MPEG-1/2 Standards: Motion-compensated video coding

Peter H.N. de With
(p.h.n.de.with@tue.nl)

MPEG Video / Temporal Prediction – (1)

Temporal redundancy reduction

∗ 1. By single-sided prediction
   – motion compensation should cover large area
   – (due to intermediate B pictures)
   – fallback coding required (for excessive motion or uncovered background)

∗ 2. Bidirectional motion compensation (interpolation)
   – assume linear interpolation of surrounding pictures
   – bidirectional prediction is more efficient than single-sided
   – more possibilities with uncovered objects
   – not used as reference for further coding: no error propagation in temporal coding

MPEG Video / Quantizer inter block – (1)

MPEG Quantization interframe data (predictive MBs)

∗ DC coefficients
   – Differential DC coefficients
   – Quantized and coded as AC coefficients

∗ AC coefficients
   – MPEG-1 decoder formula
   – \( F(u,v) = 2 \left( QF(u,v) + k \right) q\_scale W(u,v) / 16 \)
   – \( W(u,v) = 16 \) default, but new matrix can be loaded
   – \( k = \text{sign}(QF(u,v)) \) for inter-blocks
   – Mismatch control: if \( F(u,v) \) even \( \Rightarrow F(u,v) = F(u,v) - \text{sign}(QF(u,v)) \) value closest to zero
MPEG Video / Coding modes P&B – (1)

MPEG-1/2 coding modes for inter-coded images (P, B)

<table>
<thead>
<tr>
<th>Predictive (P)</th>
<th>Bidirectional (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion, no motion</td>
<td>Forward, from past, Backward, from future, interpolated (from both sides)</td>
</tr>
<tr>
<td>Intra (fallback), or non-intra (regular case)</td>
<td>Intra (fallback), or non-intra (regular case)</td>
</tr>
<tr>
<td>Coded (regular), or not-coded (skipped block)</td>
<td>Coded (regular), or not-coded (skipped)</td>
</tr>
<tr>
<td>Default quantization, new q-scale</td>
<td>Default quantization, new q-scale</td>
</tr>
</tbody>
</table>

MPEG-2 extensions for compensated coding modes

- Frame-based prediction (in both standards)
  - Equal to MPEG-1 (16x16 compensation blocks)
  - In a frame picture, either frame- or field-based prediction on MB level
- Field-based prediction (MPEG-2)
  - Results from interlaced pictures
- 16x8 motion compensation (MPEG-2)
  - Requires two motion vectors (1 for top- and 1 for bottom field)
  - In B interlaced pictures, even 4 vectors can be used

MPEG Video / Coding modes P&B – (2)

MPEG Video / Modes for P&B MBs – (3)

MPEG-2 extensions for compensated coding modes

- Special: Dual-prime prediction
  - 1 Motion Vector is coded in full resolution, 1 motion vector is a small differential vector (the dmv)
  - Field-based prediction: 2 vectors are derived from this information. The obtained fields are averaged to get the final prediction
  - Frame-based pictures: the averaging is done for both fields, yielding 4 field predictions.
  - This mode is only used for P-pictures, without B-pictures in between.

MPEG Video / Modes for P&B MBs – (4)

MPEG-2 extensions for compensated coding modes

- Frame-based prediction
  - 1 mv for P
  - 2 mv for B

MPEG Video / Modes for P&B MBs – (5)

MPEG-2 extensions for compensated coding modes

- Field-based prediction
  - 2 mv for field to frame for P
  - 2 mv for field to field for full interlacing P

MPEG Video / Modes for P&B MBs – (6)

MPEG-2 extensions for compensated coding modes

- Field-based prediction (continued)
  - 2 mv for P
  - 4 mv for B
MPEG Video / Modes for P&B MBs – (7)

MPEG-2 extensions for MC-coding
- Special field-based prediction: dual prime
  - Main and dmv vector
  - Scaling of vectors for dual prime prediction

