HEVC OVERVIEW

March 2013
Agenda

• Overview of video coding standards
• The HEVC standard
  • History, schedule, etc
  • Technical details
  • Performance and complexity analysis
• Extensions of HEVC
  • Scalable extensions of HEVC (SHVC)
• Demo: Power Aware HEVC
Video Coding Standards Overview

First Standardized

ITU-T

H.261 ISDN Videoconferencing

1996

H.262 MPEG-2 DVD, Digital TV, HDTV

H.263(+) Desktop/mobile video telephony

ISO/IEC

JPEG

MPEG-1 CD-ROM

MPEG-2
- Dominates Digital TV and DVD markets

MPEG-4

JPEG2000

H.264/AVC
- Digital TV
  - Slow uptake by cable networks in US, wider adoption in Europe
- Satellite TV service providers
- Internet Video
- Blu-Ray Disc

2003

H.264 MPEG-4 AVC

2013 (exp)

‘H.265’ HEVC

H.265 HEVC

First

Standardized

1996

2003

2013 (exp)
Agenda

• Overview of video coding standards
• The HEVC standard
  • History, schedule, etc
  • Technical details
  • Performance and complexity analysis
• Extensions of HEVC
  • Scalable extensions of HEVC (SHVC)
• Demo: Power Aware HEVC
What is HEVC?

• HEVC = High Efficiency Video Coding
• Joint project between ISO/IEC/MPEG and ITU-T/VCEG
  • ISO/IEC: MPEG-H Part 2 (23008-2)
  • ITU-T: H.265
• The JCT-VC committee
  • Joint Collaborative Team on Video Coding
  • Co-chairs: Dr. Gary Sullivan (Microsoft, USA) and Dr. Jens-Reiner Ohm (RWTH Aachen, Germany)
• Target: roughly half the bit-rate at the same subjective quality compared to H.264/AVC
• Requirements:
  • Progressive required for all profiles and levels
    • Interlaced support using field SEI message
  • Video resolution: sub QVGA to 8Kx4K, with more focus on higher resolution video content (1080p and up)
  • Color space and chroma sampling: YUV420, YUV422, YUV444, RGB444
  • Bit-depth: 8-14 bits
HEVC history

Jan 2010 • Call for proposals

Apr 2010 • Evaluation of responses

Oct 2010 • First test model (HM) and first working draft

Feb 2012 • Committee draft

Jan 2013 • Final draft international standard

Pre-HEVC technology dev.
• H.264/AVC finalized in 2004
• ITU-T VCEG established KTA in early 2005
• KTA technology development 2005 to 2009
• MPEG issued “Call for Evidence” in April 2009
• In Jan 2010, Joint CfP was issued, and JCT-VC was established
• In Apr 2010, 27 CfP responses were provided, TMuC was formed

Development of HEVC scalable extensions and multiview extensions currently on-going
MPEG-H

• High Efficiency Coding and Media Delivery in Heterogeneous Environments – a new suite of standards providing technical solutions for emerging challenges in multimedia industries

  ▪ Part 1: System, MPEG Media Transport (MMT)

    Integrated services with multiple components in a hybrid delivery environment, providing support for seamless and efficient use of heterogeneous network environments, including broadcast, multicast, storage media and mobile networks

  ▪ Part 2: Video, High Efficiency Video Coding (HEVC)

    Highly immersive visual experiences, with ultra high definition displays that give no perceptible pixel structure even if viewed from such a short distance that they subtend a large viewing angle (up to 55 degrees horizontally for 4Kx2K resolution displays, up to 100 degrees for 8Kx4K).

