rav1e

Thomas Daede
VideoLAN Dev Days Sept 22, 2018
https://github.com/xiph/rav1e
What is this

- New, experimental AV1 encoder
- Goals:
  - Faster than libaom
  - Better quality than libaom
  - Ideally at the same time
Status since my lightning talk

~15%-30% smaller than x264
~70% bigger than libaom
Background on libaom

- Derived from libvpx codebase
- Reference implementation, “sort of usable”
- Much encoder behavior is inherited from previous VPx codecs
  - multiple frame passes
  - weird rate control
Background on libaom

- SIMD is nice, but not enough

<table>
<thead>
<tr>
<th></th>
<th>Encoding Time (seconds)</th>
<th>Times Real Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV1</td>
<td>226,080</td>
<td>45,216</td>
</tr>
<tr>
<td>x265</td>
<td>289</td>
<td>58</td>
</tr>
<tr>
<td>LibVPx</td>
<td>226</td>
<td>45</td>
</tr>
<tr>
<td>x264</td>
<td>18</td>
<td>4</td>
</tr>
</tbody>
</table>

New (and old) ideas

• Accurate RDO
• Loop filters during RDO
• Alternative distortion metrics (psy-rd)
• Rate control techniques
  – mbtree
• Speed improvement
  – Alternative parallelization methods
Brief “what is RDO” slide

- Rate-distortion optimization
- How to make an encoder:
  - Measure rate
  - Measure distortion
  - Compute RDO score
  - Try again until your video is good
rav1e block order
rav1e filter inlining
Alternative distortion metrics

- Simple implementations of RDO used MSE as the distortion metric to maximize PSNR
- Swapping the distortion metric with a different one can make your video look better (at the cost of PSNR)
- x264 called this “psy-rd” and gave a tunable parameter controlling the distortion metric
• rav1e currently has a distortion metric called “cdef-dist”
  - Runs on 8x8 blocks
  - Designed to maximize PSNR-HVS-M
  - Lots of tuning work needed!
Rate control techniques

• In AV1, you control bit allocation in a frame by using segments

• AV1 has better segment prediction than VP9, take advantage of this
  – choose quantizer in RDO (“aq”)
  – temporal rdo (“mbtree”)

14
Speed

- Iterative search space pruning
  - Search modes cheaply
  - pick 5 best
  - Search with more accuracy in RDO
  - pick 3
  - Search most accurate
  - pick 1
Speed

• Ultra-granular parallelism
• Work-stealing implementation with rayon
• Considers decisions on a sub-superblock level in parallel
Questions?