MPEG Video / Decoder structure
- MPEG Video decoder hardware
  - MPEG strongly asymmetric, follows encoder decisions
  - Decoder has no ME, only MC, saves factor 3-4 in complexity

MPEG Video / Flexibility parameters
- Video sequence parameters in sequence header
  - Pixels/line, lines per picture
  - Pixel aspect ratio
  - Frame rate, bit rate
  - Required buffer size
- Conclusion MPEG-1
  - MPEG allows for a wide range of input formats
  - However, MPEG-1 is tuned to be optimal for 1.5 Mbit/s bit rate, spatial resolutions of approx. 350x250 pixels, picture rate of 20-30 frame/s, and non-interlaced pictures

MPEG Video / MPEG-1 core param'ts
- MPEG Video core parameters, purpose
  - guaranteed exchange of MPEG-coded data, which should be decodable on different systems
  - also important: bounding of encoder complexity
- MPEG-1 Core parameters
  - Pixels/line <= 720
  - Lines/frame <= 576
  - Frame rate <= 30 Hz
  - Macroblock/picture <= 396
  - Macroblock rate <= 396 x 25 Hz = 330 x 30 Hz = 9,900 Hz
  - Bit rate <= 1.86 Mbit/s
  - Buffer <= 376,832 bits

MPEG Video / Flexibility Layer Level
- GOP
  - Frame structure I,B,P, and GOP size
- Frame types
  - Intraframe I, predictive P, bidirectional B
- Slice
  - Slice size, fixed/adaptive partitioning, quantization block size
- Macroblock coding
  - Coded/skipped, motion/ no motion, intra or predicted
- Macroblock quantization
  - Adaptive or default, weighting function default or adaptive
- Motion vectors
  - One-sided, two-sided, motion estimation algorithm
MPEG Video / MPEG-2 Flexibility – (1)

- MPEG-2 extensions on flexibility
  - MPEG-2 should give a more generic set of tools for a wider range of applications
- MPEG-2 Picture formats
  - Color formats 4:2:0, 4:2:2, 4:4:4
  - Progressive, interlaced
  - More flexible frame size, more flexible pixel aspect ratio
- MPEG-2 Bit rates
  - "Composite" quality CCIR-601 at 3-5 Mbit/s
  - Component quality CCIR-601 at 8-10 Mbit/s
  - Variable bit rate, constant bit rate
  - Coded/skipped, motion/no motion, intra or predicted

MPEG Video / MPEG-2 Flexibility – (2)

- Random access
  - On slice basis, independent slice processing
- Bit stream scalability
  - Additional layering of information (partitioning)
- Compatibility
  - Backwards to MPEG-1
- Editing
  - Possible in bit stream domain
- Stability
  - Repeated coding resilience

MPEG V. / MPEG-2 Video extensions – (1)

- Interlaced video
  - Frame or field-based pictures
  - In frame case: extra MB coding options (such as frame/field motion compensation, frame or field DCT
- Hierarchical/scalable coding (optional)
  - HDTV / TV compatibility
  - MPEG-2 / MPEG-1 compatibility
  - Graceful degradation
  - Solutions: frequency scalability, spatial scalability
- Picture format
  - Parametric specification of colour sampling, colour space

MPEG V. / MPEG-2 Video extensions – (2)

- MPEG-2 extensions (cont.)
  - Coding
    - Alternate quantization tables
    - Alternate VLC tables
    - Added MB types
    - Extended precision for high-quality PQ up to HDTV

MPEG / MPEG-2 Profiles & Levels – (1)

- MPEG-2 Profiles / Levels
  - Implementation of full specification of MPEG-2 too difficult
  - Profiles serve as limited number of subsets of MPEG-2
  - Bounding of encoder/decoder complexity
- Profile
  - Limited subset of entire bit stream syntax
  - Different profiles support different features (applications)
- Level
  - Defined set of constraints imposed on the parameters in the profile bit stream

MPEG / MPEG-2 Profiles & Levels – (2)

