From VP9 to AV1 and Beyond

Debargha Mukherjee
Google

Outline

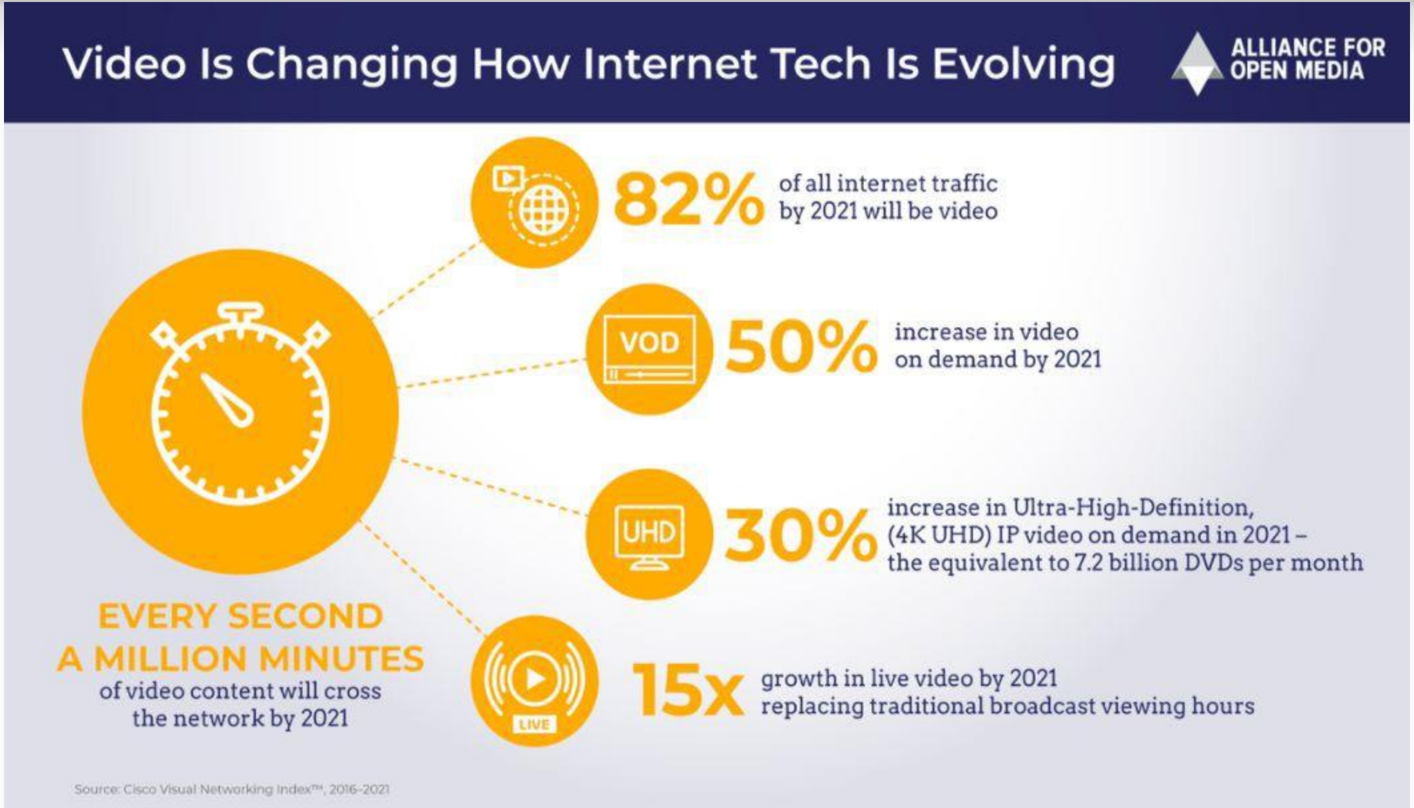
- Alliance for Open Media and AV1
- Selected Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

Outline

- Alliance for Open Media and AV1
- Selected Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

Alliance for Open Media and AV1

Projections



Fast advances in video compression technology is essential to cope with the ongoing explosion in Internet video

Alliance for Open Media and AV1

Towards royalty-free codecs

- Open technologies have aided innovation and fueled Internet growth over the last few decades.
    - Content that consumes bulk of its bandwidth (video), should be open/free too.
- Video patent landscape has been stifling innovation
    - HEVC - patent pool mess
- Google launched WebM project in 2010
    - To develop open royalty-free formats for the web
- Codecs from the WebM Project
    - **VP8** - 2010 (Hangouts)
    - **VP9** - 2013 (YouTube)
        - First serious challenge to an MPEG codec
    - **VP10** - 2014-2016, morphed into AV1

Alliance for Open Media and AV1

Founding Members

- Many other companies began to see eye-to-eye with Google after VP9
- AOM: Industry consortium to build royalty-free codecs formed in 2015

FOUNDING MEMBERS

amazon	Apple	arm	cisco
facebook	Google	IBM	intel
Microsoft	moz://a	NETFLIX	NVIDIA

Alliance for Open Media and AV1

Promoter Members

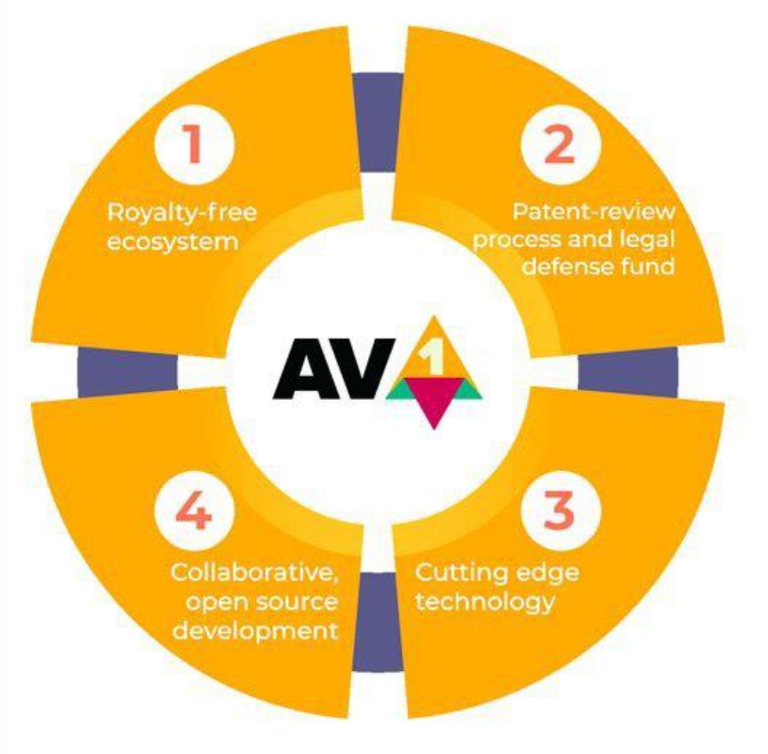
PROMOTER MEMBERS

Adobe
ALLEGRO
AMD
Amlogic
ARGON DESIGN
ateme
BBC R&D
金山云
BITMOVIN
BROADCOM
CableLabs
Chips & Media
hulu
Ittiam
iQIYI 爱奇艺
NGCODEC
Polycom
REALTEK
SIGMA DESIGNS
socionext
VeriSilicon
Vidyo
Videolan
vimeo
微帧 Visionular
XILINX

Alliance for Open Media and AV1

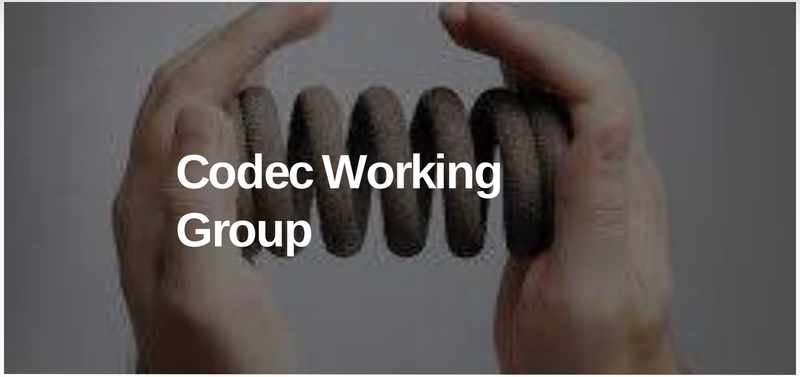
AV1

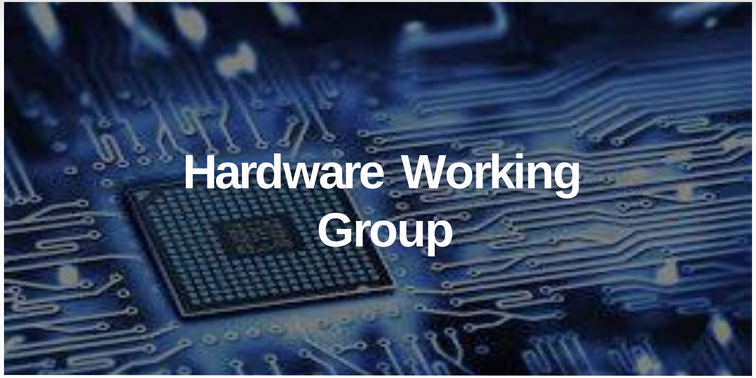
- AV1 - the first video codec from Alliance for Open Media
    - Goal - to achieve about 30% bitrate reduction over VP9 with royalty-free technologies
- Starting point was VP9+
- Tools proposed from:
    - VP10 (Google)
    - Daala (Mozilla)
    - Thor (Cisco)
    - New tools ...



