISDN picture phones and for video conferencing systems (1990)
- Image format: CIF (352 x 288 Y samples) or QCIF (176 * 144 Y samples), frame rate 7.5 ... 30 fps
- Bit-rate: multiple of 64 kbps (= ISDN-channel), typically 128 kbps including audio.
- Picture quality: for 128 kbps acceptable with limited motion in the scene
- Stand-alone videoconferencing system or desk-top videoconferencing system, integrated with PC
Image Formats

ITU-R 601

352 x 288

CIF

176 x 144

QCIF

sub QCIF
H.261 Macroblocks

- Macroblock (MB) of 16x16 pixels
- Sampling format: 4:2:0
- MB consists of 4 luminance and 2 chrominance blocks
H.261 Motion-Compensated Prediction

- Integer-pel accuracy
- One displacement vector per macroblock
- Maximum displacement vector range +/-16 horizontally and vertically
- Adaptive loop filter, separable in 1-D horizontal and vertical impulse response: \([\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]\)
- Differential encoding of motion vectors
H.261 Residual Coding

- 8x8 DCT
- Quantization
  - Uniform quantizer ($\Delta = 8$) for intra-mode DC coefficients
  - Uniform threshold quantizer ($\Delta = 2, 4, \ldots, 62$) for AC coefficients in intra-mode and all coefficients in inter-mode
- Zig-zag scan
- Run-level coding for entropy coding
  - (zero-run, value) symbols
  - zero-run: the number of coefficients quantized to zero since the last nonzero coefficient
  - value: the amplitude of the current nonzero coefficient
### H.261 Macroblock Types (VLC Table)

<table>
<thead>
<tr>
<th>Prediction</th>
<th>MQUANT</th>
<th>MVD</th>
<th>CBP</th>
<th>TCOEFF</th>
<th>VLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>0001</td>
</tr>
<tr>
<td>Intra</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>0000 001</td>
</tr>
<tr>
<td>Inter</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Inter</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>0000 1</td>
</tr>
<tr>
<td>Inter+MC</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>0000 0000 1</td>
</tr>
<tr>
<td>Inter+MC</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>0000 0001</td>
</tr>
<tr>
<td>Inter+MC</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0000 0000 01</td>
</tr>
<tr>
<td>Inter+MC+FIL</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>001</td>
</tr>
<tr>
<td>Inter+MC+FIL</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>01</td>
</tr>
<tr>
<td>Inter+MC+FIL</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0000 01</td>
</tr>
</tbody>
</table>
ITU-T Rec. H.263

- International standard for picture phones over analog subscriber lines (1995)
- Image format usually CIF, QCIF or Sub-QCIF, frame rate usually below 10 fps
- Bit-rate: arbitrary, typically 20 kbps for PSTN
- Picture quality: with new options as good as H.261 (at half rate)
- Software-only PC video phone or TV set-top box
- Widely used as compression engine for Internet video streaming
- H.263 is also the compression core of the MPEG-4 standard
H.261 vs. H.263

- Improved motion compensation
  - H.261 (1990): integer-pel accuracy, loop filter, 1 motion vector per MB
  - H.263 (1995): half-pel accuracy, no loop filter, 1 motion vector per MB
- Improved 3-D VLC for DCT coefficients *(last, run, level)*
- Reduced overhead
- Support more picture formats
- Optional features defined in annexes
  - Unrestricted motion vectors (Annex D)
  - Syntax-based arithmetic coding (Annex E)
  - Advanced prediction mode (Annex F)
    - Overlapped block motion compensation (OBMC),
    - Switch between 1 or 4 motion vectors per MB
  - PB pictures (Annex G)
- Additional optional features in H.263++. (H.263 as of 2001)
Performance of H.263 and H.261

PSNR [dB]

rate [kbps]

1) H.263
2) H.263 w/o options
3) H.261
4) H.263 w/o options, integer-pel ME
5) H.261 w/o loop filter
H.263: overlapped block motion compensation (OBMC)
Overlapped block motion compensation (OBMC)

Idea: superimpose several prediction signals, using the motion vectors from neighboring blocks also

\[
\hat{S}(x, y, t) = \sum_{b \in \text{NeighborBlocks}(x, y)} h_b(x, y) S\left[ x - d_x(b), y - d_y(b), t - \Delta t \right]
\]

[Orchard, Sullivan, 1994]
OBMC window function

- Window function usually the same for each block, i.e.

\[ h_i(x - X_i, y - Y_i) = h_j(x - X_j, y - Y_j) \]

- Reference point of blocks \( i,j \), e.g. block center

- Window function optimized for minimum m.s.e.

[Orchard, Sullivan, 1994]
Error distribution without OBMC

[Orchard, Sullivan, 1994]
Error distribution with OBMC

[Orchard, Sullivan, 1994]
**H.263: OBMC weights**

For **MV of current luminance block**:

```
4 5 5 5 5 5 5 4
5 5 5 5 5 5 5 5
5 5 6 6 6 6 5 5
5 5 6 6 6 6 5 5
5 5 6 6 6 6 5 5
5 5 6 6 6 6 5 5
5 5 5 5 5 5 5 5
4 5 5 5 5 5 5 4
```

For **remote MV of top/bottom luminance block**:

```
2 2 2 2 2 2 2 2
1 1 2 2 2 2 1 1
1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2
```

For **remote MV of left/right luminance block**:

```
2 1 1 1 1 1 1 2
2 2 1 1 1 1 2 2
2 2 1 1 1 1 2 2
2 2 1 1 1 1 2 2
2 2 1 1 1 1 2 2
2 2 1 1 1 1 2 2
2 1 1 1 1 1 1 2
2 1 1 1 1 1 1 2
```
Performance of H.263 PB-mode

1) w/o options (6.25 fps)
2) w/o options (12.5 fps)
3) PB-mode (12.5 fps)
   a) P-frames
   b) B-frames

<table>
<thead>
<tr>
<th>Network</th>
<th>System</th>
<th>Video</th>
<th>Audio</th>
<th>MUX</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSTN</td>
<td>H.324</td>
<td>H.261/3</td>
<td>G.723.1</td>
<td>H.223</td>
<td>H.245</td>
</tr>
<tr>
<td>N-ISDN</td>
<td>H.320</td>
<td>H.261</td>
<td>G.7xx</td>
<td>H.221</td>
<td>H.242</td>
</tr>
<tr>
<td>B-ISDN /ATM</td>
<td>H.321</td>
<td>H.261</td>
<td>G.7xx</td>
<td>H.221</td>
<td>Q.2931</td>
</tr>
<tr>
<td></td>
<td>H.310</td>
<td>H.261/2</td>
<td>G.7xx/MPEG</td>
<td>H.222.0/1</td>
<td>H.245</td>
</tr>
<tr>
<td>QoS LAN</td>
<td>H.322</td>
<td>H.261/3</td>
<td>G.7xx</td>
<td>H.221</td>
<td>H.242</td>
</tr>
<tr>
<td>Non-QoS LAN</td>
<td>H.323</td>
<td>H.261</td>
<td>G.7xx</td>
<td>H.225.0</td>
<td>H.245</td>
</tr>
</tbody>
</table>
H.324 Multimedia Terminals

Scope of H.324

Video I/O equipment

Audio I/O equipment

User data applications

System control

Video codec
H.261/H.263

Audio codec
G.723

Receive path delay

Data protocols
V.14, LAPM etc

Control protocol
H.245

SRP/LAPM procedures

Mux/demux
H.223

Modem
V.34/V.8

Modem control
v.25ter

PSTN network