Recent Advances of High Efficiency Video Coding

彭文孝 (W.-H. Peng)
國立交通大學
(National Chiao Tung Univ.)
資訊工程學系
(Dept. of Computer Science)
Email: wpeng@cs.nctu.edu.tw

Part 1: HEVC 標準現況介紹
JCT-VC Meetings

- 250+ Participants
- Documents: 100+ (1st & 2nd) → 300+ (3rd) → 400+ (4th) → 500+ (5th) → 700+ (6th) → 1000+ (7th) → 700+ (8th) → 500+ (9th) (3400+ in total)
- 2 Parallel Sessions, 8:00am-11:00pm, 11 Days
- 14 Ad-hoc Groups + 1 Core Experiments
- Text Specification Draft 7 + HM7.0
- Draft International Standard (2012/07)
- Final Draft International Standard (2013/01)
- International Standard (2013/04)

Common Test Conditions

- 2 tool sets (HE10, Main) x 3 prediction structures

Intra Only  Random Access  Low Delay
### HEVC Coding Tools

<table>
<thead>
<tr>
<th>Tools</th>
<th>High Efficiency (10-bit)</th>
<th>Main (8-bit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU Size</td>
<td>8x8~64x64</td>
<td>16x16~64x64</td>
</tr>
<tr>
<td>PU Partition</td>
<td>Symmetric/Asymmetric</td>
<td>Symmetric</td>
</tr>
<tr>
<td>TU Partition</td>
<td>RQT/NSQT</td>
<td>RQT</td>
</tr>
<tr>
<td>MV Prediction</td>
<td>AMVP, MRG</td>
<td>AMVP, MRG</td>
</tr>
<tr>
<td>Intra Prediction</td>
<td>DC, Planar, 33 Directions</td>
<td>DC, Planar, 33 Directions</td>
</tr>
<tr>
<td>Transform</td>
<td>DCT 4x4~32x32, DST 4x4 (Intra)</td>
<td>DCT 4x4~32x32, DST 4x4 (Intra)</td>
</tr>
<tr>
<td>Interpolation Filter</td>
<td>DCT-IF</td>
<td>DCT-IF</td>
</tr>
<tr>
<td>In-loop Filter</td>
<td>De-blocking, SAO, ALF</td>
<td>De-blocking, SAO</td>
</tr>
<tr>
<td>Entropy Coding</td>
<td>CABAC (w/ Tiles)</td>
<td>CABAC (w/ Tiles)</td>
</tr>
</tbody>
</table>

---

### HEVC (HE) vs. AVC (High)

**HM5.0 vs. JM18.2**

BD-rate Saving (%)

- **INTRA**
- **RANDOM ACCESS**
- **LOW DELAY**

Subjective Assessment

**ParkScene**

- **HM**
- **JM**

Over -50%

<table>
<thead>
<tr>
<th>Bit Rate [Kbps]</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>500</td>
<td>1.0</td>
</tr>
<tr>
<td>1000</td>
<td>2.0</td>
</tr>
<tr>
<td>1500</td>
<td>3.0</td>
</tr>
<tr>
<td>2000</td>
<td>4.0</td>
</tr>
<tr>
<td>2500</td>
<td>5.0</td>
</tr>
<tr>
<td>3000</td>
<td>6.0</td>
</tr>
<tr>
<td>3500</td>
<td>7.0</td>
</tr>
<tr>
<td>4000</td>
<td>8.0</td>
</tr>
</tbody>
</table>

**Class B (HD):** -67%

**Class C (SD):** -49%

---


---

**HM Tool Features**

- Asymmetric Motion Partitioning
- Merged Skip / Motion Merging
- Advanced MV Prediction
- DCT-based Interpolation Filter
- More Directions
- Pre-/Post-filtering
- Luma Predict Chroma
- Residual Quad-tree
- Non-Rect. Quad-tree
- Large Transform
- Adaptive Coeff. Scanning
- Deblocking Filter
- Sample Adaptive Offset
- Adaptive Loop Filter
- CABAC
- Entropy Slice
- Tiles
- Wavefront

---

**Coding Unit (CU)**

- Prediction
- Transform
- • Asymmetric Motion Partitioning
- • Merged Skip / Motion Merging
- • Advanced MV Prediction
- • DCT-based Interpolation Filter

**Transform Unit (TU)**

- • More Directions
- • Pre-/Post-filtering
- • Luma Predict Chroma

**In-loop filter**

**Entropy Coding**

- Bitstream

---

**Current Frame**

**Frame**

**Intra Prediction**

**Inter Prediction**

**IQ IDCT**

---

8
Basic Units

- Quadtree-based units for **coding, prediction and transform** (CU, PU, TU)
- Separate signaling of CU, PU, TU partitions

![Diagram of Basic Units]

Intra Prediction

- Up to 33 directions & DC/Planar modes
- Pre-filtering on reference pixels (adaptive ON/OFF according to PU size and prediction direction)
- Boundary smoothing for DC/Ver./Hor. modes
- Extra modes for Chroma
  - DM (Direct Mode)
  - LM (Linear Mode)

![Diagram of Intra Prediction]
LM Chroma Mode

- Predict Chroma from Luma
  \[ \text{Pred}_C(x,y) = \alpha \cdot \text{Rec}'_L(x,y) + \beta \]
- Estimate \( \alpha, \beta \) by the Least-Squares method

Lossless Coding

- Bypass Transform, Quantization, In-loop Filtering
- Predict pixels using nearest reconstructed pixels (Intra)
Merged Skip & Motion Merging

- B0, B1, B2
- T0, T1, A0, A1

Optional

5 Candidates at most

If LCU boundary, exclude A0 & T0

T0’s & T1’s RefIdx

A1, if 1st PU;
0, otherwise.