Alliance for Open Media and AV1

Workgroups



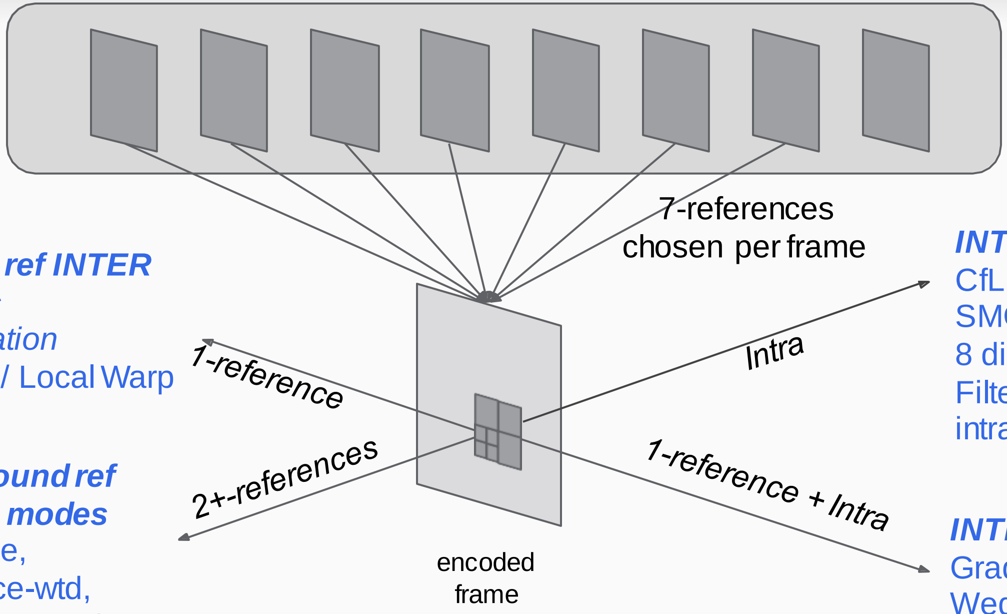


Outline

- Alliance for Open Media and AV1
- Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

Selected Coding tools

Prediction Framework



Single ref INTER modes
Translation
Global / Local Warp

Compound ref INTER modes
Average,
Distance-wtd,
Difference-wtd,
Wedge
OBMC

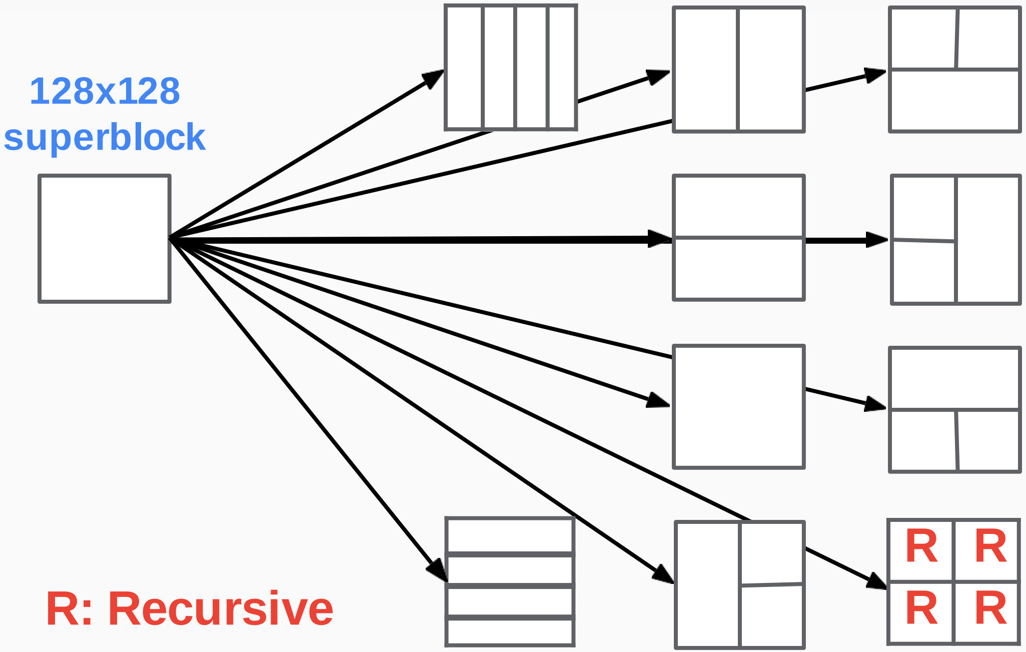
INTRA modes:
CfL, DC, PAETH,
SMOOTH, SMOOTH_H, SMOOTH_V,
8 directional modes + finer directions,
Filter-intra modes, palette,
intra-block-copy

INTER-INTRA modes:
Gradual Inter-intra
Wedge Inter-intra

Selected Coding tools

Partition Structure

- 10-way recursive partition



Selected Coding Tools

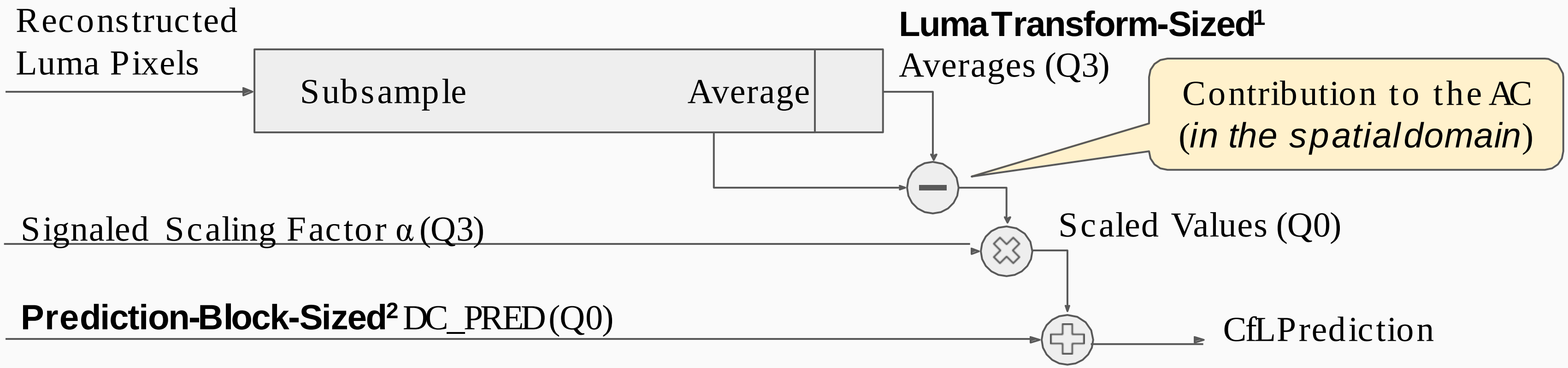
Intra Prediction

- DC Mode
- Directional Modes
- Paeth Mode
- Smooth/Smooth_h/Smooth_v Modes
- **Chroma from Luma**
- **Recursive Intra Filter**
- Intra block copy Mode
- Palette Mode

Selected Coding Tools

Intra Prediction - Chroma from Luma

- Mechanism for INTRA chroma prediction
  - Use reconstructed luma to predict chroma components



α_{Cb}, α_{Cr} signaled in bit-stream

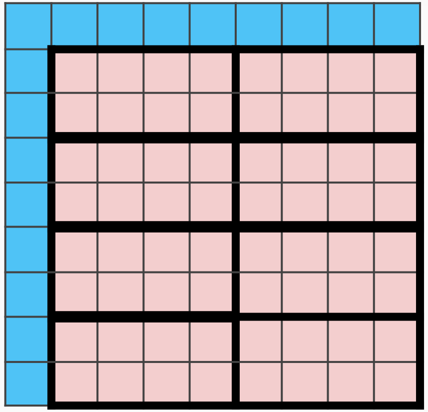
1. Luma average computed over the luma transform block
2. Chroma DC_PRED computed over prediction block

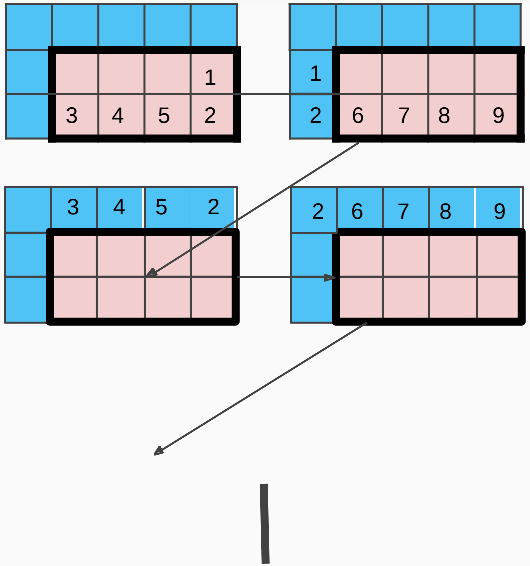
Selected Coding Tools

Intra Prediction - Recursive Intra Prediction

- Recursive Intra Prediction
  - Predict in batches of 4x2 pixels blocks
  - Apply eight 7-tap filters to get prediction

    for each of the 8 pixels
    - 5 sets of filters for 5 modes
  - Apply recursively using the predicted pixels as neighbors for adjacent blocks.





