Agenda

Video Coding Standards

• Overview and timeline
• Comparison between video (visual) coding standards
• Advanced Video Coding (AVC): H.264/ MPEG-4 Part 10
• High Efficiency Video Coding (HEVC): H.265/ MPEG-H Part 2

Other Developments
# Video Compression

## Typical applications and bit-rates

<table>
<thead>
<tr>
<th>Application</th>
<th>Bit-Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncompressed</td>
</tr>
<tr>
<td>Slow-motion video (10 frames/s) 170x120, 24bits/pixel</td>
<td>5.07 Mb/s</td>
</tr>
<tr>
<td>Video conference (15 frames/s) 352x240, 24bits/pixel</td>
<td>30.41 Mb/s</td>
</tr>
<tr>
<td>Digital video on CD-ROM (30frames/s), 352x240, 24bits/pixel</td>
<td>60.83 Mb/s</td>
</tr>
<tr>
<td>Broadcast video (30 frames/s) 720x480, 24bits/pixel</td>
<td>248.83 Mb/s</td>
</tr>
<tr>
<td>HDTV (30 frames/sec or higher) 1280x720, 24b/pixel</td>
<td>0.66 Gb/s</td>
</tr>
<tr>
<td></td>
<td>1.5 Gb/s</td>
</tr>
</tbody>
</table>

Recently, applications to 4K and 8K Ultra HD (UHD)
Video Compression

Picture formats:

<table>
<thead>
<tr>
<th>Image Format</th>
<th>Color Component</th>
<th>30 fps rate</th>
<th>25 fps rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCIR 601 (ITU-R Recommendation BT.601-2)</td>
<td>Y Cb, Cr</td>
<td>720x480</td>
<td>720x576</td>
</tr>
<tr>
<td></td>
<td></td>
<td>360x480</td>
<td>360x576</td>
</tr>
<tr>
<td>SIF (Source Input Format)</td>
<td>Y Cb, Cr</td>
<td>360x240</td>
<td>360x288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>180x120</td>
<td>180x144</td>
</tr>
<tr>
<td>CIF (Common Intermediate Format)</td>
<td>Y Cb, Cr</td>
<td>352x288</td>
<td>352x288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>176x144</td>
<td>176x144</td>
</tr>
<tr>
<td>QCIF</td>
<td>Y Cb, Cr</td>
<td>176x144</td>
<td>176x144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>88x72</td>
<td>88x72</td>
</tr>
</tbody>
</table>

Note: CCIR stands for International Consultative Committee on Radio and is now known as ITU-R, which stands for International Telecommunications Union – Radio Sector
Video Compression

DPCM approach in interframe coding:

Input

- Motion compensated prediction

Memory

+
Video Compression

Transform DCT-based Video Coding

- Variable Bit Rate (VBR) Encoding: fixed quantization scale $q$
- Constant Bit Rate (CBR) Encoding: $q$ is adapted
Video Compression

Transform DCT-based Video Coding

- Macroblock: minimum coding unit

![Diagram showing different color planes and their sizes: 4:4:4, 4:2:2, 4:2:0.](image)
Video Compression
Transform DCT-based Video Coding

• Motion Estimation and Compensation

- I frame: intra frame (also called key frame) is intra coded
- P frame: forward predicted frame from one previously decoded frame
- B frame: bi-directional predicted frame from one or two previously coded frames
- Group of Pictures (GoP): I BB P BB P BB P BB ... P BB I
Video Compression

One can also combine DPCM with subband coding (non-standard)
Agenda

Video Coding Standards

• Overview and timeline
• Comparison between video (visual) coding standards
• Advanced Video Coding (AVC): H.264/ MPEG-4 Part 10
• High Efficiency Video Coding (HEVC): H.265/ MPEG-H Part 2
• Scalable Video Coding (SVC)
• Multi-View Video Coding (MVC)
Video Compression

Compression standards

• Interoperability of different encoders and decoders.
• Reduction of high cost of video compression codecs.
• Increased focus in applied research.
Progression of Video Coding Standards

- Two main standard bodies:
  - ITU-T: International Telecommunications Union – Telecommunications Sector (was known previously as CITT)
  - ISO/IEC: International Organization of Standardization /International Electrotechnical Commission

- Video Working Groups:
  - ISO-MPEG: Motion Picture Experts Group (ISO/IEC JTC1/SC29/WG11)
  - ITU-T VCEG: Video Coding Experts Group (ITU-T SG16 Q.6)
  - Joint Video Team: MPEG and VCEG jointly working together
• ITU-T more focused on telecommunications applications (such as real time video communication applications, video conferencing, video phone).
• ISO more focused on consumer applications (such as video storage (DVD), broadcast video, and video streaming)
• Only bitstream syntax standardized (Decoder).
The Scope of Video Coding Standards

Only \textit{decoder compatible bitstream syntax} are standardized. Encoding, pre- and post- processing operations are not specified.

