Google

AOM - AV1
How does it work?

July 2017

Pascal Massimino - Google
pascal.massimino@gmail.com

who?

skal@

AOM update - Membership

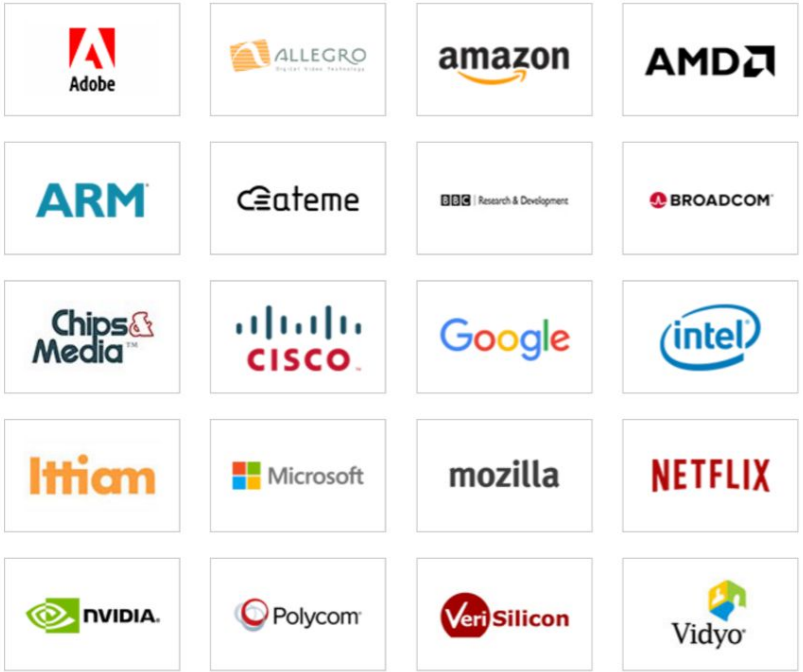
Alliance for Open Media

AOM: 25+ members and growing fast

Most recently joined: **IBM, Hulu**

Still targeting high-profile potential members

IBM hulu



AOM update - AV1

Alliance for Open Media

AOMedia launched in Q4 2015

AV1 royalty-free video codec in active development

Goal: 40-50% gain over VP9/HEVC

Rapid hardware adoption

Recent Developments

Alliance for Open Media

Successful 2-day Codec WG face-to-face in February

Attendance from 70+ engineers from member companies

AV1 currently delivering 25-35% gains across resolutions

50+ tools currently under consideration or fully adopted

Planning Towards Launch

Alliance for Open Media

Goal to achieve:

- Soft freeze of tools in Q3 2017
- Final bitstream freeze at EOY 2017

Dependencies on:

- Quality gains
- Completion of IP review
- Final approval of AOM Board

AV1 Adoption

Alliance for Open Media

YouTube will begin using AV1 immediately

Netflix will be an “early adopter” (video)

AV1 ecosystem (testing tools + SW optimizations) gearing up

vital: test streams!