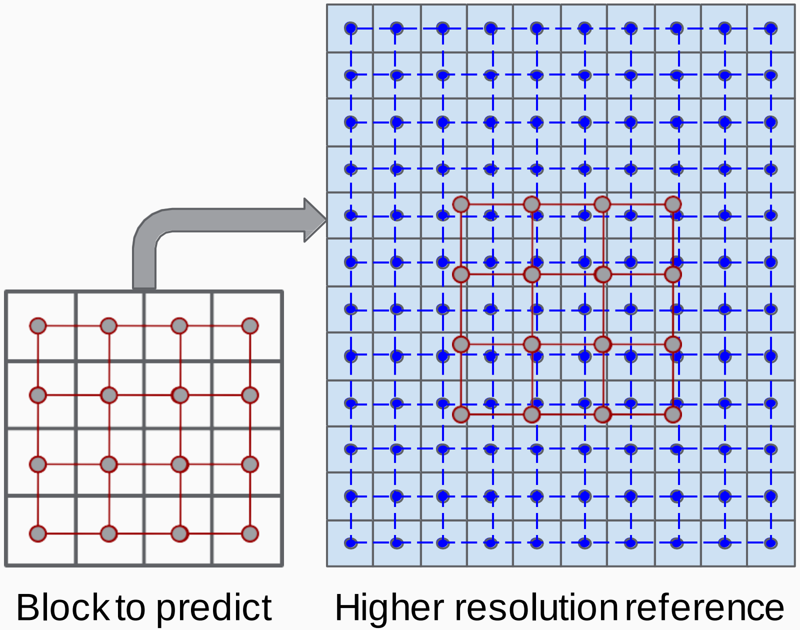
Selected Coding Tools
Inter Prediction

- Expand reference frames (up to 7 / frame) used for prediction
- Dynamic Motion Vector Referencing
  - Spatial neighborhood and Temporal neighborhood candidates
- Subpel filtering
  - Dual Switchable Subpel filters
  - High intermediate precision for filtering and compounding
- **Scaled Inter Prediction**
- **Compound Inter-Inter prediction**
- **Compound Inter-Intra prediction**
- **Overlapped Block Motion Compensation**
- **Affine warp**
  - **Global warp and Local warp**

Selected Coding Tools

Scaled INTER Prediction

- AV1 can predict from references at different resolution (limit: ½ to 2)
- Relative pixel positions hor (vert) are the same in each row (col)
    - Implies scaling can be implemented as separable filtering where step-sizes between pixels in hor (vert) directions are same
    - Need higher precision for starting offset and steps, but the same 1/16th pel filters as in unscaled prediction can be used.



Selected Coding Tools
Compound Inter Prediction

- Compound = combine two INTER predictors
  - Average weighted
  - Distance weighted
    - Uniform weight based on distance of ref frame from current frame
  - Difference weighted
    - Weight depends on difference of two prediction values per pixel
    - Average when pixel values are close, prefer one when they are different.
  - Wedge weighted
    - Wedge codebook index provides the weights

Selected Coding Tools

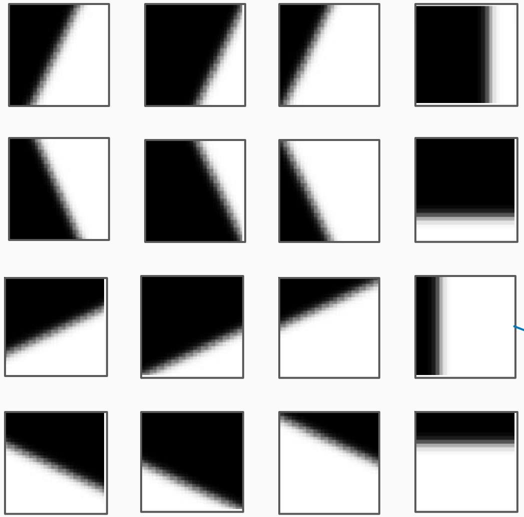
Compound Inter Prediction - Wedge Compound Prediction

Wedge codebook:

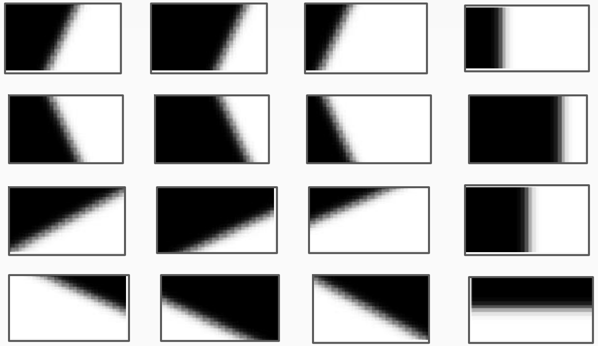
Inter-Inter
4-bit shape
1-bit sign

5 bits total

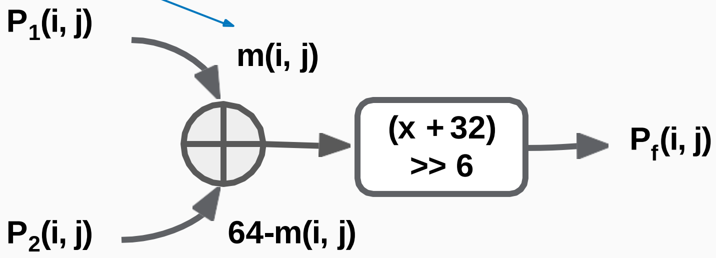
Used for 8x8 up to 32x32 sizes



Square codebook



Rectangular codebook



Selected Coding Tools

Compound Inter Prediction - Wedge Weighted

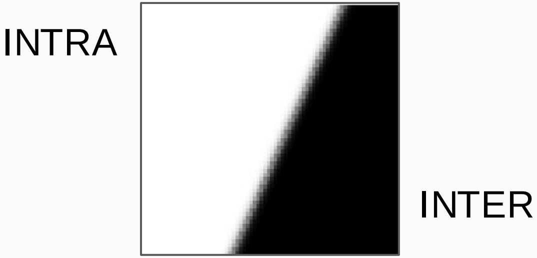
- Capture shapes of differently moving objects better
- Soft transition allows a limited number of wedges to be sufficient

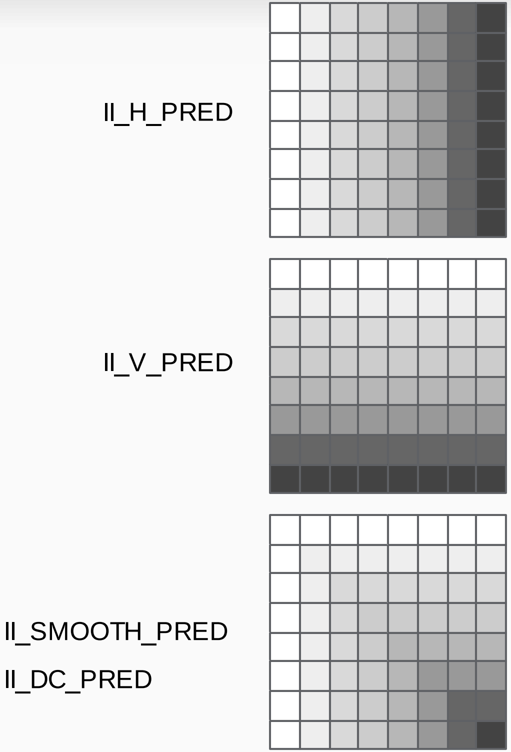
Selected Coding Tools

Inter-Intra Prediction

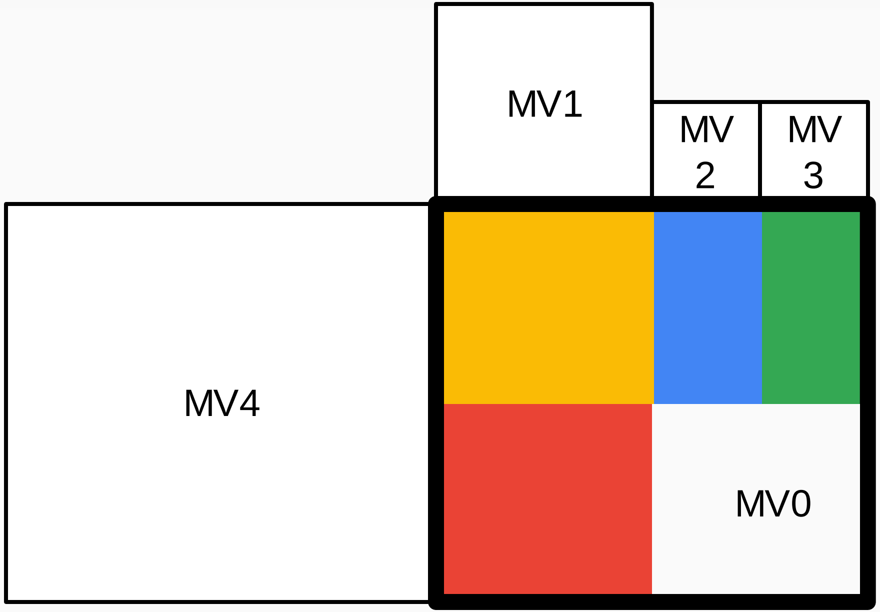
- Combination of an INTER and INTRA predictor
  - Only 4 INTRA modes allowed
- Types
  - Gradual INTER-INTRA modes
  - Wedge INTER-INTRA mode

4 modes





Selected Coding Tools
Variable block-size OBMC



- Create predictions from neighbors' MVs
- Mitigate the effect of discontinued motion field
- 2-sided causal overlapped predictor (top/left halves)
- Blend with 1-D smooth filters

Selected Coding Tools

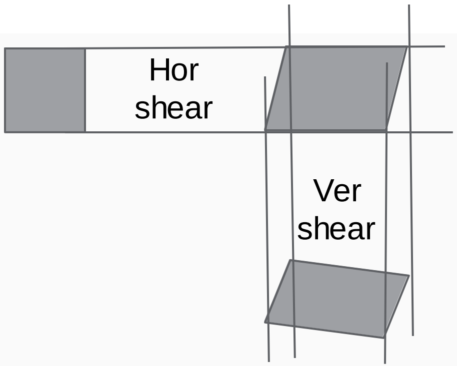
Affine warps

- True motion is never translational
- More complex motion models have not traditionally been used in video codecs
  - Parameter cost
  - Computation complexity
- AV1
  - Introduces an efficient affine motion compensation technique

