Chroma from Luma Intra Prediction for NETVC

`draft-egge-netvc-cfl-01`

Nathan E. Egge
Luc N. Trudeau
David Barr

Mozilla and the Xiph.Org Foundation

IETF 100 - Singapore
2017 Nov 15

What is Chroma from Luma?

Intra prediction tool

No dependencies on other frames

Only available to chroma planes

Predicts chroma using coincident-reconstructed luma pixels

What’s New in -01?

Based on the chroma from luma proposal for AV1

Instead of Daala implementation

No longer relies on PVQ

Prediction is done in the spatial domain

Considers only AC contribution of reconstructed luma pixels

Spatial domain equivalent of shape prediction

Uses existing chroma DC prediction for DC contribution

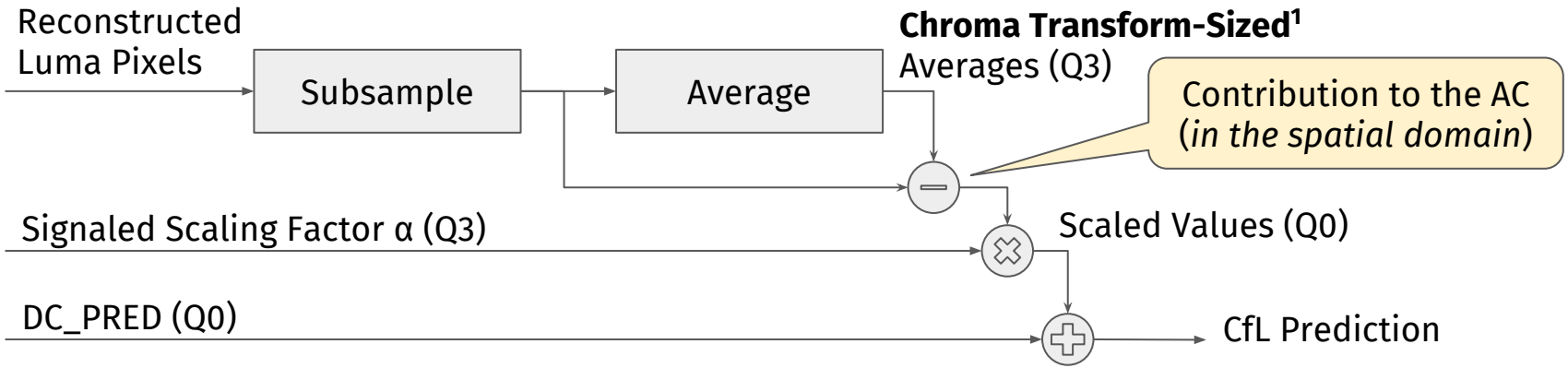
Available in AV1, requires no signaling and is more precise

What's different?

	LM Mode	Thor CfL[1]	Daala CfL[2]	**AV1 CfL**
Prediction domain	Spatial	Spatial	Frequency	**Spatial**
Bitstream signaling	No	No	Sign bit PVQ gain	**Signs + Index**
Activation mechanism	LM Mode (4x4, 8x8)	Threshold	Signaled	**CFL_PRED** **(UV-only mode)**
Requires PVQ	No	No	Yes	**No**
Encoder model fitting	Yes	Yes	Via PVQ	**Search**
Decoder model fitting	Yes	Yes	No	**No**

[1] `draft-midtskogen-netvc-chromapred-02`
[2] `draft-egge-netvc-cfl-00`

How Does it Work?



Why use Chroma DC_PRED?

β is the average chroma reference pixels for a block

$$\beta = \frac{\sum_i \sum_j C_{ij} - \alpha \sum_i \sum_j L'_{ij}}{M \times N}$$

0

AC contribution is zero mean (*it sums to 0*)

DC_PRED predicts the average value of a block

By computing the average of the neighboring pixels adjacent to the above and left borders of the block

No Signaling required

What are Scaling Factors (α_{Cb}, α_{Cr})?