- Leave enough room for optimization and innovation.
- Permits complexity reduction for implementation.

No guarantee of quality.

Reference software implementation
Agenda

Video Coding Standards

• Overview and timeline
• Comparison between video (visual) coding standards
• Advanced Video Coding (AVC): H.264/ MPEG-4 Part 10
• High Efficiency Video Coding (HEVC): H.265/ MPEG-H Part 2
• Scalable Video Coding (SVC)
• Multi-View Video Coding (MVC)
Overview

ISO-MPEG standards (MPEG: Motion Picture Experts Group)

• MPEG-1: digital storage media coding at up to 1.5 Mbps (1991).
• MPEG-3: cancelled
• MPEG-4: coding at very low bit rates (≤64Kbps)(1999)
• MPEG-H HEVC: (ITU-T recommendation H.265): High efficiency video coding targeting improved coding efficiency and high-resolution videos up to 8K (UHD, 8192x4320).
Overview

ITU-T Recommendations (VCEG: Video Coding Experts Group)

- **H.261**: video conferencing at $p \times 64$Kbps; video teleconferencing over ISDN (1990).
- **H.262/MPEG 2**: jointly with ISO
- **H.263+; H.263++**: “short term” efforts adding enhancements to H.263.
- **H.26L**: “long-term” effort aimed at developing new standard for low bit-rate visual communications.
- **H.265/MPEG-H (Part 2) HEVC (High Efficiency Video Coding)** - – Joint ITU-T/ISO effort
ITU-T H.261

- First major video compression standard
- Targeted for 2-way video conferencing and for ISDN networks that supported 40Kbps to 2Mbps.
- Supported resolutions include CIF and QCIF.
- Chrominance resolution subsampling 4:2:0
- Low complexity and low delay to support real-time communications
- Only I and P frames. No B frames.
- Full-pixel accuracy motion estimation
- 8x8 block-based DCT coding of residual
- Fixed linear quantization across all AC coefficients of DCT
- Run-length coding of quantized DCT coefficients followed by Huffman coding for DCT and motion information
- Loop filtering (simple digital filter applied on the block edges) applied to reference frames to reduce blocking artifacts.
ISO/IEC MPEG-1

• First video compression standard developed by the ISO
• Targeted for storage and retrieval of moving pictures and audio on digital media such as video CDs with target rates around 1.5 Mbps (quality same or better than VHS).
• Similar to H.261 with enhancements
• B frames in addition to I and P frames
• Adaptive perceptual quantization: separate quantization scale factor applied to each AC coefficient to optimize the human visual perception
• Only progressive (non-interlaced) video is supported.
• Intraframe coding: I frames in addition to D frames (DC frames in which only DC components from blocks, useful for fast search applications).
• Interframe coding: P and B frames.
• Wide range of input resolutions.
ISO/IEC MPEG-1

Supported picture formats and resolutions

<table>
<thead>
<tr>
<th>Image Format</th>
<th>Color Component</th>
<th>30 fps rate (NTSC)</th>
<th>25 fps rate (PAL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCIR 601</td>
<td>Y, Cb, Cr</td>
<td>720x480</td>
<td>720x576</td>
</tr>
<tr>
<td></td>
<td></td>
<td>360x480</td>
<td>360x576</td>
</tr>
<tr>
<td>SIF (Source Input Format)</td>
<td>Y, Cb, Cr</td>
<td>360x240</td>
<td>360x288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>180x120</td>
<td>180x144</td>
</tr>
<tr>
<td>CIF (Common Input Format)</td>
<td>Y, Cb, Cr</td>
<td>352x288</td>
<td>352x288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>176x144</td>
<td>176x144</td>
</tr>
<tr>
<td>QCIF</td>
<td>Y, Cb, Cr</td>
<td>176x144</td>
<td>176x144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>88x72</td>
<td>88x72</td>
</tr>
</tbody>
</table>
ISO/IEC MPEG-1

Color format supported: 4:2:0 (chrominance components have half the resolution of luminance in both directions).