Some AV1 Coding Tools

Introduction

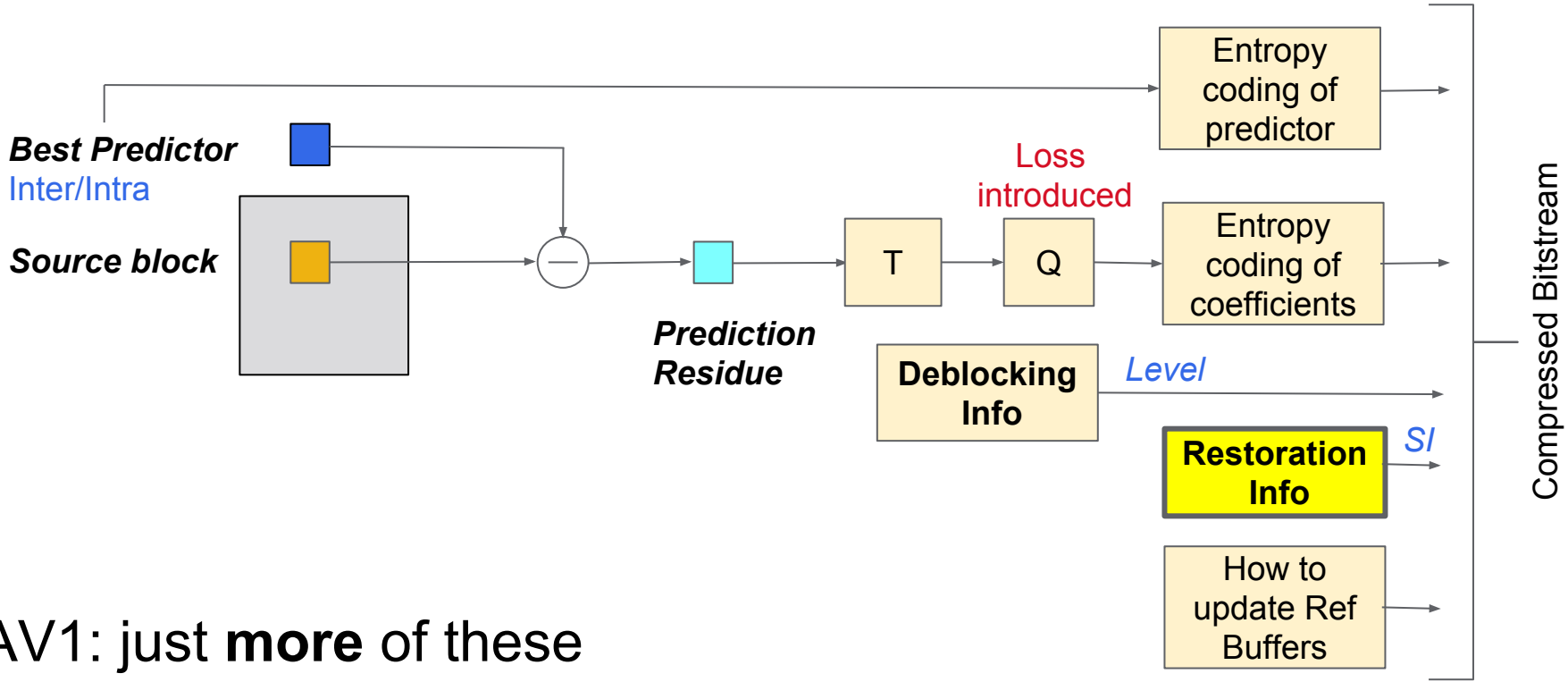
AV1 Development model

- Tools are coded as experiment, under a flag (off by default)
- Eng review (~weekly)
- IP review (including external review)
- -> decision: go/no-go.
  - Go: Flag is removed -> on by default.
  - No-go: code is removed

AV1 is not compatible with VP9, or any prior ‘VP10’.
It is an open source project:

Code: https://aomedia.googlesource.com/ (and Issue tracker)

Video compression overview



AV1: just **more** of these

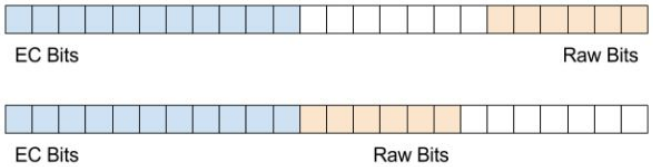
Some experiments

AV1: just **more** of these

AV1 Entropy Coding

Multi-symbol Arithmetic Range Coder*

- Allows coding of multi symbol alphabets:
  - 15-bit probabilities
  - Alphabet size up to 16 symbols
- Raw (equiprobable) bits:
  - Encoder writes from end of buffer & merges at finalize:



- One buffer per frame or tile

(longer explanation in video)

AV1 Entropy Coding

Symbol Adaptive Coding

- Binary syntax elements combined into non-binary symbols:
  - Reduces symbol throughput in the entropy coder
  - Allows (non-boolean) per-symbol probability update

- Reduced signaling overhead:
  - Explicit probability signaling not required, reduced header size
  - Better performance at lower data rates
  - Reduced overhead in error-resilient case