Affine model $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix}$

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

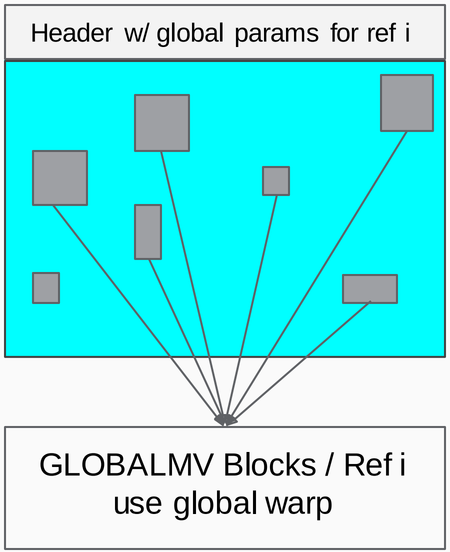
Original warping matrix   Vertical Shear   Horizontal Shear



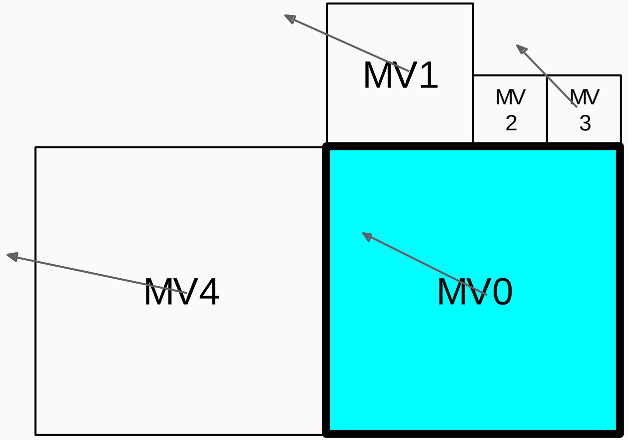
Selected Coding Tools

Affine Warps - Global and Local

- Global warp mode
  - Estimate and send affine parameters for every frame per reference
  - Invoke Global warp with GLOBALMV mode



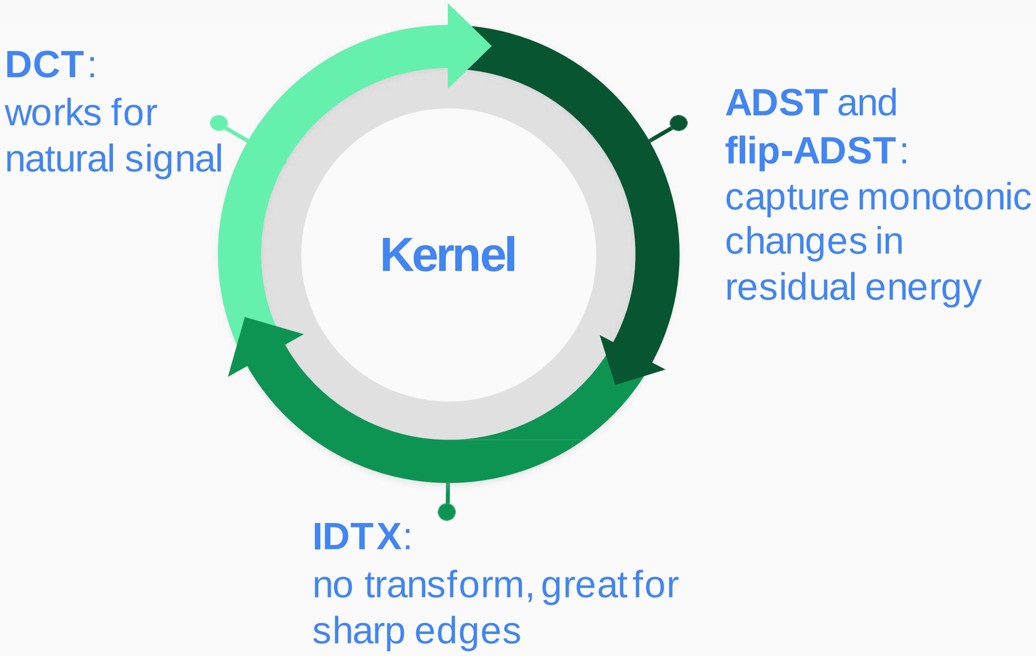
- Local warp mode
  - Estimate affine parameters from neighborhood of current block



Blocks 1, 3, 4 have same reference as current block

Selected Coding Tools
Switchable Transforms

- Expanded transforms
  - 16 possible transform kernels
  - Separable kernels where in each direction one of {DCT, ADST, FlipADST, IDTX} can be selected
  - Larger transforms reduce the number of possible transforms

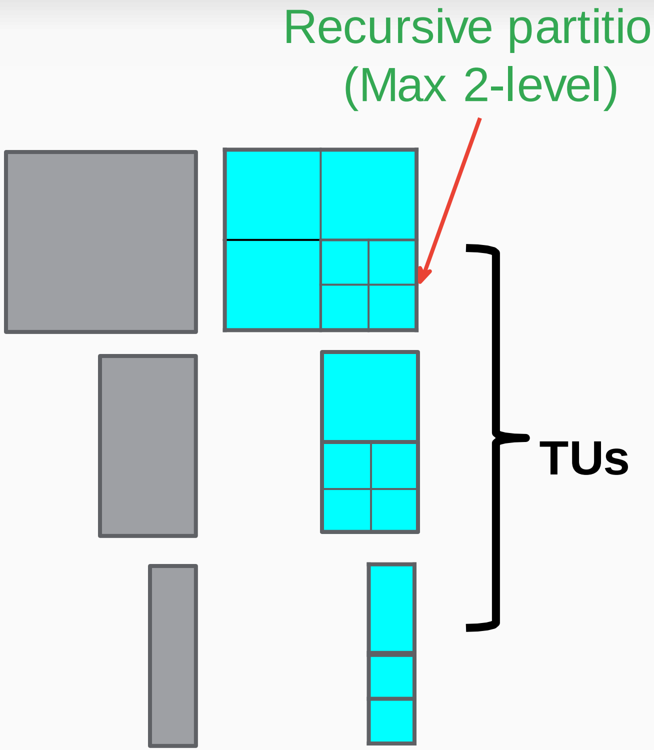


Selected Coding Tools

Transform Partitions

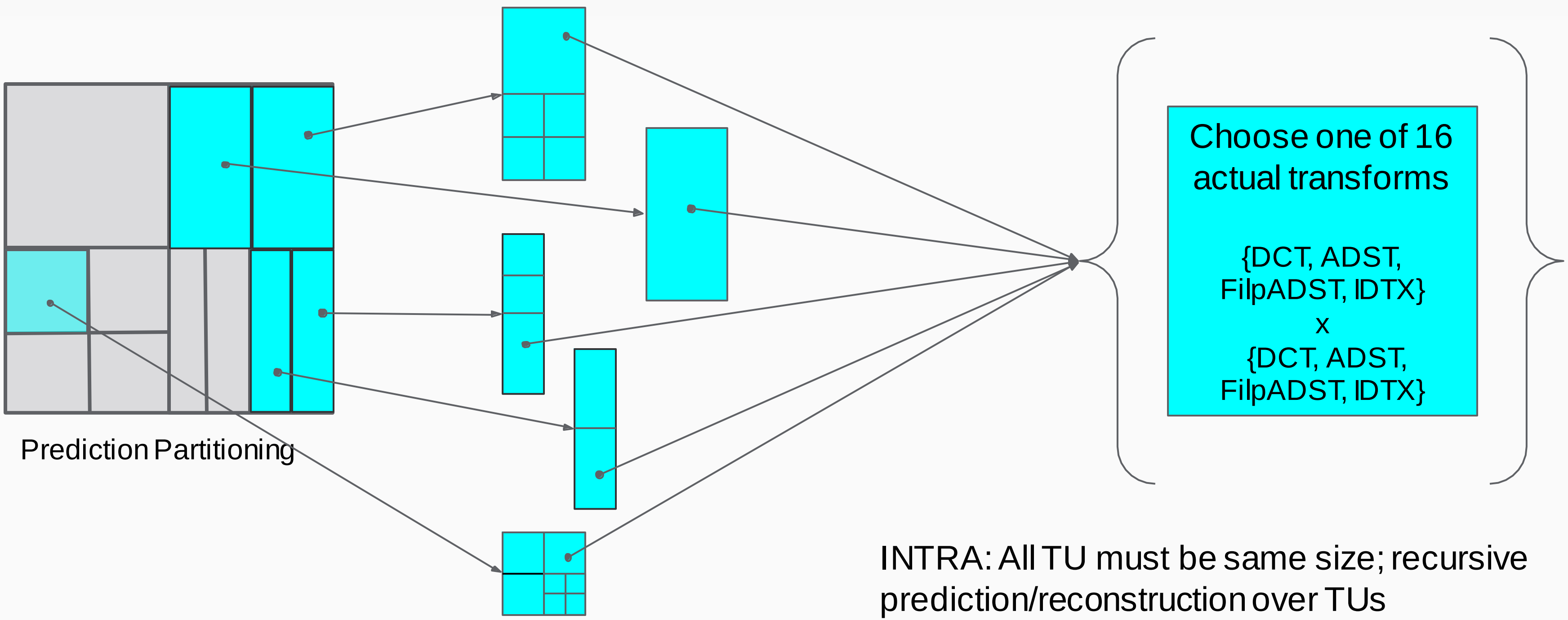
AV1

64x64, 64x32, 32x64, 64x16, 16x64
32x32, 32x16, 16x32, 32x8, 8x32
16x16, 16x8, 8x16, 16x4, 4x16
8x8, 8x4, 4x8
4x4



