AV1 Status update

Rostislav Pehlivanov
atomnuker@gmail.com

2017-02-05
What is AV1

- Interoperable and open
- Optimized for the Internet
- Scalable to any modern device at any bandwidth
- Designed with a low computational footprint and optimized for hardware
- Capable of consistent, highest-quality, real-time video delivery
- Flexible for both commercial and non-commercial content, including user-generated content
What is AV1 (decoded)

- Royalty free
- Open development
- Lots of companies who deal with video on the internet involved
  - Will see adaption
- Lots of members own patents we can use to make the codec better
- Avoiding alien IP means we have to work around patents and possibly discover better ways than the old tried and true techniques
Reference encoder

- Reference encoder based on libvpx
  - Without VP8 support
  - With some bugfixes
- Every tool added initially as an experiment after passing review
- After passing IP review it gets enabled by default and becomes part of the codec *
- After codec bitstream gets frozen experiments that didn’t make it get removed * **

* - Hasn’t happened yet
** - Won’t happen until the end of the year
A codec is only as good as its coding tools

Currently there are over 50 experiments:

- emulate
- hardware
- clpf
- dering
- var
- tx
- rect
- tx
- ref
- mv
- dual
- filter
- convolve
- round
- ext
- tx
- tx64x64
- sub8x8
- mc
- ext
- intra
- intra
- interp
- filter
- intra
- ext
- inter
- compound
- segment
- ext
- refs
- global
- motion
- new
- quant
- supertx
- ans
- ec
- multisymbol
- loop
- restoration
- ext
- partition
- ext
- partition
- types
- unpoison
- partition
- ctx
- ext
- tile
- motion
- var
- ncobmc
- warped
- motion
- entropy
- bitstream
- debug
- alt
- intra
- palette
- daala
- ec
- pvq
- cb4x4
- frame
- size
- delta
- q
- adapt
- scan
- filter
- 7bit
- parallel
- deblocking
- loopfiltering
- across
- tiles
- tile
- groups
- ec
- adapt
- temp
- mv
- signaling
- rd
- debug
- reference
- buffer
- coef
- interleave
- entropy
- stats
- masked
- tx
- daala
- dist
- tripred
Directional deringing

- Works on an 8x8 block basis
- Searches for the overall direction of the block
- Filters perpendicularly with decaying strength, looks for deviations
PVQ

- A big update on the stone-age scalar quantization
- Provides general frequency domain prediction and activity masking
PVQ Search

- Inputs: \( N \) - number of components in a vector, \( K \) - pulses, \( X \) - vector of \( N \) components, L2 normalized
- Outputs: \( y \) - a vector of \( N \) integer components, the sum of the absolute values must be equal to \( K \)
- Condition: When \( y \) is normalized to L2 the resulting vector should match \( X \) as close as possible

The gain and the shape are decoupled.
Activity masking

- Provides better resolution in low contrast areas
- Instead of blurring and reducing detail (HEVC’s SAO oil painting art), hides quantization artifacts
Chroma from Luma

- Works in the frequency domain
- Predicts chroma coefficients from luma coefficients
- Merges Y blocks using TF if chroma is subsampled
- Only works if both Luma and Chroma planes use the same transform
Rate control

\[ \text{bits\_per\_second} = \text{scale} \cdot \text{quantizer}^{-\alpha} \]

- Predicts the bit usage per frame per quantizer
- \textit{scale} gets accurately measured and updated after encoding
- Filter \textit{scale} using a second order Bessel filter
- Extendable to two pass and chunked two pass encoding
rANS

- Fast decoding speed
- Works as a stack, encoder needs to store all symbols before reversing at the end
- Windowing rANS (restarting after some bytes) has huge overhead
Other notable mentions

- ext_tx
  - More transform types + a null transform
- Adaptive coding order
  - In case zigzag isn’t necessarily the best
- 64x64 transforms
End

Questions?