AV1 Block Structure

Dynamic partitioning

Maximum coding block size 64x64 (a “Superblock”)

- Recursive sub-partitioning down to 4x4 (luma)

Coding parameters maintained at 4x4 level

Prediction:

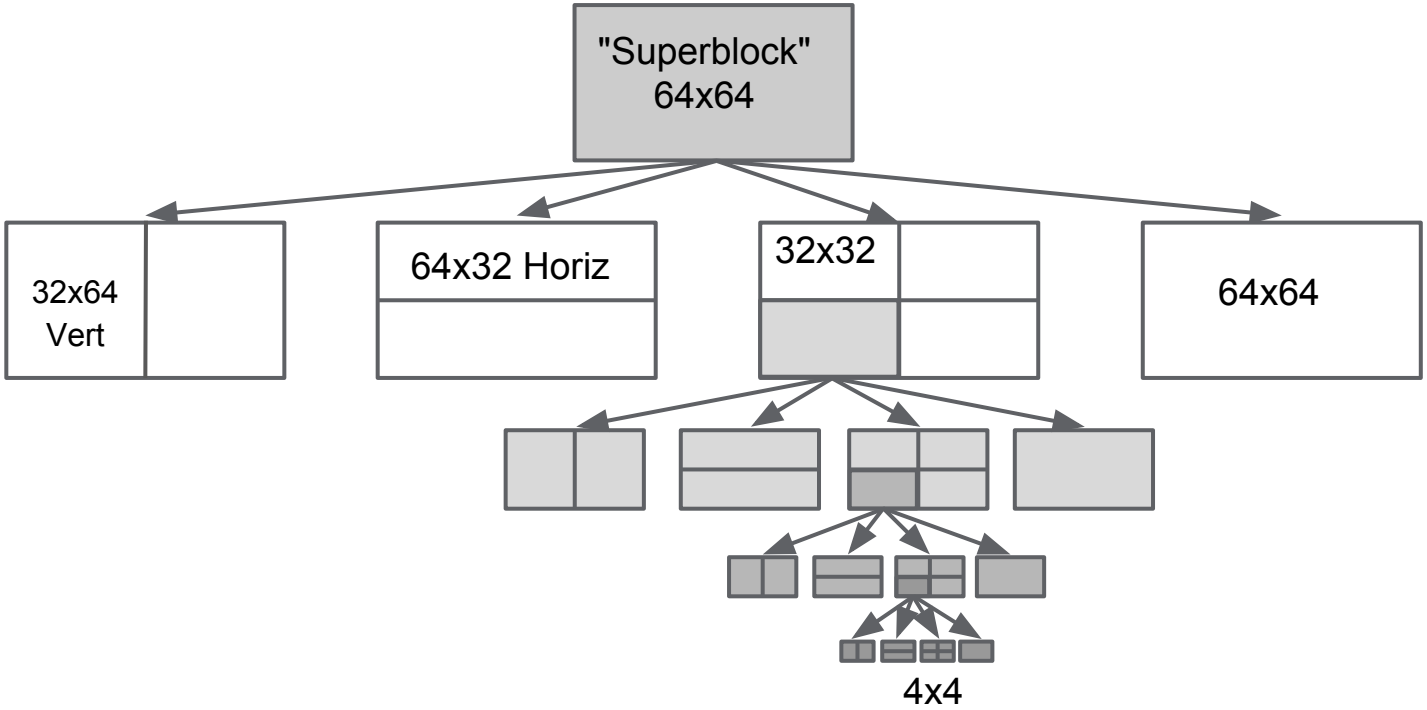
- Down to 2x2 for Chroma YUV 4:2:0 inter
- Single reference frame at 4x4, for all YUV planes

Transform:

- Minimum transform size 4x4 for intra- & inter-prediction

AV1 Block Structure

Prediction: Block Recursive Partitioning



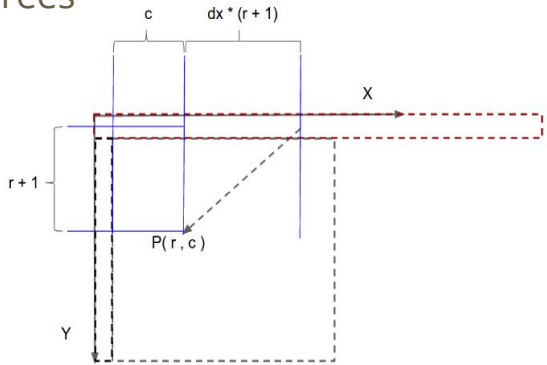
AV1 Intra Prediction

Generic Directional Predictor

prediction angle = *nominal_angle* + (*angle_delta* * *angle_step*)

nominal_angle ∈ {45, 63, 90, 117, 135, 153, 180, 207} degrees

angle_step	*angle_delta*	Total number of angles
3	[-3, +3]	8 * 7 = 56



Method:

- Back-project P(r,c) to reference pixel along selected direction
- Apply 2-tap linear interpolation filter with 1/256th pixel precision
- LUT tabulates dy corresponding to each possible dx in each available angle
  90 entries
  Each is a 16-bit integer, including 8-bits of precision

AV1 Intra Prediction

Paeth Predictor (replaces TM-pred)

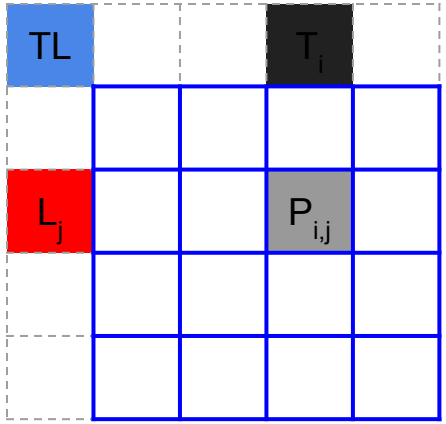
Example for a 4x4 block:

L_j = Left, T_i = Top, TL = Top-left neighbors closest to $P_{i,j}$

Paeth Predictor:

One of L_j, T_i, TL closest in value to $L_j + T_i - TL$

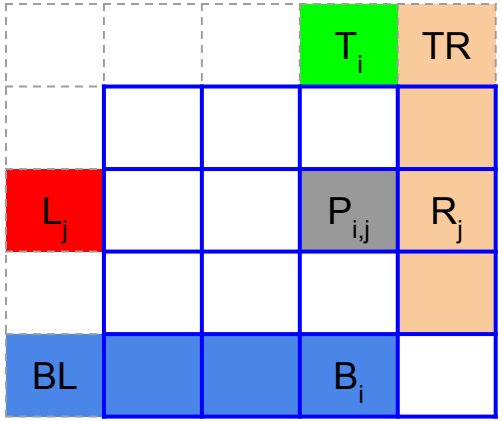
Well-defined tie-break order: L_j, T_i, TL



AV1 Intra Prediction

Smooth Predictor

- Example for a 4x4 block:
  - L_j = Left, T_i = Top, R_j = Right, B_i = Bottom, TR = Top-right, BL = Bottom Left neighbors outside of block closest to $P_{i,j}$



- Smooth Predictor:
  - Assume R_j = TR, B_i = BL
  - $P_{i,j}$ is a weighted combination of L_i, R_j, T_i, B_i
  - Weights equivalent to quadratic interpolation
  - Useful for blocks with smooth gradient

Variants: SMOOTH_H_PREV , SMOOTH_V_PRED

(0.5% intra BDRate improvement)