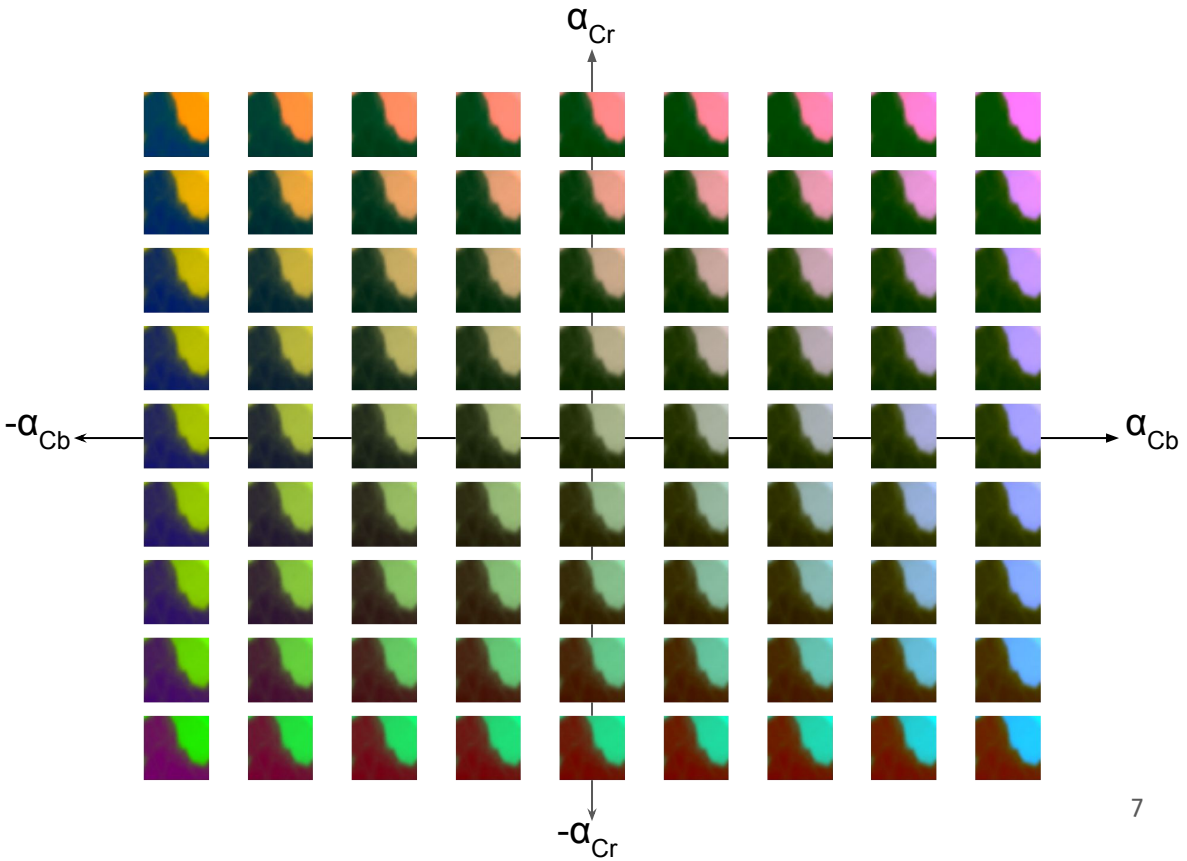
Scaling factors set the tone

Scaling factors are in Q3 and range from -2 to 2

Scaling factors are chosen by a rate-constraint search

$$\alpha = \underset{a \text{ in } A}{\operatorname{argmin}} \left(D(CfL(a)) + \lambda R(a)\right)$$

Scaling factors are signaled to the decoder



How are Scaling Factors Signaled?

A sign can either be [0, -, +]

