



AV1 Update

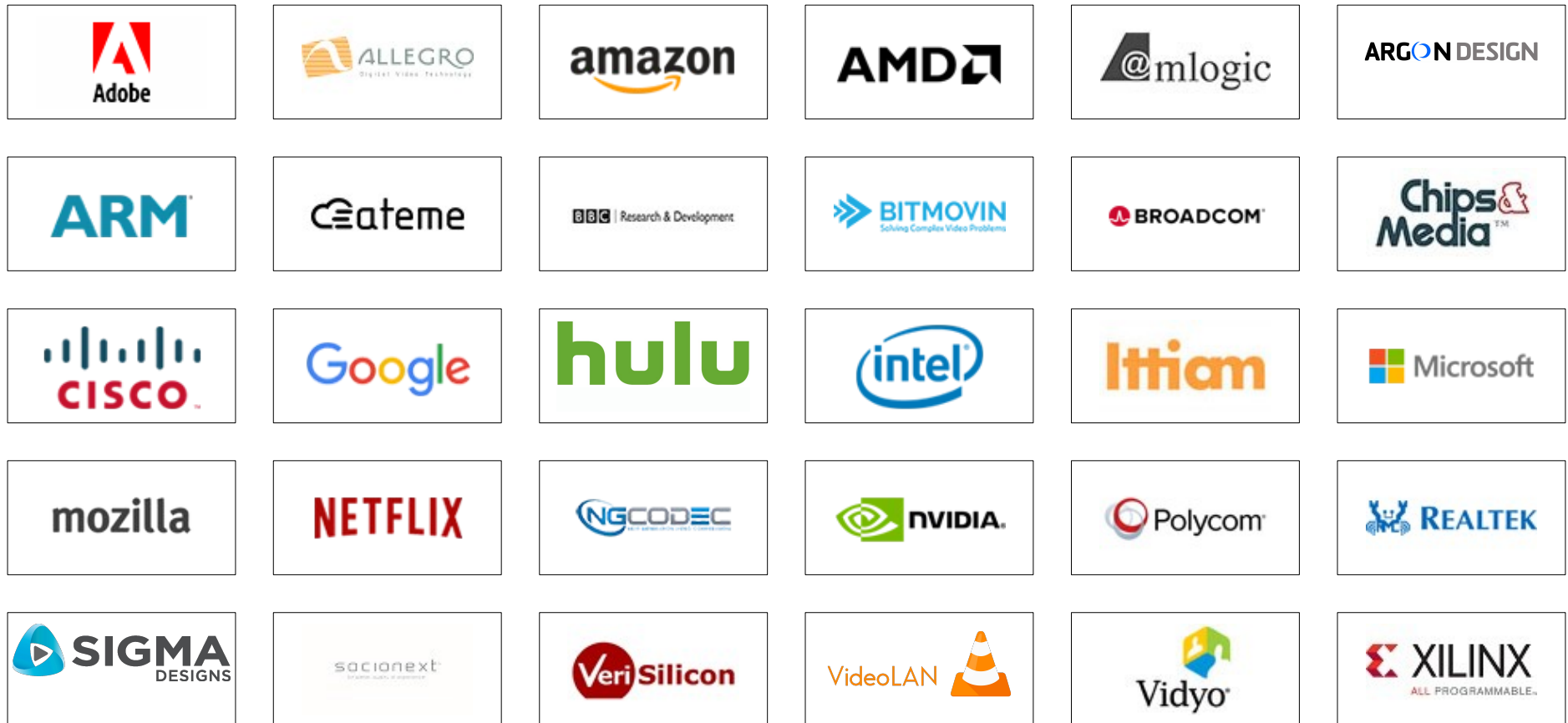
Timothy B. Terriberry



What is the Alliance for Open Media and AV1?



- Joint effort by lots of companies to develop a royalty-free video codec for the web





What is the Alliance for Open Media and AV1?



- Joint effort by lots of companies to develop a royalty-free video codec for the web





The Big Question

- Are we done yet?



The Big Question

- Are we done yet?

NO.



The Big Question

- Are we done yet?

Almost



What's left?

- Fix remaining problems with TXMG
- Final details of high-level syntax
- Last-minute changes to MV prediction
- Fix all of the bugs
- IPR analysis



Bugs



MONORAIL Project: aomedia ▾ **Issues** People Development process History [Sign in](#)

New issue Search Open issues ▾ for label:Hotlist-AV1-Normative Search [Advanced search](#)
[Search tips](#)

