Daala: One year later

Timothy B. Terriberry
Original Plan

- *Finish Daala by the end of 2015*

- Schedule driven by some strategic assumptions
  - H.264 fees can be raised in Jan. 2016
  - Assumed HEVC would be licensed cheaper than H.264 to drive adoption of the new format
Original Plan

- **Finish Daala by the end of 2015**
- Schedule driven by some strategic assumptions
  - H.264 fees can be changed in Jan. 2016
  - Assumed HEVC would be licensed cheaper than H.264 to drive adoption of the new format
- This obviously ain’t gonna happen
  - Thanks, HEVC Advance!
- VP9 appears to be good enough
- A lot more people interested in RF video than last year
Last year, we had

- No transform units or prediction units larger than 16×16
- No support for multiple reference frames, B-frames, or alt-ref equivalents
- No replacement for frequency-domain intra
- No loop filters
- No intra mode in our motion search
VDD 2014 → VDD 2015

- Last year, we had
  - No transform units or prediction units larger than 16×16
  - No support for multiple reference frames, B-frames, or alt-ref equivalents
  - No replacement for frequency-domain intra
  - No loop filters
  - No intra mode in our motion search

- This year, we have
  - 32×32 transforms and MC (64×64 in progress for both)
  - Multiple reference frames (B-frames in progress)
  - Simplified H+V intra
  - Bilinear loop filter and deringing filter
Other Major Developments

- “Fixed Lapping”
- Simplified Chroma from Luma
- Better subpel filters
- Preliminary “Screencasting” work
Old Lapping Strategy

- Filter size chosen based on size of smallest block on an edge (to prevent overlap)
- Filter order chosen to mimic a loop filter’s
  - Horizontal edges first
Old Lapping Strategy

- Filter size chosen based on size of smallest block on an edge (to prevent overlap)
- Filter order chosen to mimic a loop filter’s
  - Then vertical
  - Maximal parallelism, minimum buffering
Problem #1: Basis Weirdness
Problem #2: Block size decision

- Have to know neighbors’ block sizes to compute lapping size

- Used a heuristic based on the estimated visibility of ringing to pick block sizes up front
  - Worked “okay” for still images (at least not obviously broken)
  - Was not making good decisions for inter frames

- Wanted to try explicit block size RDO (like other encoders)...
  - But lapping dependency makes this infeasible
“Fixed Lapping”: Remove the Dependency

- Always use 8-point lapping (4 pixels on either side of an edge)
  - Except on 4×4 blocks (details in a few slides)
  - Always use 4-point lapping for chroma (because of subsampling)
New Filter Order

- Filter top/bottom superblock (32x32) edges first
New Filter Order

- Filter left/right superbloack (32×32) edges next
New Filter Order

- Splitting: Filter interior edges
New Filter Order

- Splitting: Filter interior edges

- 4×4 blocks:
  - Exterior edges use 8-point filter (from previous levels)
  - Interior edges use 4-point filter (overlaps 8-point filter)
Results

• Big boost in metrics
  – Almost all from decision
  – Used fixed lapping decision with old lapping scheme and got almost all of the gains

• Smaller lapping means less ringing but more blockiness (especially on gradients)
  – Didn’t save much on ringing: 4×4 blocks have 12-pixel support instead of 8
  – Added bilinear loop filter to compensate for blockiness

<table>
<thead>
<tr>
<th></th>
<th>RATE (%)</th>
<th>DSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>-10.36612</td>
<td>0.40904</td>
</tr>
<tr>
<td>PSNRHVS</td>
<td>-4.48956</td>
<td>0.25806</td>
</tr>
<tr>
<td>SSIM</td>
<td>-12.32547</td>
<td>0.38397</td>
</tr>
<tr>
<td>FASTSSIM</td>
<td>-5.20467</td>
<td>0.17350</td>
</tr>
</tbody>
</table>
Bilinear Loop Filter