Macroblock: minimum coding unit
ISO/IEC MPEG-1

Frame types

- P frames: based on one previous I or P frame.
- B frames: based on one previous I or P frame and one future P frame (Bi-directional Prediction).
  - The frames IBBPBBI are re-ordered and sent as IPBBPIBB.
  - Four prediction modes possible for each macroblock in B-frame: intra, forward, backward, or average of both.
## ISO/IEC MPEG-1

Prediction modes for macroblock in B frames:

<table>
<thead>
<tr>
<th>Macroblock type</th>
<th>Predictor</th>
<th>Prediction error</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra</td>
<td>$\hat{I}_1(\vec{x}) = 128$</td>
<td>$I_1(\vec{x}) - \hat{I}_1(\vec{x})$</td>
<td>M</td>
</tr>
<tr>
<td>Forward Predicted</td>
<td>$\hat{I}_1(\vec{x}) = \hat{I}<em>0(\vec{x} + \vec{m}</em>{01})$</td>
<td>$I_1(\vec{x}) - \hat{I}_1(\vec{x})$</td>
<td>M</td>
</tr>
<tr>
<td>Backward Predicted</td>
<td>$\hat{I}_1(\vec{x}) = \hat{I}<em>2(\vec{x} + \vec{m}</em>{21})$</td>
<td>$I_1(\vec{x}) - \hat{I}_1(\vec{x})$</td>
<td>M</td>
</tr>
<tr>
<td>Average</td>
<td>$\hat{I}_1(\vec{x}) = \frac{1}{2} \left[ \hat{I}<em>0(\vec{x} + \vec{m}</em>{01}) + \hat{I}<em>2(\vec{x} + \vec{m}</em>{21}) \right]$</td>
<td>$I_1(\vec{x}) - \hat{I}_1(\vec{x})$</td>
<td>M</td>
</tr>
</tbody>
</table>

- $\hat{I}_1$: prediction of current frame $I_1$
- $\hat{I}_0$: previous decoded I or P frame
- $\hat{I}_2$: next decoded P or I frame
- $\vec{x}$: coordinate of picture element
- $\vec{m}_{01}$: motion vector relative to the reference frame $\hat{I}_0$
- $\vec{m}_{21}$: motion vector relative to the reference frame $\hat{I}_2$
ISO/IEC MPEG-1
Bi-directional Prediction

• Advantages
  – Higher coding efficiency.
  – No uncovered background problems.
  – Effect of noise can be decreased by averaging between past and future reference frames.
  – No prediction error propagation.
  – Increased frame rate with few extra bits.

• Disadvantages
  – Increasing the number of B frames between references results in a decreased correlation between reference frames and between B frames and reference frames.
    » For a large class of scenes, references places at 1/10th second interval resulting in 2 B frames between the reference I and P frames; example: IBBPBBPBB..PBBI.
  – At least two frames need to be stored at decoder
  – Increased delay
Motion Vectors are coded differentially relative to previous adjacent macroblock, then difference is variable-length coded. B-macroblock type is also sent.
ISO/IEC MPEG-1

Decoder

Step size
Buffer -> VLC decoder -> Q⁻¹ -> IDCT

Motion vectors
Previous frame <- Buffer + Future frame

MacroBlock Type
Motion compensation
Mux

Decoded image

©2007-2015 by Lina J. Karam
Intro to Image and Video Compression
ISO/IEC MPEG-2 / ITU-T H.262

• Targeted for digital television and was deployed in consumer video applications, digital satellite, digital cable, DVDs, and, lately, HDTV.
• Support for both progressive and interlaced video.
• Backward compatible with MPEG-1.
• Support for higher video resolution including standard television resolutions (such as interlaced 720x480 at 60 fields per second for NTSC and 720x576 at 50 fields per second for PAL and other).
• Support for wider motion estimation search ranges for higher resolution video with half-pixel accuracy
• Optimized for applications at 10 Mbps, but can be successfully used for HDTV applications at 80 to 100 Mbps.
• Color formats supported: 4:2:0, 4:2:2, 4:4:4.
• Frame-based and field-based coding
• Range of coding support divided in profiles (set of algorithmic tools) and levels (constraints on parameters sets such as picture size and bit-rates).
ISO/IEC MPEG-2 / ITU-T H.262

• Profiles defined for scalable video applications with scalable coding tools to allow multiple layer video coding, including temporal, spatial, and SNR scalability, and data partitioning.

• MPEG-2 Main Profile supports single layer coding (non scalable) and is the one that is widely deployed.