Selected Coding Tools

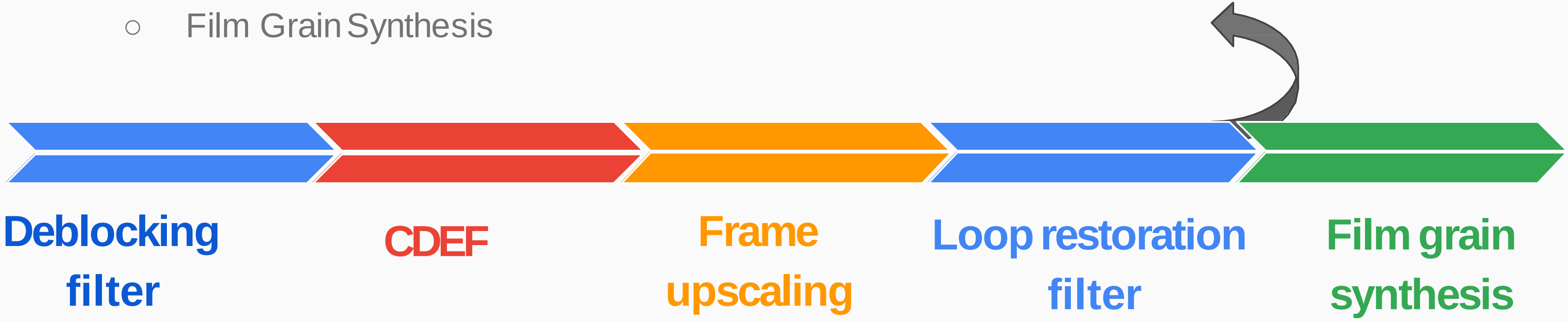
Transform Summary



INTRA: All TU must be same size; recursive prediction/reconstruction over TUs

Selected Coding Tools
In-loop filtering and postfilter

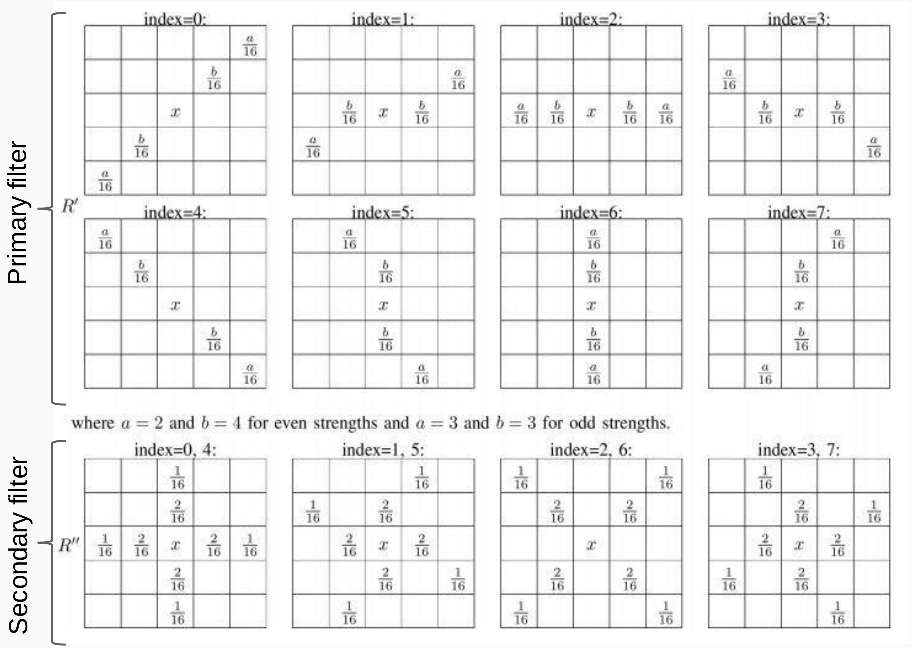
- AV1 pushes the envelope in in-loop filtering
  - Deblocking
  - CDEF (Constrained Directional Enhancement Filter)
  - Upscaling
  - Loop-restoration
- Also adds an out-of-loop filter
  - Film Grain Synthesis



Selected Coding Tools

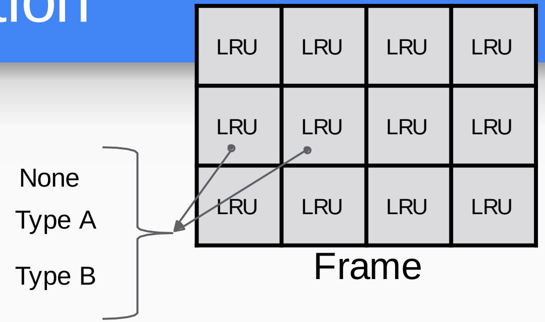
In-loop filtering / postfilter - CDEF

- CDEF:
  - Combination of Daala Deringing filter and Cisco's Constrained Low Pass Filter
  - Adapts filtering to direction of edges and patterns in an image
- Direction estimation:
  - Conducted at 8x8 level by minimizing variances along predefined lines
- Use a nonlinear directional filter
  - 5x5 support region
  - Combined Primary + Secondary filter



Selected Coding Tools

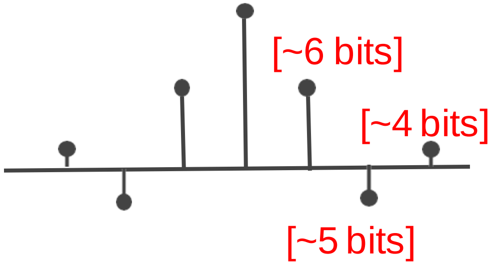
In-loop filtering / postfilter - Loop Restoration



- Switchable restoration filter in blocks referred to as Loop-restoration unit (LRU)
  - Filter type and corresponding parameters

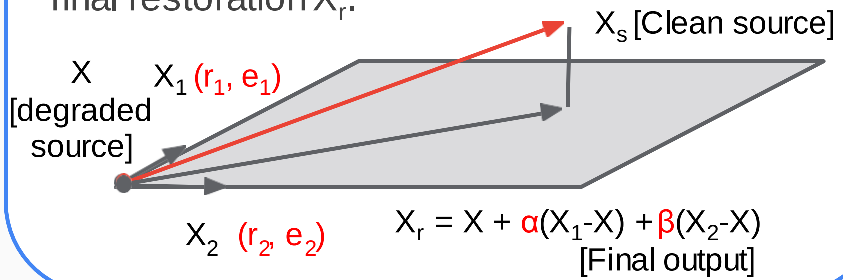
Type A: Wiener filter

Separable (horz + vert filter) 7-tap, symmetric, normalized



Type B: Dual Self-guided filters

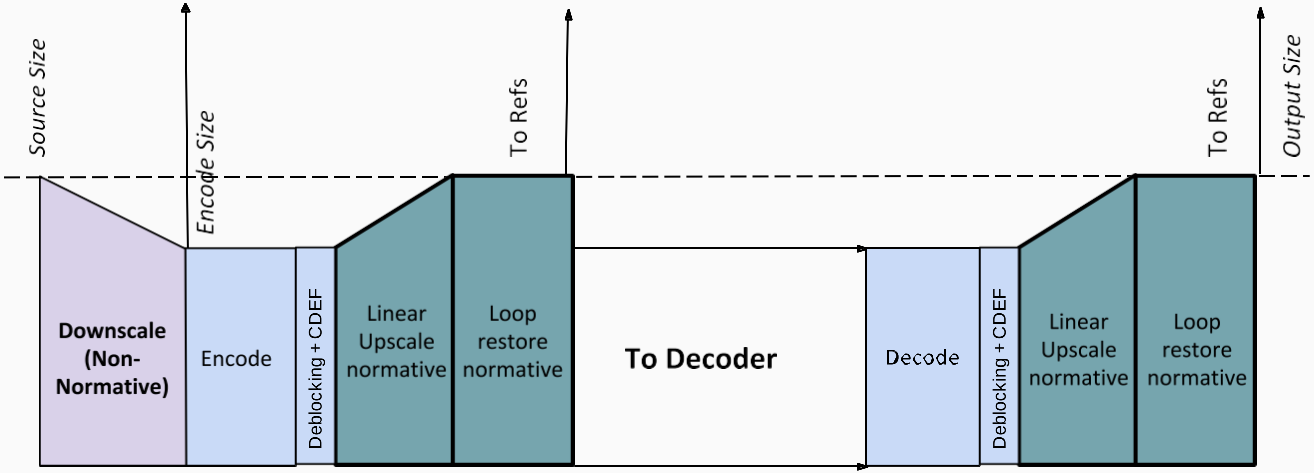
X_1 and X_2 are cheap restored versions. Subspace projection can yield a much better final restoration X_r.



Selected Coding Tools

In-loop filtering / postfilter - Frame superresolution

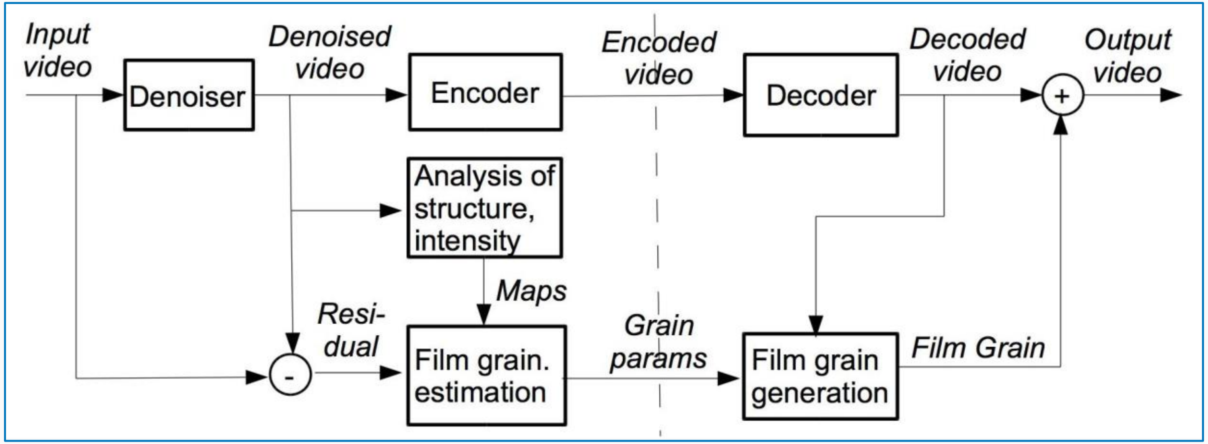
- AV1 supports predicting blocks from references at different resolutions
- Loop-restoration filter can super-resolve
- Frame-superresolution depends on the above two properties
    - Only horizontal upscaling supported to control line-buffer