  ▪ Part 3: Audio, 3D-Audio

    Highly immersive audio experiences in which the decoding device renders a 3D audio scene. This may be using 10.2 or 22.2 channel configurations or much more limited speaker configurations or headphones, such as found in a personal tablet or smartphone.
Agenda

• Overview of video coding standards
• The HEVC standard
  • History, schedule, etc
  • Technical details
  • Performance and complexity analysis
• Extensions of HEVC
  • Scalable extensions of HEVC (SHVC)
• Demo: Power Aware HEVC
HEVC Encoder

- **Split into CTU**

- **Quad-tree split (CU, PU, TU)**

- **Core transform: 4x4 to 32x32**
  - **4x4 DST for intra**
  - **Transform skip mode**

- **CABAC only**
  - **Sign hiding**
  - **Parallelization (tiles, WPP, slices)**

- **Deblocking filter**
  - **Sample Adaptive Offsets (SAO)**

- **33 directional modes, DC and planar**
- **Adaptive ref sample smoothing**
- **Mode dependent scanning**
- **Boundary smoothing**
- **And more (mode coding, sample substitution, etc)**

- **DCT-IF for fractional pixel interpolation**
- **Asymmetric Motion Partition (AMP)**
- **Merge**
- **Advanced Motion Vector Pred (AMVP)**

- **Input Video**

- **Transform & Quantization**

- **Quantized coefficients**

- **Packing and Entropy coding**

- **Coded bitstream**

- **Mode Decision and Other Encoder Control Logic**

- **Prediction block**

- **Coding mode, prediction mode, motion**

- **Loop Filters (Control and analysis)**

- **Decoded Picture Buffer**
HEVC Decoder
Block Structure in HEVC

- **Coding Tree Units (CTU)**
  - Corresponds to macroblocks in earlier coding standards (H.264, MPEG2, etc)
  - Luma and chroma Coding Tree Blocks (CTB)
  - Quadtree structure to split into Coding Units (CUs)
  - 16x16, 32x32, or 64x64, signaled in SPS

- **Coding Units (CU)**
  - Luma and chroma Coding Blocks (CB)
  - Rooted in CTU
  - Intra or inter coding mode
  - Split into Prediction Units (PUs) and Transform Units (TUs)

- **Prediction Units (PU)**
  - Luma and chroma Prediction Blocks (PB)
  - Rooted in CU
  - Partition and motion info

- **Transform Units (TU)**
  - Rooted in CU
  - 4x4, 8x8, 16x16, 32x32 DCT, and 4x4 DST
Intra Prediction (1)

- 35 intra modes: 33 directional modes + DC + planar
- For chroma, 5 intra modes: DC, planar, vertical, horizontal, and luma derived
Intra Prediction (2)

- Adaptive reference sample filtering
  - 3-tap filter: $[1 \ 2 \ 1]/4$
  - Not performed for 4x4 blocks
  - For larger than 4x4 blocks, adaptively performed for a subset of modes
    - Modes except vertical/near-vertical, horizontal/near-horizontal, and DC

- Mode dependent adaptive scanning
  - 4x4 and 8x8 intra blocks only
  - All other blocks use only diagonal upright scan (left-most scan pattern)
Intra Prediction (3)

- Boundary smoothing:
  - Applied to DC, vertical, and horizontal modes, luma only
  - Reduces boundary discontinuity

- For DC mode, 1st column and row of samples in predicted block are filtered
  \[
  \text{predSamples}[0, 0] = (p[-1, 0] + 2*DCVal + p[0, -1] + 2) >> 2 \\
  \text{predSamples}[x, 0] = (p[x, -1] + 3*DCVal + 2) >> 2, \text{ with } x = 1..nS-1 \\
  \text{predSamples}[0, y] = (p[-1, y] + 3*DCVal + 2) >> 2, \text{ with } y = 1..nS-1 \\
  \text{predSamples}[x, y] = DCVal, \text{ with } x, y = 1..nS-1
  \]

- For Hor/Ver mode, first column/row of pixels in predicted block are filtered
  
  **Vertical:**
  \[
  \text{predSamples}[0, y] = \text{clip}(p[0, -1] + (d[y] >> 1)), d[y] = p[-1, y] - p[-1, -1]
  \]

  **Horizontal:**
  \[
  \text{predSamples}[x, 0] = \text{clip}(\text{p}[-1, 0] + (d[x] >> 1)), d[x] = p[x, -1] - p[-1, -1]
  \]
Inter Prediction (1)