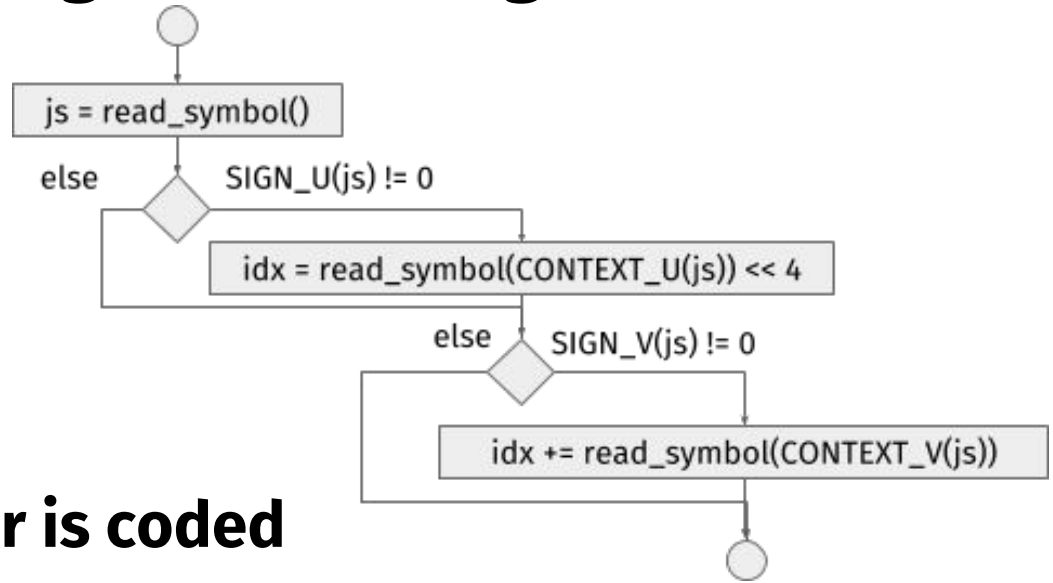
Signs are jointly coded

using an 8-value[1] CDF

Each non-zero scaling factor is coded

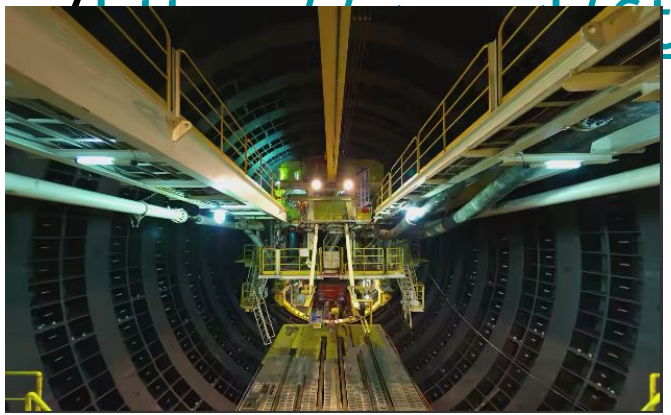
using a 16-value CDF (0,2]