- Not a standard deblocking filter
  - Doesn’t look outside of current block!
  - Compare decoded block to bilinear interpolation of corner pixels, blend with optimal Wiener filter gain

\[ w = \min\left(1, \frac{\alpha Q^2}{12 \sigma^2}\right) \]
Bilinear Loop Filter

- Not a standard deblocking filter
  - Doesn’t look outside of current block!
  - Compare decoded block to bilinear interpolation of corner pixels, blend with optimal Wiener filter gain

\[ w = \min\left(1, \frac{\alpha Q^2}{12 \sigma^2}\right)^2 \]
Less Lapping

• In July, moved to 4-point lapping everywhere
  – Not sure this is a good change
    • Much less ringing
    • Less detail preservation
  – Small perceptual metrics changes, big visual changes
  – Help doing visual tests welcome!

• Thor deblocking experiment
  – Can now use 4-point deblocking filter from Thor instead of lapping (by flipping a #define)
  – Help testing this also welcome

<table>
<thead>
<tr>
<th></th>
<th>RATE (%)</th>
<th>DSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>-3.64821</td>
<td>0.12193</td>
</tr>
<tr>
<td>PSNRHVS</td>
<td>-1.36706</td>
<td>0.06994</td>
</tr>
<tr>
<td>SSIM</td>
<td>-2.39326</td>
<td>0.06152</td>
</tr>
<tr>
<td>FASTSSIM</td>
<td>0.35914</td>
<td>-0.01095</td>
</tr>
</tbody>
</table>
Deringing

- May remember “paint deringing” from last year

- Complex and not SIMDable
Thor’s “Constrained Low Pass Filter”

• Super-simple filter:

\[ X' = X + \text{clamp} \left( -1, \frac{A + B + C + D - 4X + 2}{4}, 1 \right) \]

• Solved long-standing “quilting artifact” problem
Thor’s “Constrained Low Pass Filter”

- Super-simple filter:

\[ X' = X + \text{clamp}\left(-1, \frac{A+B+C+D-4X+2}{4}, 1\right) \]

- Solved long-standing “quilting artifact” problem
Thor’s “Constrained Low-Pass Filter”

- Also cleans up a lot of noise
  - But would still like a proper deringing filter...
“Directional Deringing”  
(Paint Deringing Take 2)

- Estimate orientation of each 8×8 block
  - Taken from paint deringing (this part SIMDs)
- Apply \( \{2,2,3,2,3,2,2\}/16 \) filter along direction
  - If pixel differs from center by a threshold, just use center pixel instead
  - Same filter on whole 8×8 block
  - No crazy blending weights → SIMD!
- Apply \( \{1,1,1\}/3 \) filter along orthogonal direction
  - Much tighter threshold (don’t want to blur edges)
Directional Deringing
Example
Directional Deringing Example
Progress and Metrics
PCS2015
Still Image Challenge

The Xiph.Org Foundation & The Mozilla Corporation
FastSSIM Progress
Jan. 2014 to Sep. 2015

Reduced rate by 82.75%

up and left is better

HQ YouTube
LQ Video Conference
H.265
Jan
Jun
May
Apr
Feb
Sep
Nov

The Xiph.Org Foundation & The Mozilla Corporation
PSNR-HVS Progress: Jan. 2014 to Sep. 2015

Reduced rate by 80.91%

HQ YouTube
LQ Video Conference
H.265

Jan
May
Jun
Sep
Apr
Nov

The Xiph.Org Foundation & The Mozilla Corporation
IETF NETVC Working Group

- Had a successful BoF in March
- First working group meeting in July
- Also have contributions from Cisco (Thor)
  - And we are already stealing from it
- Planned spec by May 2017
Alliance for Open Media

- Amazon, Cisco, Google, Intel, Microsoft, Mozilla, Netflix so far (more to come!)
  - Just a few weeks old
  - Still figuring out how this will work
- Forum for sharing IPR research
  - Can’t do this in an SDO
- W3C-style patent commitments
- Xiph and Mozilla still committed to working in the open and publishing a spec in the IETF
Questions?