- Fractional sample interpolation
  - ¼ pixel precision for luma
- DCT based interpolation filters:
  - 8/-7- tap for luma
  - 4-tap for chroma
- Supports 16-bit implementation with non-normative shift
- High precision interpolation and bi-prediction

- DCT-IF design
  - Forward DCT, followed by inverse DCT

\[
p(\alpha) = \sum_{k=0}^{2M-1} C_k W_k(\alpha) = \frac{1}{M} \sum_{k=1}^{2M-1} \left( \sum_{l=-M+1}^{M} p(l) d_k (l) \right) W_k(\alpha) = \frac{1}{M} \sum_{l=-M+1}^{M} p(l) \left( \sum_{k=1}^{2M-1} W_k(\alpha) d_k (l) \right)
\]

\[
d_k (l) = \cos \left( \frac{(2l-1+2M)}{4M} k \pi \right)
\]

\[
W_k(x) = \cos \left( \frac{(2x-1+2M)}{4M} k \pi \right)
\]
Inter Prediction (2)

- Asymmetric Motion Partition (AMP) for Inter PU

- Merge
  - Derive motion (MV and ref pic) from spatial and temporal neighbors
  - Which spatial/temporal neighbor is identified by merge_idx
  - Number of merge candidates (≤ 5) signaled in slice header
  - Skip mode = merge mode + no residual

- Advanced Motion Vector Prediction (AMVP)
  - Use spatial/temporal PUs to predict current MV
Transforms

- Core transforms: DCT based
  - 4x4, 8x8, 16x16, and 32x32
- Square transforms only
- Support partial factorization
- Near-orthogonal
- Nested transforms

Alternative 4x4 DST
- 4x4 intra blocks, luma only
- Transform skipping mode
  - By-pass the transform stage
  - Most effective on “screen content”
- 4x4 TBs only
HEVC Decoder

- CABAC only
- Sign hiding
- High level parallelization

- Deblocking filter
- Sample Adaptive Offsets (SAO)
Loop Filters: Deblocking

- Simpler deblocking filter in HEVC (vs H.264)

<table>
<thead>
<tr>
<th></th>
<th>H.264</th>
<th>HEVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid</td>
<td>4x4</td>
<td>8x8</td>
</tr>
<tr>
<td>Boundary strength</td>
<td>0-4</td>
<td>0-2</td>
</tr>
<tr>
<td>Chroma</td>
<td>Filtered if Bs ≠ 0</td>
<td>Filtered only if Bs = 2</td>
</tr>
</tbody>
</table>

- Deblocking filter boundary strength is set according to
  - Block coding mode
  - Existence of non zero coefficients
  - Motion vector difference
  - Reference picture difference
Loop Filters: SAO

- **SAO: Sample Adaptive Offsets**
  - New loop filter in HEVC
  - Non-linear filter

- For each CTB, signal SAO type and parameters

- Encoder decides SAO type and estimates SAO parameters (rate-distortion opt.)

<table>
<thead>
<tr>
<th>EdgedIdx</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>c is same as 2 neighbors</td>
</tr>
<tr>
<td>1</td>
<td>c &lt; 2 neighbors</td>
</tr>
<tr>
<td>2</td>
<td>c &lt; 1 neighbor &amp;&amp; c == 1 neighbor</td>
</tr>
<tr>
<td>3</td>
<td>c &gt; 1 neighbor &amp;&amp; c == 1 neighbor</td>
</tr>
<tr>
<td>4</td>
<td>c &gt; 2 neighbors</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Classification</th>
<th>Parameters</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No SAO</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td>Band Offset (BO)</td>
<td>BO offsets for 4 consecutives bands</td>
<td>Reduce “banding artifacts”</td>
</tr>
<tr>
<td>2</td>
<td>Edge Offset (EO)</td>
<td>EO class index, EO offsets</td>
<td>Enhance edge sharpness</td>
</tr>
</tbody>
</table>

- **SAO: Sample Adaptive Offsets**
  - For each CTB, signal SAO type and parameters
  - Encoder decides SAO type and estimates SAO parameters (rate-distortion opt.)
Entropy Coding