Parallel Motion Merging

- CU
- Parallel Group
- Possible Merge Dir

PUs/CUs within a group run in parallel
Advanced MV Prediction (AMVP)

1. Find First Available
   - A1
   - A0

2. Find First Available
   - B2
   - B1
   - B0

3. Optional
   - T1

2 Candidates at most

If LCU boundary, exclude A0 & T0

Spatial AMVP Derivation

1. MV0_L0
2. MV1_L1
3. MV1_L0
4. MV0_L1

L0, Ref1, L0, Ref0, L1, Ref1, Current, L1, Ref0
Combined Reference Frame List

- Combine List 0 and List 1 into a single list to remove uni-prediction signaling overhead.

![Diagram showing reference frames and mapping]

Weighted Prediction

\[ P_0(x, y) \]
\[ W_0 \]
Ref 0

\[ P_1(x, y) \]
\[ W_1 \]
Ref 1

\[ V_0 \]
Current

\[ V_1 \]

Uni: \[ \frac{w_0 P_0(x, y) + 2^{\text{shift}-1}}{2^{\text{shift}}} + o_0 \]

Bi: \[ \frac{w_0 P_0(x, y) + w_1 P_1(x, y) + (o_0 + o_1 + 1)^{\text{shift}}}{2^{\text{shift}+1}} \]
Motion Data Storage

- Compressed MV, RefIdx, PredMode
- Temporal MVp Derivation

Motion Field $\rightarrow$ Duplication $\rightarrow$ Storage

### DCT-based Interpolation Filter

- Interpolation by Inverse DCT (DCT-IF)

$$p(\alpha) = \frac{C_0}{2} + \sum_{k=1}^{2M-1} C_k \cos \left( \frac{\pi \left( 2\alpha - 1 + 2M \right) k}{4M} \right)$$

$C_0$ and $C_k$ are DCT coefficients for sub-pel position $\alpha$. The spatial domain samples are $\{P_{-3}, P_{-2}, \ldots, P_4\}$, and the DCT domain coefficients are $\{C_{-3}, C_{-2}, \ldots, C_4\}$.
Significance Map Scanning

<table>
<thead>
<tr>
<th>TU Size</th>
<th>Prediction Type</th>
<th>Scanning Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>4x4, 8x8</td>
<td>Intra (mode-dependent)</td>
<td>Vertical, Horizontal</td>
</tr>
<tr>
<td>All</td>
<td>Intra (mode-dependent), Inter</td>
<td>4x4 Sub-diagonal</td>
</tr>
</tbody>
</table>

- **Vertical**
- **Horizontal**
- **4x4 Sub-diagonal**

**Sample Adaptive Offset (SAO)**

- Band Offset (BO): intensity classification + DC comp.
- Edge Offset (EO): pixel classification + DC comp.

**LCUs in a Frame**
LCU-based SAO Representation

Repeat Row
Run=2
Merge Up
EO Off
EO
BO
EO

Sign of SAO Edge Offset

Positive Offset

![Positive Offset Diagram]

Negative Offset

![Negative Offset Diagram]
Adaptive Loop Filter (ALF)

- In-loop filtering for image restoration
- 16 filters (at most) for a LCU
- Similar syntax representation to SAO offsets

Choose one from 16 filters

Entropy Slices (ES)

Break slice data into parts

Prediction can cross ES boundary
Wavefront

- Pass CABAC states to next Wave
- [2 Waves in a Slice]
- Flush CABAC States

Tiles

- Prediction can’t cross Tiles
- Parallel
- [4 Tiles in a Slice/Frame]
Core Experiments & Ad-hoc Groups

- CE1: Intra transform mode dependency simplifications
- AHG1: JCT-VC project management
- AHG2: HEVC Draft and Test Model editing
- AHG3: Software development and HM software technical evaluation
- AHG4: High-level parallelism
- AHG5: Entropy coding improvements
- AHG6: In-loop filtering
- AHG7: Memory bandwidth restrictions in motion comp.
- AHG8: Loss robustness
- AHG9: High-level syntax
- AHG10: Hooks for scalable coding
- AHG11: Lossless coding
- AHG12: Support for range extensions
- AHG13: Reference picture buffering and list construction
- AHG14: Study on HEVC conformance requirements
Information


- [http://phenix.int-evry.fr/jct/](http://phenix.int-evry.fr/jct/) (Website)