1 - 8 of 8 [List](#) [Grid](#)

| ID ▾ | Type ▾ | Status ▾ | Priority ▾ | Milestone ▾ | Owner ▾ | Summary + Labels ▾ | ... |
|----------------------|-------------|----------|------------|-------------|-----------------------------|---|-----|
| 1311 | Defect | New | Medium | --- | debargha@google.com | [mono video] read_palette_mode_info() shouldn't read luma palette info in monochrome mode | |
| 1309 | Defect | Assigned | Medium | --- | huisu@google.com | z1 top-right pixel check for sb128 right half | |
| 1292 | Defect | New | Medium | --- | david.ba...@argondesign.com | [wedge/compound-segment] Multiple rounding in masked blend | |
| 1291 | Enhancement | Assigned | Medium | --- | debargha@google.com | Can LRU size be smaller than super block size? | |
| 1288 | Enhancement | Assigned | Medium | --- | jingning@google.com | [opt_ref_mv] Concern about hardware implementation | |
| 1287 | Defect | New | Medium | --- | joeyoung@google.com | [txmg] AV1 iadst16 doesn't match VP9 iadst16 | |
| 1263 | Defect | New | Medium | --- | david.ba...@argondesign.com | [horzonly-frame-superres] Inconsistencies in upscale process + narrow tiles | |
| 1191 | Defect | Assigned | Medium | --- | urvang@google.com | can the normative divisions in rect_tx{,_ext} dc_pred be removed? | |

1 - 8 of 8



Specification

<https://aomedia.google.com/av1-spec/>



What's Changed?

Very technical details

Adaptive Multisymbol Entropy Coding (1)



- Even smaller multiplies
 - Replaced $8 \times 15 \rightarrow 23$ bit with $8 \times 9 \rightarrow 17$ bit multiply
 - 15-bit CDFs (probabilities) shifted down before multiply
 - Probability adaptation still happens in 15 bits
 - Reducing it causes larger losses than reducing the multiply
 - Problem: Probabilities can underflow to 0
 - Solution: Reserve small space in each interval for each symbol (costs 1 addition)
 - Bonus: No need for CDF adaptation to maintain minimum probability (cheaper adaptation)

Adaptive Multisymbol Entropy Coding (2)



- Simplified backwards adaptation
 - Used to average together CDFs from all tiles
 - Hardware didn't like buffering all of this data
 - Now just use the CDFs from the biggest tile (most coded bytes)
 - Performs basically the same



Transforms (1)

- Transforms with 4:1 or 1:4 ratio added
 - 4x16, 16x4, 8x32, 32x8
- 64-point transforms added
 - 64x64, 32x64, 64x32, 16x64, 64x16
 - Only upper-left 32x32 region allowed to be non-zero
 - Or 16x32/32x16 for 4:1/1:4 transforms
- `daala_tx` was not adopted
 - Sorry. We tried really hard



Transforms (2)

- Many problems raised by daala_tx now being addressed in TXMG
 - Order of row/column transforms now consistent
 - VP9's 4-point ADST restored
 - But it has 64-bit overflows
 - Type IV DSTs now consistent between DCT and ADST transforms (can now reuse them)
 - Extra scaling for rectangular transforms now done consistently
 - Many changes to scaling/dynamic range
- Current state:
 - Overflow handling unclear: None of C code, SIMD, or spec match



Coefficient Coding

- VP9-style token coding replaced by `lv_map`
- Code position of last non-zero coefficient up front
- Scan coefficients in multiple passes
 1. 0, ± 1 , ± 2 , $\pm 3+$
 - One 4-value symbol, special case last coeff. (non-zero)
 2. Signs of non-zero values
 3. Large values (3+)
 - More 4-value symbols, escape to Golomb code if very large
- Much smaller number of contexts/probabilities



Intra Block Copy

- New intra prediction mode
- Copies contents of *current* decoded frame
 - Location specified by “motion” vector
 - Source must be more than two superblocks prior
 - To allow pipelining in hardware decode
 - Loop filters are disabled
 - To prevent having to write back to reference frame memory twice



Motion Vector Coding (1)

- VDD 2017 recap
 - Super-complicated entropy coding scheme to indicate which predictor to use and if there's a delta
- Current status
 - Exactly the same situation, but all details changed
 - More changes possible to reduce hardware latency



Motion Vector Coding (2)

- Added “MFMV”
 - Project motion vectors from reference frames to the current frame (scaled by temporal distance)
 - Gather candidates that intersect each 8x8 block
 - Processes three 64x64 superblocks from each ref frame
 - Co-located 64x64 plus left/right neighbors
- Changed warped motion sample selection
 - Add upper-right block to list of samples
 - Remove samples very different from current MV



“Extended” Skip Mode

- When current frame has one adjacent forward and backwards reference
 - Can mark a block as an “extended” skip
 - Inter coded
 - No residual (VP9’s “skip”)
 - Compound mode
 - Using the one forward and one backward reference
 - Using best predicted motion vector for each reference
- I.e., works like the skip mode in other codecs



Loop Filtering

- Deblocking modifies 1 fewer line
 - Eliminates line buffers in subsequent CDEF and Loop Restoration filters
 - Changes to offset of Loop Restoration processing blocks and handling of superblock boundaries
 - To align them with CDEF output
 - No changes to CDEF required
- Loop Restoration: Simplified Self-Guided Filter
 - Computes self-guided filter parameters on a reduced set of pixels and interpolates
- Total line buffers for all filters: 16 (same as VP9)



Frame Super-resolution

- Not actual super-resolution
- Instead
 - Code at reduced resolution
 - Run deblocking and CDEF, but not Loop Restoration
 - Upsample with simple upscaler
 - Run Loop Restoration filter at full resolution
- Only horizontal resolution reduction allowed
 - Simplifies hardware (no new line buffers)