• MPEG-2 non-scalable (single layer) profiles
  – Simple profile: no B frames for low-delay applications
  – Main profile: support for B frames; can also decode MPEG-1 video

• MPEG-2 scalable profiles
  – SNR profile: adds enhancement layers for DCT coefficient refinement
  – Spatial profile: adds support for enhancement layers carrying the coded image at different spatial resolutions (sizes)
  – High profile: adds support for coding a 4:2:2 video signal and includes scalability tools of SNR and spatial profiles
ISO/IEC MPEG-2 / ITU-T H.262

Color formats and macroblocks

Frame representation
ITU-T H.263: Main Features

Enhancement of H.261

Baseline algorithm

- Half-pixel accuracy motion estimation and compensation.
- MV differentially coded, median MV prediction.
- 8 by 8 discrete cosine transform and uniform quantization.
- Variable length coding of DCT coefficients.

Four optional modes

- Unrestricted Motion Vector (UMV) mode
  - increased motion vector range with frame boundary extrapolation.
- Advanced Prediction (AP) mode
  - 4 MVs per macroblock.
  - Overlapped Block Motion Compensation (OBMC).
- PB frame mode: bi-directional prediction.
- Arithmetic coding mode.
ITU-T H.263: Video Source Format

Support for five standard picture formats (more than H.261).

<table>
<thead>
<tr>
<th>Picture Format</th>
<th>Y columns</th>
<th>Y rows</th>
<th>C columns</th>
<th>C rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-QCIF</td>
<td>128</td>
<td>96</td>
<td>64</td>
<td>48</td>
</tr>
<tr>
<td>QCIF</td>
<td>176</td>
<td>144</td>
<td>88</td>
<td>72</td>
</tr>
<tr>
<td>CIF</td>
<td>352</td>
<td>288</td>
<td>176</td>
<td>144</td>
</tr>
<tr>
<td>4CIF</td>
<td>704</td>
<td>576</td>
<td>352</td>
<td>288</td>
</tr>
<tr>
<td>16CIF</td>
<td>1408</td>
<td>1152</td>
<td>704</td>
<td>576</td>
</tr>
</tbody>
</table>

Y: luminance component, \( C_b \) \( C_r \): chrominance component. Color conversion: \( Y \ C_b \ C_r \\leftrightarrow R \ G \ B \).

Chrominance sampling 4:2:0 (same as H.261).
Without compression, 30 frames/s QCIF needs bandwidth: 
\[ 30 \times 176 \times 144 \times 1.5 \times 8 = 9,123,840 \text{ bits/s}. \]
700MB CD can only store 10.7 minutes QCIF, 2.7 minutes CIF size video at 30 frames/s.
ITU-T H.263

Performance

• About 3 to 4 dB PSNR improvement over H.261 at bit-rates less or equal to 64Kbits/s.
• 30% saving in bit-rate as compared to MPEG-1.

Design flexibility (things not specified by standard)

• H.263 standard inherently has the capability to adapt to varying input video content.
• Frame level: Intra or Inter or skipped.
• Macroblock (MB) level:
  – Intra, Inter or Un-coded.
  – One MV or 4 MVs
  – Quantizer parameter, QP value.
    ▪ Constant QP $\Rightarrow$ almost constant quality, variable bit-rate.
    ▪ Varying QP $\Rightarrow$ variable quality, try to achieve almost constant bit-rate.
ITU-T H.263: Encoder Diagram

Coding Control

Transform → Quantization → Quantized Index

Inverse Quantization → Inverse Transform

Frame Store

Motion Estimation → Motion Compensation

Predicted Frame

Video Source → Inter

Intra

Bit Stream → Entropy Coding

Motion Vectors
ITU-T H.263

Differential Coding of MVs

- The predictor is the median value of MV1, MV2, MV3.
- Horizontal and vertical component are coded separately.
- If neighboring MB Intra coded, or was not coded, the candidate predictor is set to 0.
- If neighboring MB outside frame or GOB boundary, the candidate predictor is set to 0.
ITU-T H.263

Two additional modes for MEC

• Unrestricted Motion Vector (UMV) mode
  – Motion vectors allowed to point outside frame boundary.
    ▪ Frame boundary extension.
    ▪ Significant gain for movement along frame edges.
  – Extension of motion vector range to [-31.5, 31.5].
    ▪ Improve coding efficiency for high motion videos.

• Advanced Prediction (AP) mode
  – Four motion vector per MB: one for each 8*8 Y block.
  – MV allowed to point outside frame as in UMV mode.
  – Overlapped Block Motion Compensation (OBMC) is used for four 8*8 Y block.
    ▪ Significant improvement in subjective quality by reducing blocking artifacts.
ITU-T H.263
Quantization and Entropy Coding

Quantization: uniform scalar quantization

- Intra DC: \( \text{LEVEL} = \frac{\text{COF}}{8} \)
- Intra AC: \( \text{LEVEL} = \frac{|\text{COF} - \text{QP}/2|}{2 \times \text{QP}} \)
- Inter MB: \( \text{LEVEL} = \frac{|\text{COF}|}{2 \times \text{QP}} \)

- QP value: up to 5 bits
- Frame header (PQUANT), GOB header (GQUANT): 5 bits
- MB header (DQUANT): 2 bits, -1, -2, 1, 2.