- MPEG-2 profiles / levels
  - Example: MP@ML
- Main Profile
  - sampling 720 x 576, 4:2:0 standard
  - DCT based, frame/field DCT, frame/field MC, B frames
- Simple profile
  - no B pictures are used
- Next profile
  - scalability
    - 4:2:0 or 4:2:2 sampling
MPEG Video / MPEG-2 Profile Table

**Profile**

<table>
<thead>
<tr>
<th>Syntaxic element</th>
<th>Simple</th>
<th>Main</th>
<th>SNR</th>
<th>Spatial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>YUV format</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0, 4:2:2</td>
</tr>
<tr>
<td>Frame rate extens.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Frame rate extens. d</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Repeat first field</td>
<td>constrained</td>
<td>constrained</td>
<td>no constrain</td>
<td>no constrain</td>
<td>no constrain</td>
</tr>
<tr>
<td>Intra frame</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Scalable mode</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Spatial scalable extens.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Interlace</td>
<td>8, 9, 10</td>
<td>restricted</td>
<td>8, 9, 10</td>
<td>8, 9, 10</td>
<td>restricted</td>
</tr>
</tbody>
</table>

**Level**

<table>
<thead>
<tr>
<th>Level</th>
<th>Simple</th>
<th>Main</th>
<th>SNR</th>
<th>Spatial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
</tr>
<tr>
<td>High-1440</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
</tr>
<tr>
<td>Main</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
<td>M2P@ML</td>
</tr>
<tr>
<td>High-1440</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
<td>M2P@ML, M2P@HL</td>
</tr>
</tbody>
</table>

MPEG Video / MPEG-2 Profile-Level

Combined overview of MPEG-2 profiles and levels

**Profile**

<table>
<thead>
<tr>
<th>Syntaxic element</th>
<th>Simple</th>
<th>Main</th>
<th>SNR</th>
<th>Spatial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>YUV format</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0, 4:2:2</td>
</tr>
<tr>
<td>Frame rate extens.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Frame rate extens. d</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Repeat first field</td>
<td>constrained</td>
<td>constrained</td>
<td>no constrain</td>
<td>no constrain</td>
<td>no constrain</td>
</tr>
<tr>
<td>Intra frame</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Scalable mode</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Spatial scalable extens.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Interlace</td>
<td>8, 9, 10</td>
<td>restricted</td>
<td>8, 9, 10</td>
<td>8, 9, 10</td>
<td>restricted</td>
</tr>
</tbody>
</table>

MPEG Video / MPEG-2 Level Table

**Level**

<table>
<thead>
<tr>
<th>Syntaxic element</th>
<th>High-1440</th>
<th>High-1440</th>
<th>High-1440</th>
<th>High-1440</th>
<th>High-1440</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal vector range</td>
<td>-512</td>
<td>-1024</td>
<td>-2048</td>
<td>-3072</td>
<td>-4096</td>
</tr>
<tr>
<td>Vertical vector range</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td>-512, +509.5</td>
<td></td>
</tr>
<tr>
<td>Vertical range (field)</td>
<td>-32, +31.5</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td></td>
</tr>
<tr>
<td>Sample rate / line</td>
<td>512</td>
<td>576</td>
<td>1024</td>
<td>1152</td>
<td>2048</td>
</tr>
<tr>
<td>Frame / second</td>
<td>30</td>
<td>90</td>
<td>300</td>
<td>720</td>
<td>1440</td>
</tr>
<tr>
<td>Sample rate / frame</td>
<td>-64, +63.5</td>
<td>1024, +1023.5</td>
<td>2048, +2047.5</td>
<td>4096, +4095.5</td>
<td></td>
</tr>
<tr>
<td>Sample rate / field</td>
<td>-32, +31.5</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td></td>
</tr>
<tr>
<td>VBV buffer size (Mbit)</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
</tr>
</tbody>
</table>