- [http://mailman.rwth-aachen.de/mailman/listinfo/jct-vc](http://mailman.rwth-aachen.de/mailman/listinfo/jct-vc) (Subscribe)

- [https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/) (SVN, Software Manual JCTVC-F634)

- Text Specification Draft 6 (JCTVC-H1003)

---

Part 2: Text Reading
Part 3: HM Software
**Preparation**

- HM Software
  - [https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/)
  - Software files:
    - .settings
    - config
    - HM.xcodeproj
    - .project
    - .svn
    - compat
    - source
    - build
    - doc
    - .project
    - COPYING
    - README-newconfig.txt

- Testing Sequences (Private)

**Platform**

- VS2005: .build/HM_vc8.sln
- VS2008: .build/HM_vc9.sln
- Gcc: .build/linux/makefile
- xCode: .HM.xcodeproj/project.pbxproj

**Output**

- .bin/vc9/Win32/Release (E.g. VS2008)

---

**Make Binary Executables**

- **Platform**
  - VS2005: .build/HM_vc8.sln
  - VS2008: .build/HM_vc9.sln
  - Gcc: .build/linux/makefile
  - xCode: .HM.xcodeproj/project.pbxproj

- **Output**
  - .bin/vc9/Win32/Release (E.g. VS2008)
Encoding & Decoding

- **Encoding Commands**
  
  ```
  SET EncoderBin=\bin\vc9\win32\release\TAppEncoder.exe
  SET TestCfg=\cfg\encoder_lowdelay_main.cfg
  SET SeqCfg=\cfg\per-sequence\RaceHorses.cfg
  SET InputYuv=\RaceHorses.yuv
  SET QP=37
  "%EncoderBin%" -c "%TestCfg%" -c "%SeqCfg%" -i %InputYuv% -q %QP%
  
  -- Output file: str.bin, rec.yuv
  ```

- **Decoding Commands**
  
  ```
  SET DecoderBin=\bin\vc9\win32\release\TAppDecoder.exe
  SET Bitetream=\str.bin
  SET OutputYuv=\dec.yuv
  "%DecoderBin%" -b "%Bitetream%" -o "%OutputYuv%"
  
  -- Output file: dec.yuv
  ```

Per-frame coding result

<table>
<thead>
<tr>
<th>SUMMARY</th>
<th>Total Frames</th>
<th>Bitrate</th>
<th>Y-PSNR</th>
<th>U-PSNR</th>
<th>U-PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>a</td>
<td>576.6400</td>
<td>29.8087</td>
<td>35.4138</td>
</tr>
<tr>
<td>I Slices</td>
<td>1 i</td>
<td>1096.8000</td>
<td>30.8046</td>
<td>35.7445</td>
<td>35.8751</td>
</tr>
<tr>
<td>P Slices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 p</td>
<td>-1.#IND</td>
<td>-1.#IND</td>
<td>-1.#IND</td>
<td>-1.#IND</td>
</tr>
<tr>
<td>B Slices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 b</td>
<td>96.4000</td>
<td>28.0128</td>
<td>35.0830</td>
<td>35.1582</td>
</tr>
</tbody>
</table>

RUN: 0.000
Bytes written to file: 5045 (605.400 kbps)
Total Time: 3.130 sec.
Per-frame coding result

Summarize report

Summary

<table>
<thead>
<tr>
<th>Total Frames</th>
<th>Bitrate</th>
<th>V-PSNR</th>
<th>U-PSNR</th>
<th>V-PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Slices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.576.6400</td>
<td>29.0987</td>
<td>35.4318</td>
<td>35.5167</td>
</tr>
<tr>
<td>P Slices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1076.0000</td>
<td>30.0046</td>
<td>35.7445</td>
<td>35.0751</td>
</tr>
<tr>
<td>B Slices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>96.4800</td>
<td>28.8128</td>
<td>35.0830</td>
<td>35.1582</td>
</tr>
</tbody>
</table>

SUM: 0.000
Bytes written to file: 5045 (605.400 kbps)
Total Time: 3.130 sec

Encoding time

Decoding time

MD5 Checksum

Collect Coding Result

<table>
<thead>
<tr>
<th>Reference</th>
<th>QP</th>
<th>Slice</th>
<th>kbps</th>
<th>Y psnr</th>
<th>U psnr</th>
<th>V psnr</th>
<th>Enc T[s]</th>
<th>Dec T[s]</th>
<th>Enc Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>RaceHorses</td>
<td>26</td>
<td>2557</td>
<td>90.08</td>
<td>41.24</td>
<td>42.04</td>
<td>2220.38</td>
<td>3.40</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>684.11</td>
<td>35.65</td>
<td>38.41</td>
<td>39.53</td>
<td>2072.83</td>
<td>2.78</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>321.64</td>
<td>25.12</td>
<td>36.07</td>
<td>37.14</td>
<td>1022.53</td>
<td>2.20</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>150.31</td>
<td>28.29</td>
<td>33.95</td>
<td>34.93</td>
<td>1033.93</td>
<td>1.67</td>
<td>0.51</td>
<td></td>
</tr>
</tbody>
</table>

[Anchor]

Rate saving

[Summary]
~Thank You~