Entropy coding: Variable Length coding (VLC)

- Zig-zag scan of quantized DCT coefficients.
- Three dimensional run-length VLC (LAST, RUN, LEVEL).
- MV difference and various control info are also VLC coded.
ITU-T H.263: Spatial Division of Frame
Block $\rightarrow$ Macro Block $\rightarrow$ Group of blocks $\rightarrow$ Frame

- Block: 8*8
- Macro block (MB): 4 $Y$ block, 1 $C_r$ block, 1 $C_b$ block.
- Group of blocks (GOB): integer number of MB rows.
  - 18, 9, 6 GOBs for CIF, QCIF, and sub-QCIF, respectively.

![Diagram of GOB structure]

<table>
<thead>
<tr>
<th>GOB 1</th>
<th>GOB 2</th>
<th>GOB 3</th>
<th>GOB 4</th>
<th>GOB 5</th>
<th>GOB 6</th>
<th>GOB 7</th>
<th>GOB 8</th>
<th>GOB 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_1 )</td>
<td>( Y_2 )</td>
<td>( Y_3 )</td>
<td>( Y_4 )</td>
<td>( C_b )</td>
<td>( C_r )</td>
<td>( Y_1 )</td>
<td>( Y_2 )</td>
<td>( Y_3 )</td>
</tr>
</tbody>
</table>
ITU-T H.263 Expansions

H.263+: new modes to support new features.

- Improve compression efficiency: advanced intra-coding, de-blocking filter, alternate inter VLC, modified quantization, improve PB frames.
- Custom and flexible video formats.
- Scalability: SNR (fidelity), spatial (resolution), temporal.

H.263++: more new modes for compression efficiency improvement and error resilience.
ISO/IEC MPEG-4

• Targeted for low bit-rate applications, increased error robustness to support wireless networks
• ITU-T H.263 and its expansions (annexes) formed the core for many of the coding tools in MPEG-4.
• New tools for coding video objects (irregular-shaped regions) and to support merging graphic objects with video: most of these tools have not yet gained deployment in products
• MPEG-4 Visual include 19 defined profiles: 15 profiles for “natural” video coding, 3 profiles for “synthetic” video, and one hybrid profile.
• MPEG-4 Visual profile for “natural” video range from the Simple profile (coding of rectangular video frames) through profiles for arbitrary-shaped and scalable object coding (including Fine Granular Scalability Profile) to profiles for coding of studio-quality video.
ISO/IEC MPEG-4

- MPEG-4 simple profile is built based on ITU-T H.263 with added tools for improved compression
  - Unrestricted Motion Vectors (support prediction when objects partially move outside the frame)
  - Variable block size motion compensation at either 16x16 or 8x8
  - Context-adaptive intra DCT prediction: DC/AC DCT coefficients can be predicted from neighboring blocks (to left or above of current block)
  - Extended dynamic range for DCT AC coefficients to support high-fidelity video: [-2047, 2047] versus [-127,127] in H.263
ISO/IEC MPEG-4

• MPEG-4 error resiliency features to support packet loss recovery
  – Slice resynchronization
  – Data partitioning: separating motion part from DCT part with a unique motion boundary marker
  – Reversible VLC: allows decoding backwards as well as forwards
  – New prediction (NEWPRED): decoder can use a feedback channel to request additional information from encoder in case of packet losses.
ISO/IEC MPEG-4

• MPEG-4 Advanced Simple Profile (ASP) starts from simple profile and adds
  – B frames
  – Interlaced tools (for Level 4 and up)
  – Quarter-pixel accuracy motion estimation
  – Option for global motion compensation

• MPEG-4 Deployment
  – Internet streaming, Apple’s Quicktime 6
  – MPEG-4 simple profile used for mobile streaming
  – MPEG-4 ASP forms the foundation of the proprietary DivX codec and open source Xvid codec
ISO/IEC MPEG-4

Segmentation into VOPs
ITU-T H.26L – precursor of H.264/AVC

Intra prediction
- 16*16 (4 modes) or 4*4 (6 modes)

Inter prediction
- Variable block size: 15% bit-rate savings
- 1/4 (QCIF) and 1/8 (CIF) pixel accuracy motion estimation: 20% savings
- Multiple reference picture selection: 5-10% savings

4*4 integer transform and scalar quantization.