- One entropy coder, CABAC
  - Reuse H.264 CABAC core algorithm
  - More friendly to software and hardware implementations
  - Easier to parallelize, reduced HW area, increased throughput

- Context modeling:
  - Reduced # of contexts
  - Increased use of by-pass bins
  - Reduced data dependency

- Coefficient coding
  - Adaptive coefficient scanning for intra 4x4 and 8x8
    - Diagonal upright, horizontal, vertical
  - Processed in 4x4 blocks for all TU sizes
  - Sign data hiding:
    - Sign of first non-zero coefficient conditionally hidden in the parity of the sum of the non-zero coefficient magnitudes
    - Conditions: 2 or more non-zero coefficients, and “distance” between first and last coefficient > 3
High Level Parallelism

- Independently decodable packets
- Sequence of CTUs in raster scan
- Error resilience
- Parallelization

Slices

Tiles

- Independently decodable (re-entry)
- Rectangular region of CTUs
- Parallelization (esp. encoder)
- 1 slice = more tiles, or 1 tile = more slices

WPP: Wavefront Parallel Proc

- Rows of CTUs
- Decoding of each row can be parallelized
- Shaded CTU can start when gray CTUs in row above are finished
- Main profile does not allow tiles + WPP combination
Profiles, Levels and Tiers

- Historically, profile defines collection of coding tools, whereas Level constrains decoder processing load and memory requirements.

- The first version of HEVC defined 3 profiles:
  - **Main Profile:** 8-bit video in YUV4:2:0 format
  - **Main 10 Profile:** same as Main, up to 10-bit
  - **Main Still Picture Profile:** same as Main, one picture only

- Levels and Tiers:
  - Levels: max sample rate, max picture size, max bit rate, DPB and CPB size, etc.
  - Tiers: “main tier” and “high tier” within one level.

<table>
<thead>
<tr>
<th>Level</th>
<th>Max luma sample rate</th>
<th>Max luma samples/sec</th>
<th>MaxBR (kbps)</th>
<th>Min Compression Ratio</th>
<th>MinCr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>552 960</td>
<td>128</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3 686 400</td>
<td>1 500</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>7 372 800</td>
<td>3 000</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16 588 800</td>
<td>6 000</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>33 177 600</td>
<td>10 000</td>
<td>-</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>66 846 720</td>
<td>12 000</td>
<td>30 000</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>133 693 440</td>
<td>20 000</td>
<td>50 000</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>267 386 880</td>
<td>25 000</td>
<td>100 000</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>534 773 760</td>
<td>40 000</td>
<td>160 000</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>1 069 547 520</td>
<td>60 000</td>
<td>240 000</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1 069 547 520</td>
<td>60 000</td>
<td>240 000</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>2 139 095 040</td>
<td>120 000</td>
<td>480 000</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>4 278 190 080</td>
<td>240 000</td>
<td>800 000</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Format nickname</th>
<th>Luma width</th>
<th>Luma height</th>
<th>Luma picture size</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGA</td>
<td>640</td>
<td>480</td>
<td>327 680</td>
</tr>
<tr>
<td>720p HD</td>
<td>1280</td>
<td>720</td>
<td>983 040</td>
</tr>
<tr>
<td>1080 HD</td>
<td>1920</td>
<td>1080</td>
<td>2 088 960</td>
</tr>
<tr>
<td>4Kx2K</td>
<td>4096</td>
<td>2048</td>
<td>8 388 608</td>
</tr>
<tr>
<td>8192x4096</td>
<td>8192</td>
<td>4096</td>
<td>33 554 432</td>
</tr>
</tbody>
</table>
Agenda

• Overview of video coding standards
• The HEVC standard
  • History, schedule, etc
  • Technical details
  • Performance and complexity analysis
• Extensions of HEVC
  • Scalable extensions of HEVC (SHVC)
• Demo: Power Aware HEVC
HEVC Compression Performance

- Compression performance: PSNR based
  - Bit rate savings
  - Tool-by-tool performance
  - Example operating R-D curves

- Compression performance: subjective quality based
  - Bit rate savings
  - Example curves