Selected Coding Tools

In-loop filtering / postfilter - Film Grain Synthesis

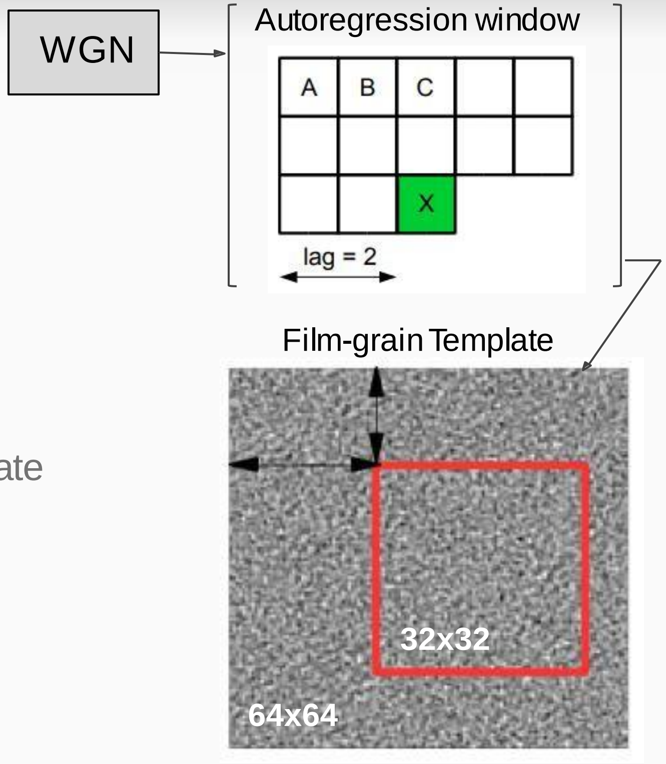
- Film-grain is hard to compress, but present in most commercial content.
  - Need to be preserved as part of creative intent
- AV1 supports film grain synthesis via a normative post-processing outside the coding loop.



Selected Coding Tools

In-loop filtering / postfilter - Film Grain Synthesis

- Generate WGN sequence
  - Generated by taking random samples from a 2048-ary predefined Gaussian noise sequence
- Generate Film-grain template
  - Filter AWGN sequence with autoregressive parameters sent in the bitstream
- For each 32x32 block:
  - Take pseudo random offsets from the Film-grain template
  - Scale based on intensity of reconstruction using piecewise Lut specified in the bitstream
  - Add scaled grain to reconstruction and clip
  - 2-line overlapped grain addition to prevent sudden changes in grain



Outline

- Alliance for Open Media and AV1
- Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

Latest Coding Results

AWCY

- www.arewecompressedyet.com
  - Official test infrastructure for AV1
  - Test set - Objective-1-fast, constant quality encoding, 60 frames, resolutions 360p-1080p
- Compare VP9 vs. HEVC vs. AV1
  - VP9 - tip-of-tree Oct 6, 2018, hash 4a47ef814b57d16787e6331e4ac1bd9dc093459e
  - AV1 - tip-of-tree Oct 6, 2018, hash 9b21428c86af1c081ae87cc546a360eaefa8ba8a
  - HEVC - X.265 hash e293b13373b72dfd88a91e196fbb595b027da3a3; HM 16.17

Latest Coding Results

Constant Quality BDRATE - AWCY Objective-1-Fast

- VP9 vs. X.265 vs. AV1
  - Testset: Objective-1-Fast on www.arewecompressedyet.com;
    - 60 frames, single keyframe, constant quality
  - Baseline: VP9 (libvpx); Tests: X.265 (--preset placebo --no-wpp --tune psnr), AV1 (libaom)

Codec	PSNR-Y	PSNR-HVS	SSIM	CIEDE2000	PSNR-Cb	PSNR-Cr
X.265	-4.04	-1.35	+7.34	+10.76	+20.51	+19.70
AV1 (cpu-used=0)	-30.16	-29.70	-30.46	-31.85	-33.01	-34.67
AV1 (cpu-used=1)	-28.75	-28.40	-29.21	-30.94	-32.76	-34.16
AV1 (cpu-used=2)	-26.10	-25.68	-26.67	-28.54	-30.61	-31.77

Latest Coding Results

Constant Quality BDRATE - AWCY Objective-1-Fast

- After recent VP9 improvements with pyramid structure
- Baseline VP9* (--auto-alt-ref=5); Tests: X.265, AV1

Codec	PSNR-Y	PSNR-HVS	SSIM	CIEDE2000	PSNR-Cb	PSNR-Cr
X.265	-1.60	+1.92	+7.74	+10.12	+18.20	+17.21
AV1 (cpu-used=0)	-27.74	-26.88	-29.28	-31.68	-33.58	-35.25
AV1 (cpu-used=1)	-26.31	-25.55	-28.03	-30.78	-33.28	-34.76
AV1 (cpu-used=2)	-23.55	-22.69	-25.41	-28.33	-31.11	-32.31

Latest Coding Results
Constant Quality BDRATE - Google testsets

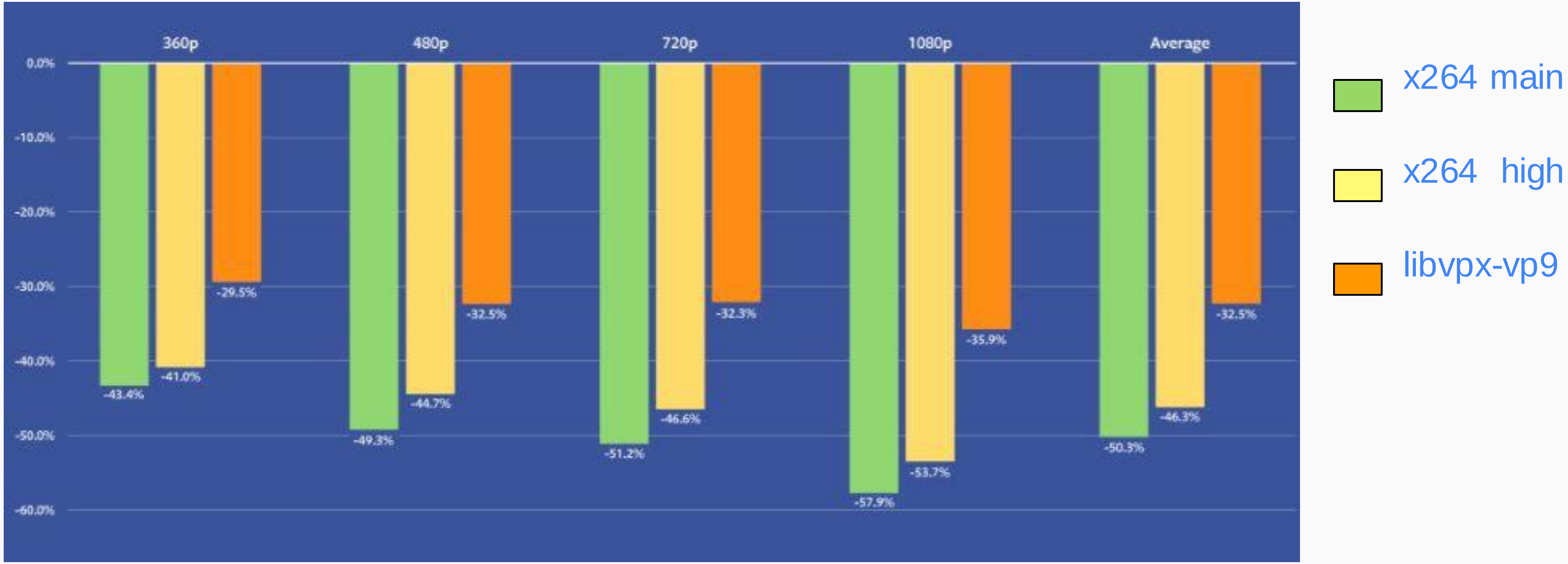
- HM16.17 vs. VP9 vs. AV1
  - Testset Google's internal (Lowres, Midres, Hdres), 120 frames
    - Single keyframe, constant quality mode
  - Baseline: HM16.17 (-ip -1 --ConformanceWindowMode=1); Tests: VP9 (libvpx), AV1 (libaom)

	Av. PSNR	**SSIM**
Lowres (120 frames), VP9	+10.448%	+5.437%
Lowres (120 frames), AV1	-19.578%	-22.703%
Midres (120 frames), VP9	+12.215%	+7.477%
Midres (120 frames), AV1	-19.840%	-23.015%
Hdres (120 frames), VP9	+7.637%	+5.240%
Hdres (120 frames), AV1	-23.979%	-26.152%