Entropy coding:

<table>
<thead>
<tr>
<th>Mode 1</th>
<th>Mode 2</th>
<th>Mode 3</th>
<th>Mode 4</th>
<th>Mode 5</th>
<th>Mode 6</th>
<th>Mode 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>One 16x16 block</td>
<td>Two 8x16 blocks</td>
<td>Two 16x8 blocks</td>
<td>Four 8x8 blocks</td>
<td>Eight 4x8 blocks</td>
<td>Eight 8x4 blocks</td>
<td>Sixteen 4x4 blocks</td>
</tr>
<tr>
<td>One motion vector</td>
<td>Two motion vectors</td>
<td>Two motion vectors</td>
<td>Four motion vectors</td>
<td>Eight motion vectors</td>
<td>Eight motion vectors</td>
<td>Sixteen motion vectors</td>
</tr>
</tbody>
</table>

0 1 0
1 2 3
4 5 6 7
8 9 10 11
12 13 14 15
ITU-T H.26L – precursor of H.264/AVC

Entropy coding:

- Universal Variable-Length Coding (UVLC)
- Context-Adaptive Binary Arithmetic Coding (CABAC) (8% savings)
Milestones in Video Coding

Variable block size (16x16 – 4x4) + quarter-pel + multi-frame motion compensation (H.26L, 2001)

Variable block size (16x16 – 8x8) (H.263, 1996) + quarter-pel motion compensation (MPEG-4, 1998)


Frame Difference coding (H.120 1988)

Bit-rate Reduction: 75%

PSNR [dB]

35 36 37 38 39 40

Rate [kbit/s]

0 100 200 300

Intraframe DCT coding (JPEG, 1990)

Integer-pel motion compensation (H.261, 1991)

Conditional Replenishment (H.120)

Foreman 10 Hz, QCIF 100 frames

©2007-2015 by Lina J. Karam
Intro to Image and Video Compression 45
Agenda

Video Coding Standards

• Overview and timeline
• Comparison between video (visual) coding standards
• Advanced Video Coding (AVC): H.264/ MPEG-4 Part 10
• High Efficiency Video Coding (HEVC): H.265/ MPEG-H Part 2
• Scalable Video Coding (SVC)
• Multi-View Video Coding (MVC)
Adopted by DVB, ATSC, 3GPP(2), DMB, MediaFLO, HD-DVD, Blu-Ray, video streaming (e.g., Amazon, Hulu. Netflix, YouTube)

Important profiles:
- Baseline profile: low-delay, low-complexity applications (videophones)
- **Main profile**: high coding efficiency applications (broadcast, storage)
- **High profile** or **Fidelity Range extension (FRExt)**: HD video (broadcast, storage)

New and improved encoding concepts:
- I, P, B slices
- Spatial intra prediction
- 4x4 integer transform (plus 8x8 integer transform for High profile)
- In-loop deblocking filter
- Variable block sizes for MCP with quarter-pixel precision (16x16, 16x8, 8x8, 8x4, 4x4)
- **Generalized B frames**
- Multiple reference frames
- Unequal weighting of prediction blocks
- Context Adaptive Binary Arithmetic Coding (CABAC)
- Context Adaptive Variable Length Coding (CAVLC)
- Several error-resiliency tools (FMO, ASO, RS, DP)
- Support for interlaced video
JVT (ITU-T/ISO) H.264/AVC/MPEG 4 Part 10

©2007-2015 by Lina J. Karam
Intro to Image and Video Compression
Two layers:

- **Video Coding Layer (VCL)**
  - Coded video data

- **Network Abstraction Layer (NAL)**
  - VCL data mapped into NAL units
  - NAL units can be transmitted over packet-based networks or a bitstream transmission link or stored in a file.
  - Each NAL unit contains a NAL header and a Raw Byte Sequence Payload (RBSP), which consists of coded video data or header information

<table>
<thead>
<tr>
<th>NAL header</th>
<th>RBSP</th>
<th>NAL header</th>
<th>RBSP</th>
<th>NAL header</th>
<th>RBSP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequence of NAL units
JVT (ITU-T/ISO) H.264/AVC/MPEG 4 Part 10

Main Features

• Transform
  - 4x4/8x8 integer transform
  - 2x2/4x4 secondary Hadamard transform

• Inter-frame Prediction:
  - variable block-size, quarter-pixel motion estimation
  - multiple frames (long-term memory)

• Context-Adaptive Entropy Coding: CAVLC and CABAC

• In-loop deblocking filter

• Data partitioning: frame can be partitioned into *slices*.
Multiple Slice Groups (known as Flexible Macroblock Ordering or FMO in previous drafts)

A slice group may contain one or more slices. Each slice contains a subset of Macroblocks

Baseline profile: I slices (only intra-coded macroblock predicted from sample within same slice) and P slices (inter-, intra-coded, skipped macroblocks).