AV1 Intra Prediction

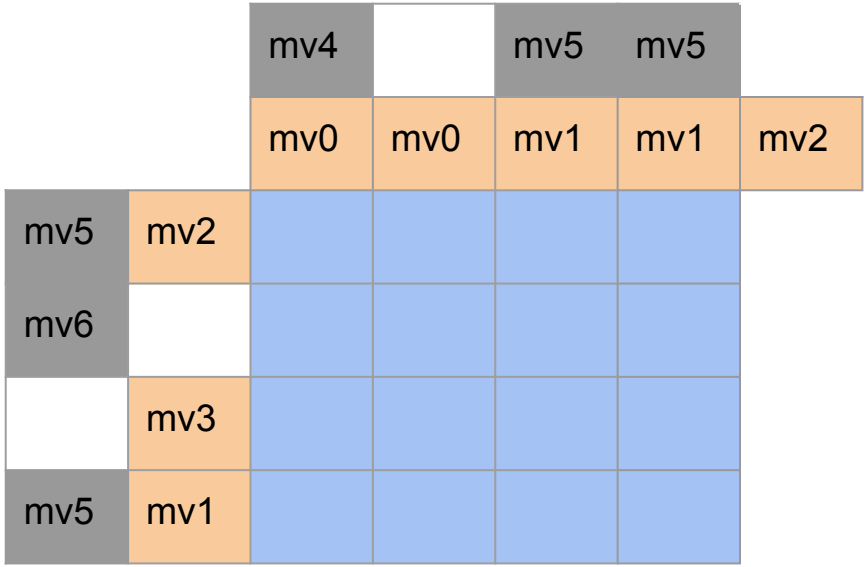
Palette **Predictor** - For Screen Content

- Palette sizes of 2 to 8 ‘base’ color entries supported
- Separate palettes for Y, U, V planes
- Palette mode available only when block size >= 8x8:
- Palette size (N) and the N ‘base’ entries are encoded in the bitstream
- A palette index is encoded, using a context model, for each pixel in the block
- Pixels predicted in ‘wavefront’ order to allow parallel computation

AV1 Inter Prediction

Dynamic Reference Motion Vector Prediction (REFMV)

- Specify the index of a MV from a list of MVs computed from neighbors
  - Examine MVs at the 8x8 block level
- Rank the MVs in the list by:
  - Distance from current block
  - Amount of overlap with current block
- Encode the index of the selected MV