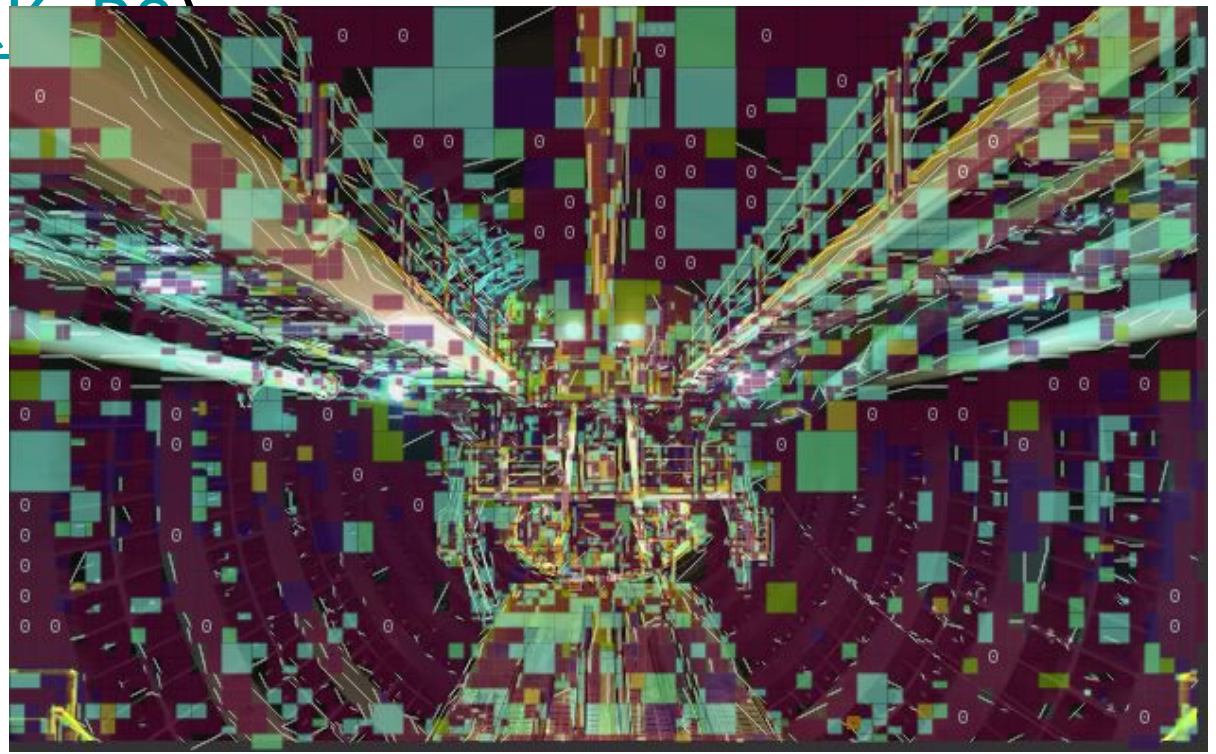
Joint sign used as context



[1]. (0,0) is not a valid code as it is equivalent to DC_PRED

UV Mode Selection Example





CFL_PRED 17%

DC_PRED 44.36%

TM_PRED 7.98%

SMOOTH_PRED 4.85%

Ohashi0806shield.y4m
QP = 55

Results (AWCY High Latency)

Subset1

	BD-Rate (%)						
	PSNR	PSNR-HVS	SSIM	**CIEDE2000[1]**	PSNR Cb	PSNR Cr	MS SSIM
Average	-0.46	-0.29	-0.33	**-4.65**	-12.99	-10.84	-0.32

Ref: https://arewecompressedyet.com/?job=master%402017-07-26T10%3A40%3A11.180Z&job=cfl-baseline%402017-07-29T00%3A04%3A47.130Z

Objective-1 fast

	BD-Rate (%)						
	PSNR	PSNR-HVS	SSIM	**CIEDE2000[1]**	PSNR Cb	PSNR Cr	MS SSIM
Average	-0.43	-0.42	-0.38	**-2.41**	-5.85	-5.51	-0.40
1080p	-0.32	-0.37	-0.28	**-2.52**	-6.80	-5.31	-0.31
1080p Screen	-1.82	-1.72	-1.71	**-8.22**	-17.76	-12.00	-1.75
360p	-0.15	-0.05	-0.10	**-0.80**	-2.17	-6.45	-0.04
720p	-0.12	-0.11	-0.07	**-0.52**	-1.08	-1.23	-0.12

Ref: https://arewecompressedyet.com/?job=master%402017-09-13&job=cfl-inter%402017-09-13T14%3A13%3A13.918Z

1. CIEDE2000 is the only metric that combines luma and chroma plane (*The distance measured is more perceptually uniform*)

Awesome for Gaming (Twitch dataset)

	BD-Rate (%)						
	PSNR	PSNR-HVS	SSIM	**CIEDE2000[1]**	PSNR Cb	PSNR Cr	MS SSIM
Average	-1.01	-0.93	-0.90	**-5.74**	-15.55	-9.88	-0.81

Ref: https://arewecompressedyet.com/?job=no-cfl-twitch-cpu2-60frames%402017-09-18T15%3A39%3A17.543Z&job=cfl-inter-twitch-cpu2-60frames%402017-09-18T15%3A40%3A24.181Z

Notable Mentions

	BD-Rate (%)						
	PSNR	PSNR-HVS	SSIM	**CIEDE2000[1]**	PSNR Cb	PSNR Cr	MS SSIM
Minecraft	-3.76	-3.13	-3.68	**-20.69**	-31.44	-25.54	-3.28
GTA V	-1.11	-1.11	-1.01	**-5.88**	-15.39	-5.57	-1.04
Starcraft	-1.41	-1.43	-1.38	**-4.15**	-6.18	-6.21	-1.43

Minecraft

MINECRAFT_10_120f.y4m

GTA V

GTAV_0_120f.y4m

Starcraft

STARCRAFT_10_120f.y4m