MPEG Video / MPEG-2 Profile-Level

Combined overview of MPEG-2 profiles and levels

**Profile**

<table>
<thead>
<tr>
<th>Syntaxic element</th>
<th>Simple</th>
<th>Main</th>
<th>SNR</th>
<th>Spatial</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>YUV format</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0</td>
<td>4:2:0, 4:2:2</td>
</tr>
<tr>
<td>Frame rate extens.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Frame rate extens. d</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Repeat first field</td>
<td>constrained</td>
<td>constrained</td>
<td>no constrain</td>
<td>no constrain</td>
<td>no constrain</td>
</tr>
<tr>
<td>Intra frame</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Scalable mode</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Spatial scalable extens.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Interlace</td>
<td>8, 9, 10</td>
<td>restricted</td>
<td>8, 9, 10</td>
<td>8, 9, 10</td>
<td>restricted</td>
</tr>
</tbody>
</table>

MPEG Video / MPEG-2 Level Table

**Level**

<table>
<thead>
<tr>
<th>Syntaxic element</th>
<th>Low</th>
<th>Main</th>
<th>High-1440</th>
<th>High-1440</th>
<th>High-1440</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal vector range</td>
<td>-512</td>
<td>-1024</td>
<td>-2048</td>
<td>-3072</td>
<td>-4096</td>
</tr>
<tr>
<td>Vertical vector range</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td>-512, +509.5</td>
<td></td>
</tr>
<tr>
<td>Vertical range (field)</td>
<td>-32, +31.5</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td></td>
</tr>
<tr>
<td>Sample rate / line</td>
<td>512</td>
<td>576</td>
<td>1024</td>
<td>1152</td>
<td>2048</td>
</tr>
<tr>
<td>Frame / second</td>
<td>30</td>
<td>90</td>
<td>300</td>
<td>720</td>
<td>1440</td>
</tr>
<tr>
<td>Sample rate / frame</td>
<td>-64, +63.5</td>
<td>1024, +1023.5</td>
<td>2048, +2047.5</td>
<td>4096, +4095.5</td>
<td></td>
</tr>
<tr>
<td>Sample rate / field</td>
<td>-32, +31.5</td>
<td>-64, +63.5</td>
<td>-128, +127.5</td>
<td>-256, +255.5</td>
<td></td>
</tr>
<tr>
<td>VBV buffer size (Mbit)</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
<td>2048</td>
</tr>
</tbody>
</table>

5LSE0 - Mod 10

Part 3

Towards the future..., Lower bit rate and Video Objects

MPEG Outlook / MPEG-4 – (1)

* Not only bit streams and bit maps
* Abstract object-oriented multimedia
* A/V Programs as SW programs
* Elements can be described independently and combined only at playback time
* Elements can include stills, digital video, 3D graphics, text, speech
* Elements can be combined intelligently – video texture on 3D objects

MPEG Outlook / MPEG-4 Objects – (2)

* Going from images to individual objects...
MPEG Outlook / MPEG-4 aspects – (3)

* System aspects and manipulation
  - Demultiplex elementary streams
  - Decompression
  - Downstream data, user events, etc.

MPEG Coding / Conclusions – (1)

* MPEG-1 provides a suitable platform for 1 Mbit/s applications, whereas MPEG-2 enables TV up till HDTV coding and contains many extensions for interlaced images
* MPEG Video Compression is based on motion-compensated DCT coding, with extensive VLC usage of various signal components
* The complete specification for audio, video and data, together with system has resulted in wide acceptance of MPEG-1 and MPEG-2 for many applications
* MPEG-4 will be the next important step and is based on several developments: Internet, Blu-ray disc, mobile phone

MPEG Coding / Conclusions – (2)

* MPEG-4 has proven to be the next important step and is more pluriiform
* Resulted in several standards due to Internet development
  - 1. MPEG-4 AVC / H.264 for HDTV-optimized coding on Blu-ray Disc (stream-based decoding)
    • Again Motion-Compensated DCT Coding but optimized
    • Halved bit rate of MPEG-2
  - 2. MPEG-4 Object-Oriented Coding, for low bit-rates and interactive or conditional access to individual parts
    • Objects on Internet (not yet broadly applied)