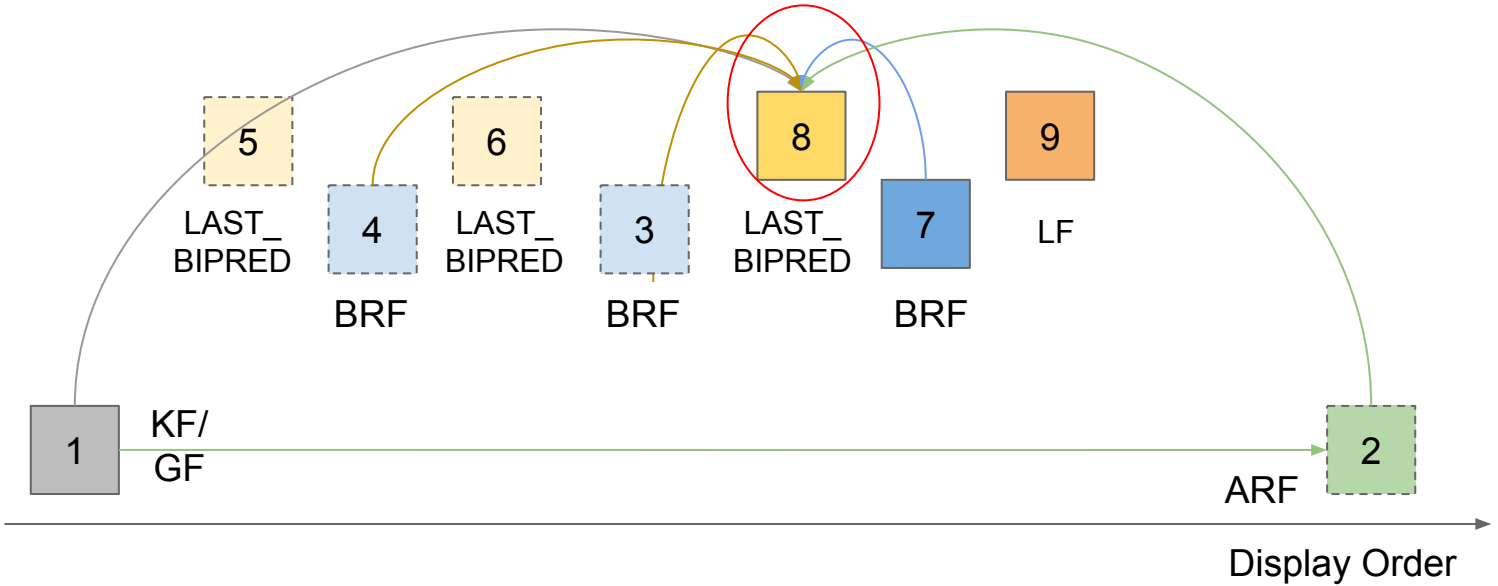


Category 1	Category 2
mv1, mv0, mv2, mv3	mv5, mv4, mv6

AV1 Inter Prediction

More Reference frames

- Upto 6 reference frames per block (VP9 specified 3)

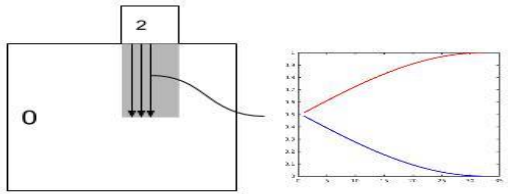


AV1 Inter Prediction

Overlapped Block Motion Compensation

- Use neighbors' predictors to refine the prediction of the current block
- Potential benefits:
  - Better prediction near to the block boundary
  - Smoother predictions / residues
- For each neighboring inter block:
  - Build a predictor, $p_2(x, y)$, using its MV
  - Combine $p_0(x, y)$ and $p_2(x, y)$: $p_{ob}(x, y) = w_0(x, y)p_0(x, y) + w_2(x, y)p_2(x, y)$
    - *Where w_0 and w_2 are defined by a 1-D raised cosine function*

COMPLEX!



raised cosine window based weighting

AV1 Inter Prediction

Global & Warped Motion

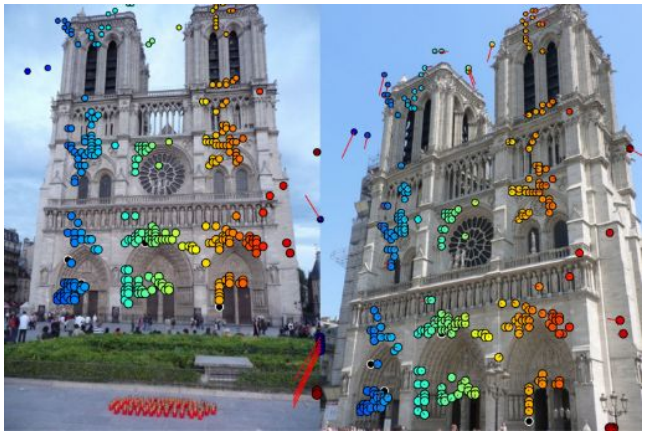
- Global Motion:
  - GM parameters computed for each reference frame using feature matching + RANSAC
  - Supported motion models (& DoF):
    - Translation (2), similarity (4), affine (6), homographic (8)
- Warping:
  - Implemented as 2 shears:
  - Using 8-tap filters
- Motion parameters coded in the bitstream

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \gamma & 1+\Delta \end{bmatrix} \begin{bmatrix} 1+\alpha & \beta \\ 0 & 1 \end{bmatrix}$$