HEVC Performance: PSNR based

<table>
<thead>
<tr>
<th>Applications</th>
<th>Bit rate savings (HEVC vs. previous standards)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>H.264/AVC HP</strong></td>
</tr>
<tr>
<td>Entertainment</td>
<td>35.4%</td>
</tr>
<tr>
<td>Interactive</td>
<td>40.3%</td>
</tr>
</tbody>
</table>

**Notes**
- Rate savings calculated using Bjontegaard Delta rate (BD-rate)
- $\text{PSNR}_{YUV} = (6 \cdot \text{PSNR}_Y + \text{PSNR}_U + \text{PSNR}_V) / 8$
- Official reference software + improved encoder control used to generate the results

<table>
<thead>
<tr>
<th></th>
<th>Entertainment</th>
<th>Interactive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test seq.</td>
<td>5 HD, 4 WVGA seq.</td>
<td>6 video conf. seq.</td>
</tr>
<tr>
<td>Prediction structure</td>
<td>Hier-B with GOP=8 (HEVC and H.264), iBBB (otherwise)</td>
<td>Hier-B (HEVC), Hier-P (H.264, H.263), IPPP (otherwise)</td>
</tr>
<tr>
<td>Intra period</td>
<td>Approx. 1 sec</td>
<td>First I only</td>
</tr>
<tr>
<td>Num. ref</td>
<td>4 (HEVC, H.264, H.263)</td>
<td>4 (HEVC, H.264, H.263)</td>
</tr>
<tr>
<td>QP</td>
<td>Qstep inc 12% P/B vs. I, and inc 12% per level</td>
<td></td>
</tr>
</tbody>
</table>
Performance of Individual Coding Tools

Entertainment Applications

<table>
<thead>
<tr>
<th></th>
<th>Up to 16x16 CTB</th>
<th>Up to 8x8 Transform</th>
<th>RQT depth =1</th>
<th>TMVP off</th>
<th>SAO off</th>
<th>AMP off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>28.2%</td>
<td>12.2%</td>
<td>0.8%</td>
<td>2.6%</td>
<td>2.4%</td>
<td>0.6%</td>
</tr>
<tr>
<td>Class B</td>
<td>18.4%</td>
<td>9.3%</td>
<td>1.1%</td>
<td>2.2%</td>
<td>2.4%</td>
<td>0.7%</td>
</tr>
<tr>
<td>Class C</td>
<td>8.5%</td>
<td>4.2%</td>
<td>1.1%</td>
<td>2.4%</td>
<td>1.7%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Class D</td>
<td>4.2%</td>
<td>2.4%</td>
<td>1.1%</td>
<td>2.7%</td>
<td>0.5%</td>
<td>0.9%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>11.0%</strong></td>
<td><strong>5.4%</strong></td>
<td><strong>1.0%</strong></td>
<td><strong>2.5%</strong></td>
<td><strong>1.6%</strong></td>
<td><strong>0.9%</strong></td>
</tr>
</tbody>
</table>

Interactive Applications

<table>
<thead>
<tr>
<th></th>
<th>Up to 16x16 CTB</th>
<th>Up to 8x8 Transform</th>
<th>RQT depth =1</th>
<th>TMVP off</th>
<th>SAO off</th>
<th>AMP off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class B</td>
<td>19.2%</td>
<td>9.7%</td>
<td>1.4%</td>
<td>2.5%</td>
<td>2.6%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Class C</td>
<td>10.3%</td>
<td>5.5%</td>
<td>1.5%</td>
<td>2.8%</td>
<td>2.9%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Class D</td>
<td>5.7%</td>
<td>3.1%</td>
<td>1.4%</td>
<td>2.4%</td>
<td>1.3%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Class E</td>
<td>39.2%</td>
<td>10.6%</td>
<td>0.8%</td>
<td>2.4%</td>
<td>3.3%</td>
<td>1.7%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>17.4%</strong></td>
<td><strong>7.2%</strong></td>
<td><strong>1.3%</strong></td>
<td><strong>2.5%</strong></td>
<td><strong>2.5%</strong></td>
<td><strong>1.2%</strong></td>
</tr>
</tbody>
</table>
Compression Performance (Entertainment)