Latest Coding Results

Tested by third-parties

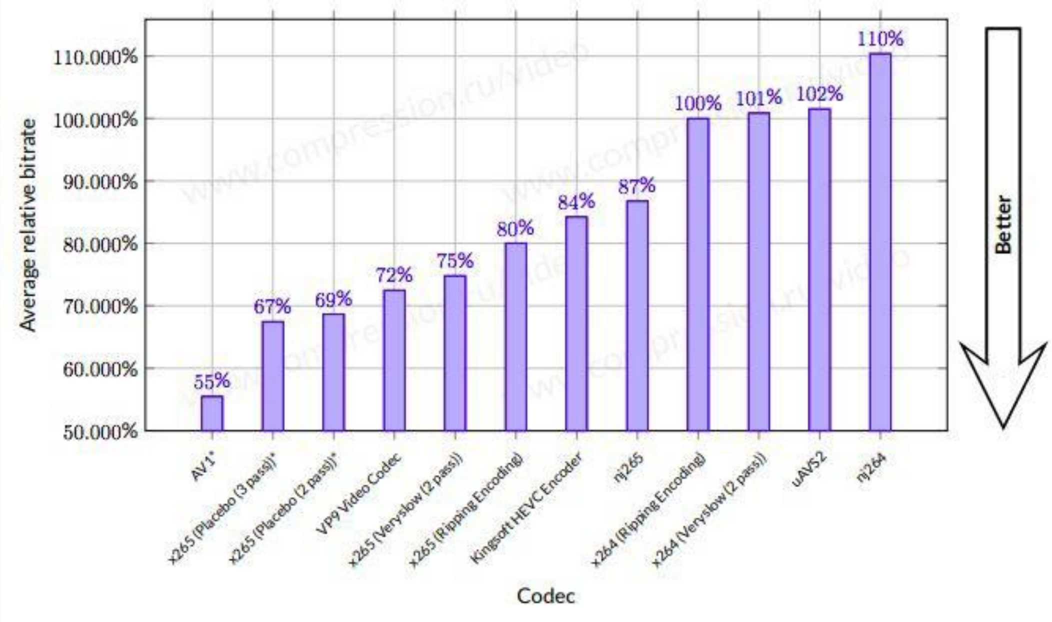
- Facebook: *"AV1 beats x264 and libvpx-vp9 in practical use case"*



[1]. https://code.fb.com/video-engineering/av1-beats-x264-and-libvpx-vp9-in-practical-use-case/

Latest Coding Results
Tested by third-parties

- MSU Video codec comparison

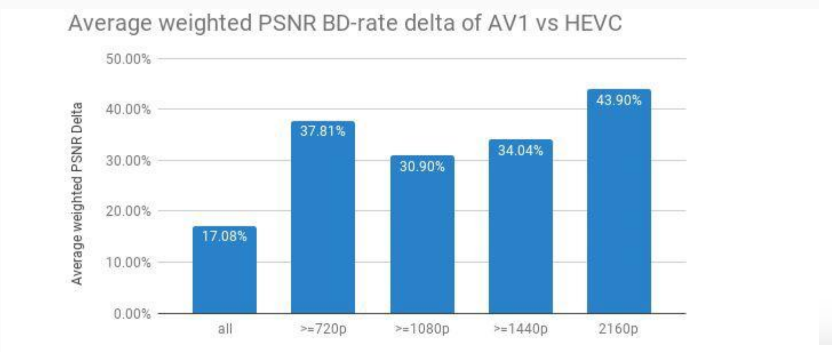


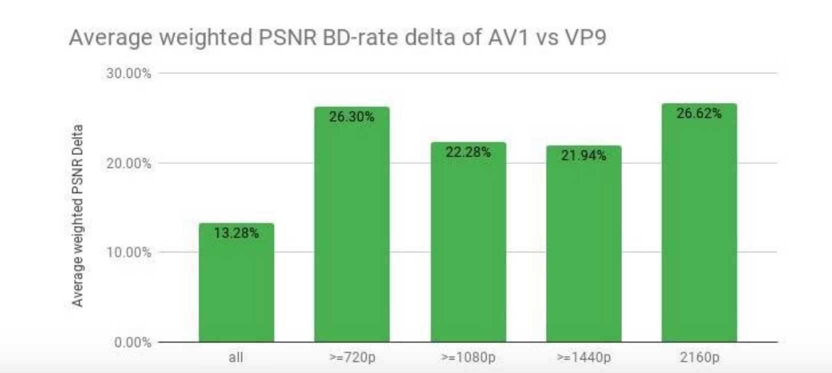
[1].http://www.compression.ru/video/codec_comparison/hevc_2017/MSU_HEVC_comparison_2017_P5_HQ_encoders.pdf

Latest Coding Results

Tested by third-parties

- Bitmovin: "*AV1 is able to outperform VP9 and even HEVC by up to 40%*"





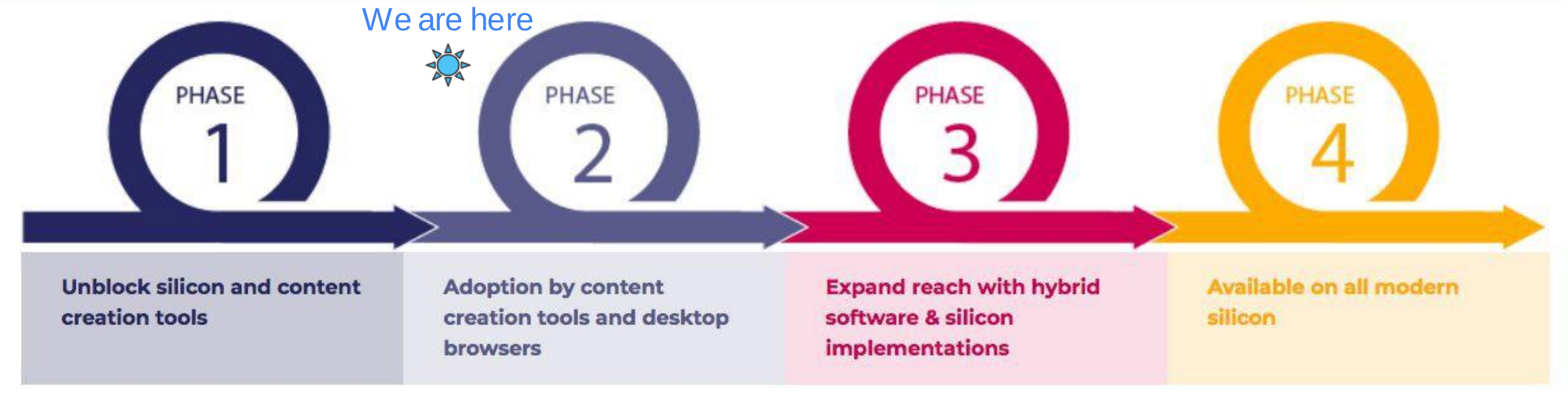
[[1]. https://bitmovin.com/av1-multi-codec-dash-dataset/](https://bitmovin.com/av1-multi-codec-dash-dataset/)

Outline

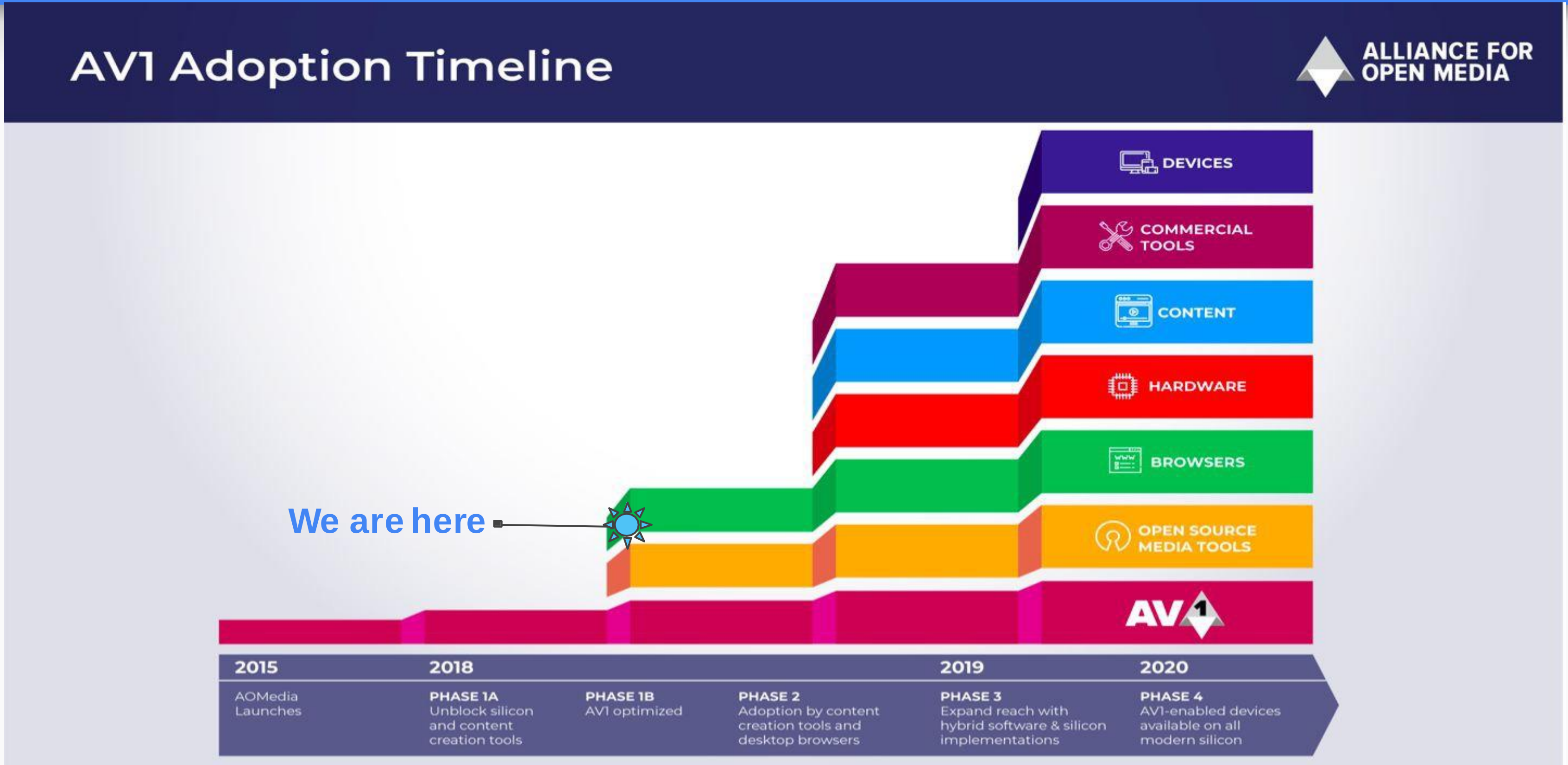
- Alliance for Open Media and AV1
- Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