Spatial Segmentation



- New spatial prediction for segmentation labels
 - Used to change quantizer/loop filter on block-by-block basis
- Predictor given by majority vote of left, up-left, up neighbors (if 3-way tie use left)
- Re-orders label list so predictor comes first, nearby labels follow
 - No redundancy in encoding
- No longer required to code a segment label for skipped blocks (with no residual)
 - Unless you're using segments to signal skips or to hard-code the reference frame
 - Greatly reduces signaling overhead for adaptive quantization (activity masking) and/or temporal RDO (MB-Tree)



Other Changes

- Updated rules on cross-tile dependencies in a tile group
 - Allow low-latency encoding and re-packetizing tiles into different tile groups
- Decoder rate model
 - Constrains usage of hidden frames (alt-refs) to allow hardware to guarantee decoding without a fixed re-ordering depth (B-frames)
- CICC colorspace metadata
- Support for mono video



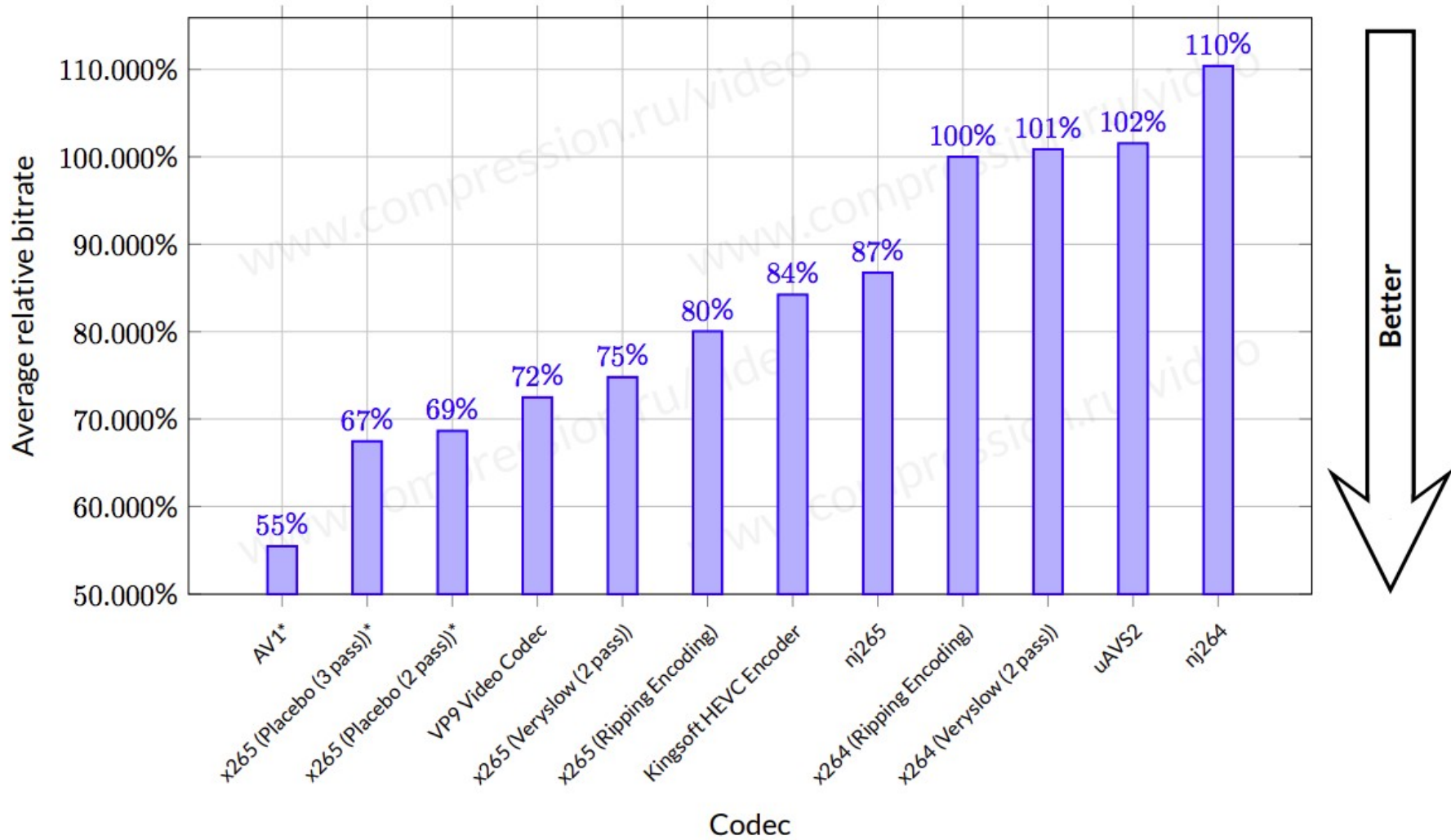
Metrics



Moscow State University (SSIM – June 29)



http://www.compression.ru/video/codec_comparison/hevc_2017/MSU_HEVC_comparison_2017_P5_HQ_encoders.pdf





Questions?