Park Scene, 1920x1080, 24Hz

![Graph showing compression performance](graph.png)
Compression Performance (Entertainment)
Compression Performance (Interactive)
Compression Performance (Interactive)
### HEVC Performance: Subjective Quality based

**DSIS (Double Stimulus Impairment Scale) according to ITU-R BT.500**

<table>
<thead>
<tr>
<th>Sequences</th>
<th>1080p</th>
<th>Bit rate savings (HEVC vs. H.264/AVC HP)</th>
<th>WVGA (720x480)</th>
<th>Bit rate savings (HEVC vs. H.264/AVC HP)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Original 1 sec</td>
<td>Coded 1 sec</td>
<td>Original 1 sec</td>
</tr>
<tr>
<td>1080p</td>
<td></td>
<td>BQTerrace</td>
<td>BasketballDrive</td>
<td>Kimono</td>
</tr>
<tr>
<td></td>
<td></td>
<td>63.1%</td>
<td>66.6%</td>
<td>55.2%</td>
</tr>
<tr>
<td>WVGA (720x480)</td>
<td></td>
<td>BQMall</td>
<td>BasketballDrill</td>
<td>PartyScene</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41.6%</td>
<td>44.9%</td>
<td>29.8%</td>
</tr>
</tbody>
</table>

**Notes:**
- Entertainment applications
- Random Access config
- HEVC settings: HM5.0, QP = \{31, 34, 37, 40\}
- H.264/AVC: JM18.2* (imp encoder control), QP = \{27, 30, 33, 34\}
Compression Performance: Subjective Quality based
Compression Performance: Subjective Quality based

![Graph showing bit rate saving relative to H.264/MPEG-4 AVC HP for Kimono1, 1920x1080, 24Hz. The graph is labeled HEVC MP. Key points: 380 kbps at subjective quality of 0.0 and 1270 kbps at subjective quality of 8.0.](image-url)
Compression Performance: Subjective Quality based
Compression Performance: Subjective Quality based

Cactus, 1920x1080, 50Hz

bit rate saving relative to H.264/MPEG-4 AVC HP

subjective quality (MOS)

1100 kbps

3170 kbps

HEVC MP
HEVC Complexity Analysis (1)

- Software-based HEVC decoder capabilities (published by NTT Docomo)
  - Single-threaded: 1080p@30 on ARMv7 (1.3GHz), 1080p@60 decoding on i5 (2.53GHz)
  - Multi-threaded: 4Kx2K@60 on i7 (2.7GHz), 12Mbps, decoding speed up to 100fps
- Other independent software-based HEVC real-time decoder implementations published by Samsung and Qualcomm during HEVC development
- Decoder complexity not substantially higher
  - More complex modules: MC, Transform, Intra Pred, SAO
  - Simpler modules: CABAC and deblocking

HEVC Complexity Analysis (2)

Performance and complexity comparison between HEVC AVC

<table>
<thead>
<tr>
<th></th>
<th>AI</th>
<th>RA</th>
<th>LD-B</th>
<th>LD-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate red. @ same PSNR</td>
<td>23%</td>
<td>35%</td>
<td>40%</td>
<td>35%</td>
</tr>
<tr>
<td>Encoder complexity</td>
<td>3.2x</td>
<td>1.2x</td>
<td>1.5x</td>
<td>1.3x</td>
</tr>
<tr>
<td>Decoders complexity</td>
<td>2.0x</td>
<td>1.6x</td>
<td>1.5x</td>
<td>1.4x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>HEVC</th>
<th>H.264/AVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference Software</td>
<td>HM 6.0</td>
<td>JM18.0</td>
</tr>
<tr>
<td>QP</td>
<td>$Q_{pHM} = {22, 27, 32, 37}$</td>
<td>$Q_{pM}$ chosen for each test point (per setting, seq, $Q_{pHM}$) to match HEVC PSNR</td>
</tr>
<tr>
<td>Other parameters</td>
<td>Prediction structure, intra period, num ref, ME search range, etc, all kept the same</td>
<td></td>
</tr>
</tbody>
</table>