Original warping matrix | Vertical Shear | Horizontal Shear

Combined GM + WM + OBMC

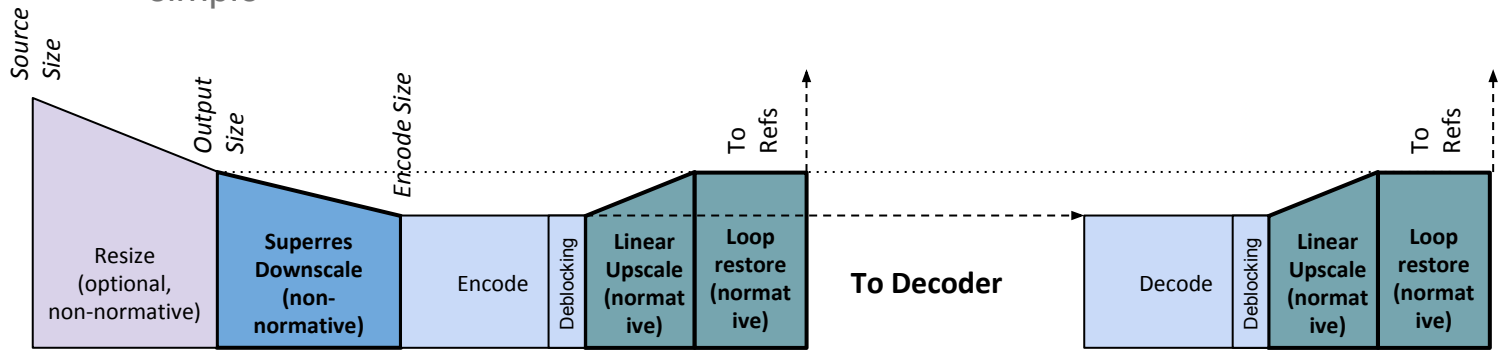
-> 2.0 - 6.0% gain (esp. Handheld camera)



AV1 Inter Prediction

Other Tools

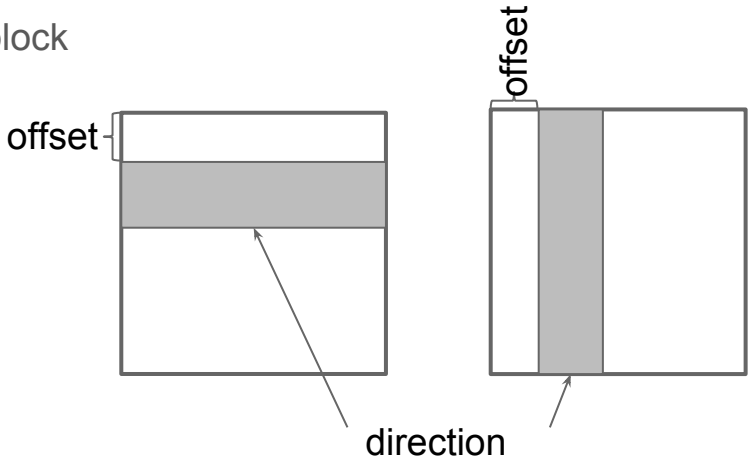
- Guided restoration (filters and upscaling):
  - Downscale before encoding, upscale before updating reference
  - Parameters transmitted in the bitstream (learned Wiener filters)
  - Super-resolution = simple linear upscale + loop-restoration
  - 10-30% rate reduction for super-res
  - ~3% average gain for switchable restoration 7x7 filters
  - simple



AV1 Inter Prediction

Other Tools

- Dual filter:
  - Independent selection of horizontal and vertical interpolation filters
- Masked transform:
  - Only transform a portion of a square block



AV1 Transform

Extended Transform Set

4 Transform types:

DCT, ADST, Flipped (reverse) ADST, Identity

Upto 16 transform combinations with independent horizontal & vertical selection

Reduce available combinations based on block size (for efficiency)

Identity transform useful for screen content coding

Rectangular transforms added:

4x8, 8x4, 8x16, 16x8, 16x32, 32x16

Used only on rectangular partitions of corresponding size

AV1 Quantization

Non-Linear Quantization & Delta QP Signaling

Non-linear quantization matrix:

- Can also support a linear model

Delta QP values can be signaled at the Superblock level (64x64)

AV1 Hardware IP

Google’s codec HW group developing a free AV1 HW decoder

Target features

- Full hardware acceleration
- AV1 and/or VP9 decode support
- 8 and 10 bits per pixel
- 2160p 60 fps at 600 MHz / core
- 3 Mgates logic estimated silicon area

AV1 Hardware IP

Licensing

Availability

- Early pre-releases available from Q4 / 2017
- Final release at the time of AV1 bitstream freeze

Licensing

- royalty free license
- Includes C and Verilog source code
- Modification rights

Thanks!

questions?