Main Profile: B slices

©2007-2015 by Lina J. Karam Intro to Image and Video Compression
4x4 Luma Intra-prediction Modes

9 intra prediction modes

©2007-2015 by Lina J. Karam
Examples of 4x4 Luma Intra-prediction Modes

Mean of available decoded samples (shaded).
16x16 Luma and 8x8 Chroma Intra-prediction Modes

• 4 modes available
  - Vertical extrapolation
  - Horizontal extrapolation
  - DC: mean of upper and left samples
  - Plane or linear fitting between upper and left sample (diagonal direction \( \backslash / \))
JVT (ITU-T/ISO) H.264/AVC/MPEG 4 Part 10

Inter-Prediction

- Segmentation of Macroblocks for Motion Estimation

<table>
<thead>
<tr>
<th>M Types</th>
<th>16x16</th>
<th>16x8</th>
<th>8x16</th>
<th>8x8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8x8 mode</td>
<td>8x8</td>
<td>8x4</td>
<td>4x8</td>
<td>4x4</td>
</tr>
<tr>
<td>8x8 Types</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

- Corresponding chroma components have half the horizontal and vertical resolutions of luma
- Chroma motion vectors are obtained by dividing by 2 the horizontal and vertical components of the luma motion vectors.
Inter-Prediction

• quarter-pixel accuracy motion vectors for luma
  - Need to interpolate to get samples at sub-sample positions
  - Nearby coded samples used in interpolating sub-samples
• eighth-pixel accuracy motion vectors for chroma
• Motion vector (MV) prediction
  - High correlation between motion vectors of neighboring partitions
  - Predicted motion vector formed from previously calculated neighboring MVs and difference is encoded
  - Prediction mode depends on partition size and available nearby motion vectors. Typically, predicted MV is median of neighboring MVs or the MV of a neighboring partition
Multiframe Motion Compensation
In-Loop Deblocking Filter

• Filtering applied after inverse transform to horizontal or vertical edges of 4x4 block

• Filtering strength depends on gradient of image samples across boundary and on boundary strength: BS = 4 (Strong filtering) to 0 (no filtering). BS depends on block type and location of boundary

Filter only if

1) BS > 0 and
2) |p0-q0| < alpha
   and |p1-p0| < beta
   and |q1-q0| < beta

alpha and beta specified by standard and depend on QP
Variable block size
(16x16 – 4x4) +
quarter-pel +
multi-frame
motion compensation+
R-D Optimization
( H.264, 2004)
Variable block size
(16x16 – 4x4) +
quarter-pel + multi-frame
motion compensation + R-D Optimization
(H.264 2004)
Variable block size
(16x16 – 4x4) +
quarter-pel + multi-frame
motion compensation + R-D Optimization
(H.264 2004)
Agenda

Video Coding Standards

- Overview and timeline
- Comparison between video (visual) coding standards
- Advanced Video Coding (AVC): H.264/ MPEG-4 Part 10
- High Efficiency Video Coding (HEVC): H.265/ MPEG-H HEVC
JVT (ITU-T/ISO) H.265/HEVC

Main Features
• 30% to 50% bit-rate savings as compared to H.264/AVC for same visual quality
• Built-in parallelism allowing higher throughput and low power implementations
  - enables support for up to 8K UHD at 120 frames per second
• Adopted for Ultra HD 4K streaming (e.g., Netflix) and Samsung Galaxy S4 and higher models
JVT (ITU-T/ISO) H.265/HEVC

More prediction modes (35 vs 10): DC, Planar, and 33 angular directions

High-Throughput CABAC

More and larger

Larger blocks (8x8 vs 4x4) Parallelism

Less signaling overhead with larger blocks

CTU can have a mix of intra and inter-coded blocks

N= 64, 32, 16

N= 64, 32, 16

Larger and flexible NxN blocks: Coding Tree Unit (CTU)

N= 64, 32, 16

Larger and flexible NxN blocks: Coding Tree Unit (CTU)
Coding Tree Unit

- Divided into Coding Units (CU) using Quatree (two levels shown here)
- Each CU is divided into Predictions Units (PU)