“The future trend is that the processing performance will continue to develop faster than transmission and storage technologies. This trend will further promote HEVC because of its capability to almost halve the bit rate. Due to this reason, we forecast rapid proliferation of HEVC in the next-generation video products and services.”
Agenda

• Overview of video coding standards
• The HEVC standard
  • History, schedule, etc
  • Technical details
  • Performance and complexity analysis
• Extensions of HEVC
  • Scalable extensions of HEVC (SHVC)
• Demo: Power Aware HEVC
## Scalable Codec Overview

<table>
<thead>
<tr>
<th>Feature</th>
<th>Standard Support</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SVC  MVC  HEVC extensions</td>
<td>Base layer  Enhance layer</td>
</tr>
<tr>
<td>Spatial</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Temporal</td>
<td>X</td>
<td>X (in HEVC)</td>
</tr>
<tr>
<td>SNR</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Standard</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>View</td>
<td>X</td>
<td>X (in 3DV)</td>
</tr>
<tr>
<td>Bit-depth</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>Chroma format</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>Aspect ratio</td>
<td>X*</td>
<td></td>
</tr>
</tbody>
</table>

X*: not in the 1st phase of scalable extensions

- Historically, older scalable video codecs had not been adopted by markets
  - This includes scalable extensions of MPEG1, MPEG2 and MPEG4 part 2
- The most recent SVC (H.264 Annex G), finalized in 2007, has limited market adoption:
  - Vidyo software products for video conferencing
  - Broadcom BCM7422 chipset
  - IMTC SVC AG: Ericsson, LifeSize, Magor, Polycom, Radvision, and Vidyo
- Another scalable codec MVC (H.264 Annex H), has enjoyed wider and faster market adoption: 3D Blu-Ray Disc
Use Cases for Scalable Coding

Video calling and multi-point video conferencing

Live Streaming

Cloud gaming

Increased heterogeneity in networks and in clients, increased wireless penetration → scalable coding
SHVC Timeline

- First version of HEVC (single layer coding) finalized in Jan 2013

- HEVC extensions currently underway
  - JCT-VC: Scalable extensions of HEVC (SHVC)
  - JCT-3V: Multi-view extensions of HEVC
    - MV-HEVC and 3D-HEVC

- At the 11\textsuperscript{th} JCT-VC meeting (Oct 2012), 20 SHVC CfP responses were submitted
Our SHVC solution

• The “ref_idx” approach
  • Similar to multi-view extensions such as MVC (H.264 Annex H) and MV-HEVC
  • Inter layer prediction achieved by inserting inter layer picture into EL DPB
  • Maximally re-use existing HEVC encoder/decoder implementations

• Design benefits:
  • Reduced implementation cost
  • Coherent architecture to support different scalabilities
  • Easy to combine spatial/quality with 3D in the future
  • Agnostic to underlying single layer codecs
    • HEVC→HEVC, AVC→HEVC, MPEG2 →HEVC, etc
## SHVC Current Status

- In Jan 2013, first test model and first working draft were established
- 2 parallel tracks in the test model

<table>
<thead>
<tr>
<th>Features</th>
<th>RefIdx solution</th>
<th>IntraBL solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-layer prediction signaling</td>
<td>Use existing syntax ref_idx</td>
<td>Create new syntax IntraBL flag</td>
</tr>
<tr>
<td>History</td>
<td>MVC-like</td>
<td>SVC-like</td>
</tr>
<tr>
<td>Change EL codec block level logics?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Easy to combine with view scalability?</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison w/ simulcast</th>
<th>Luma BD rate</th>
<th>Encoding time</th>
<th>Decoding time</th>
</tr>
</thead>
<tbody>
<tr>
<td>* BL &amp; EL coded independently</td>
<td>-18.7%</td>
<td>103.5%</td>
<td>94.3%</td>
</tr>
<tr>
<td></td>
<td>-19.1%</td>
<td>105.4%</td>
<td>93.3%</td>
</tr>
</tbody>
</table>
THANK YOU!!

QUESTIONS?