AV1 Deployment
Phases

- Codec Deployment Phases
  - Bit-stream freeze is simply phase 1
- Productization (phase 2) efforts in full swing
  - Encoder/Decoder speed-ups, encoder advancements



AV1 Deployment
Adoption Timeline



AV1 Deployment
Rollout plans in Google

- AV1 Rollout plans in Google

1. Spec Published (Completed)
2. Chrome Support (Q4 2018)
3. WebRTC Integration (H2 2019)
4. Android Q Support (Q3 2019)
5. SoC Hardware (2020)

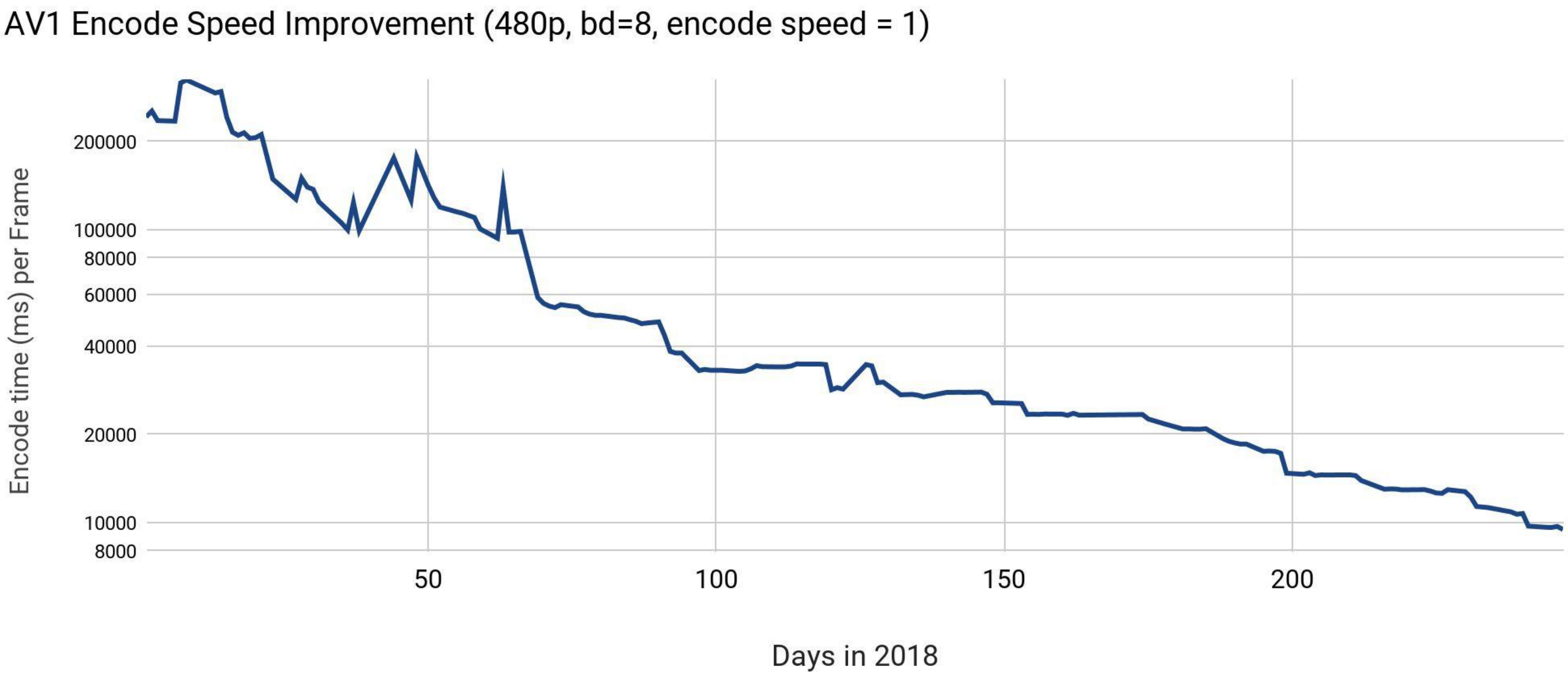
AV1 Deployment

Software speed-ups

- Early deployment of AV1 to depend on SW decode and encode
- SW Encoder speed-up
  - Encoder was very slow 6 months ago 1000x of VP9 cpu-used=0
  - Substantially faster today
    - AV1 cpu-used=0 (libaom): 40x of VP9 encode time at cpu-used=0 (libvpx)
    - AV1 cpu-used=1 (libaom): 16x of VP9 encode time at cpu-used=0 (libvpx) +1.5% PSNR
    - AV1 cpu-used=2 (libaom): 10x of VP9 encode time at cpu-used=0 (libvpx) +4.0% PSNR
  - Extensive use of machine learning based mode/partition decisions, early terminations, etc.
- SW Decoder speed-up
  - SIMD optimizations
  - Today 3.4x of VP9 decoder (single-threaded)
- Ongoing work …more speed-ups expected by end of the year.

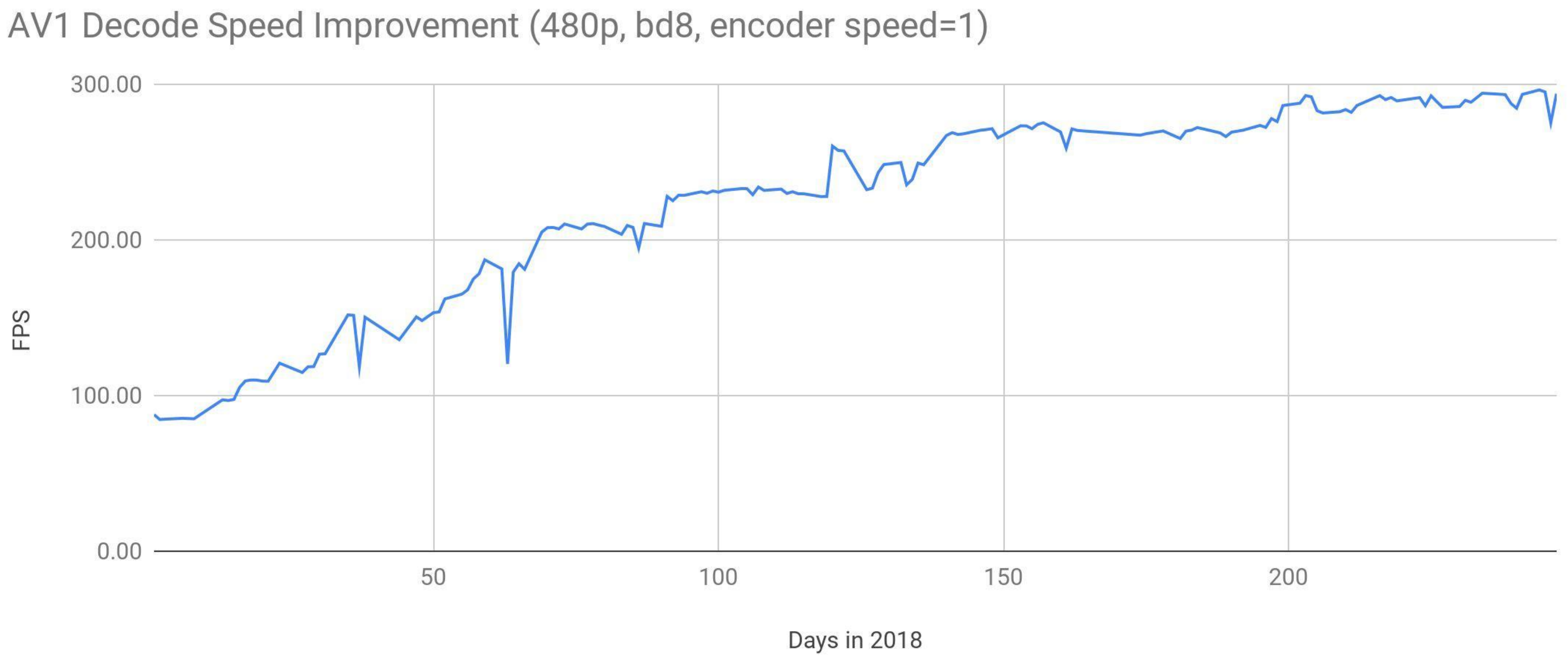
AV1 Deployment

Software Encoder speed progression



AV1 Deployment

Software Decoder Speed Progression



AV1 Deployment
Quality Improvements

- Improvement in frame management
  - Better hierarchical arrangement and quality control
  - 2+% gain beyond the bit-stream freeze.
- Adaptive quantization
  - Perceptual quality improvement - validated with perceptualtests
- Forward reference keyframe implementation
  - Substantial improvement in short key-frame interval scenarios for trick-play
  - 4+% improvement for 30-frame key frame interval.
- Coming:
  - Intelligent use of superresolution mode
    - Already shows gains at low bitrate for key-frames.

Outline

- Alliance for Open Media and AV1
- Coding Tools
- Latest Coding Results
- AV1 Deployment
- Beyond AV1

Beyond AV1
Trend

- AV1 is undoubtedly the most advanced video codec in the world today
  - ~30% better than VP9, and no slower than ~3x of VP9 decoding
  - ~20% better than HEVC
- May not remain the best in a couple of years - with VVC coming
  - A 2-horse race?
- Advancement in codec technology will continue
  - AV2

Beyond AV1
Promising areas

- Traditional - better prediction modes, transforms, etc.
- Motion
  - Non-translational motion models
- Restoration for coding
  - Throw away information that can be restored
  - Multiple restoration modes
  - Joint search among prediction, transform, restoration modes for a block
- Learned image compression
  - Can yield nonlinear transforms, can inform us on the best non-linear processing to compress and reconstruct
  - May be too complex for practical use.

Questions

Questions ?