- Intra-Coded CU can only be divided into square PUs: divide into 4 or keep undivided as one PU
- Inter-Coded CU can be divided into square or non-square partitions
- Each side of a PU should be greater or equal to 4 pixels
JVT (ITU-T/ISO) H.265/HEVC

Transform
• 32x32, 16x16, 8x8, and 4x4 integer transforms
  - Two types of 4x4 transforms: DST for intra-coding and DCT for inter-coding
  - DCT for all other sizes
• Implementation with increased parallelism
• 5% to 10% increase in coding efficiency
• Higher complexity relative to H.264/AVC
JVT (ITU-T/ISO) H.265/HEVC

• Profiles: each profile defines specs and coding tools for target application
  - Main: 8 bits/sample
  - Main 10: 10 bits/sample
  - Main still picture (intra-frame or still picture coding)

• Levels: each level sets max supported resolution and frame rate
  - Levels 1 to 6 (including sub-levels e.g. 6.1, 6.2)
  - Example: Level 5 supports 4K (4096x2160) at 30 frames per second
Other Coding Standards

- Audio-Video Standard (AVS) – China
- Open Video Coding Standards: VP9/VP10
Agenda

Other Developments

• Scalable video coding (SVC)
• Multi-view video coding (MVC)
• Distributed video coding
• Texture-based video coding
H.264 Scalable Video Coding (SVC) Extension

Scalability modes:

• Temporal (frame rates)
• Spatial (resolution)
• SNR: coarse (CGS), medium (MGS), fine (FGS) granular scalability
• Combined spatial-temporal-SNR scalability

Important concepts:

• **Hierarchical B frames** (temporal scalability and high R-D efficiency)
• Single-loop decoding
• Inter-layer prediction for spatial and CGS (Motion, Residual, Intra prediction)
• Extended spatial scalability (arbitrary resolution factors, cropping)
• MGS and FGS based on drift controlled prediction structure
• MGS refinements based on CGS syntax with limited extraction points
Temporal scalability prediction structures

• Hierarchical B frames (SVC and compatible with H.264/AVC)

• Classical B frames (by default in H.264/AVC, MPEG-4 Part 2)
H.264 Scalable Video Coding (SVC) Extension

- Spatial inter-layer prediction (SVC)
- SNR scalability or Quality scalability
  - Base layer followed by enhancement layers
  - Coarse Grain Scalability (CGS): each enhancement layer corresponds to a quality layer
  - Medium Grain Scalability (MGS): allows several quality layers within an enhancement layer through progressive layered coding of the DCT coefficients
Agenda

Other Developments

- Scalable video coding (SVC)
- Multi-view video coding (MVC)
- Distributed video coding
- Texture-based video coding
Multi-View Video Coding (MVC)

Need for Compression:
Raw multi-view video - highly data intensive
Trouble in storage, communication and processing

Removal of redundancies:
Inter-frame redundancies – Motion compensation
Inter-view redundancies – Disparity compensation
Intra-frame redundancies – Transform coding
Agenda

Other Developments

• Scalable video coding (SVC)
• Multi-view video coding (MVC)
• Distributed video coding
• Texture-based video coding
What is distributed video coding?
Distributed Video Coding - Motivation

Wireless sensor networks
• Many encoders to one decoder application
• Low power wireless sensor
• Low complexity video encoder

Conventional video compression
• MPEG or the ITU-T H.26x
• High complexity video encoder and low complexity video decoder
• Motion estimation at the encoder

©2007-2015 by Lina J. Karam
Distributed Video Coding - Objectives

Intraframe encoding and interframe decoding
- Remove motion estimation from encoder to decoder
- Achieve interframe compression rate-distortion performance

Distributed source coding
- Compress two consecutive frames separately by using two different encoders.
- Decode the frames jointly at the decoder.
Agenda

Other Developments

• Scalable video coding (SVC)
• Multi-view video coding (MVC)
• Distributed video coding
• Texture-based video coding
Texture-based Video Coding


Input Image/Frame → Non-Texture → Standard Image/video encoding
Other Sources of Information


Other Sources of Information


Other Sources of Information

• International Telecommunication Union - [http://www.itu.int/](http://www.itu.int/)
  (download the H.264/AVC and H.265/HEVC standards – search for H.264 and for H.265)

• MPEG homepage - [http://www.chiariglione.org/mpeg/](http://www.chiariglione.org/mpeg/)


• H.265 Reference software [https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/)

• Transform-Domain Distributed Video Coding with Rate-Distortion Based Adaptive Quantization, Wei-Jung Chien and Lina J. Karam, IET Image Processing Journal, Special Issue on Distributed Video Coding, pages 340-354, vol. 3, no. 6, December 2009.