Thor update

High Efficiency, Moderate Complexity Video Codec using only RF IPR

(https://datatracker.ietf.org/ipr/2636/)

draft-fuldseth-netvc-thor-02

Steinar Midtskogen (Cisco)

IETF 96 – Berlin, DE – July 2016

Topics for this update

- Most important changes since IETF95/April 2016
  - Temporal skip motion vectors
  - Support for YUV 4:4:4 input
  - Improved chroma prediction

Temporal skip motion vectors

- Only used in B frames
- Replaces one spatial skip mv candidate with a candidate based on temporal mv
- Gives nearly half of the gain of interpolated reference frames (presented at IETF94), but eliminates the motion search and the overhead of storing an extra reference frame
  - Significantly reduces decoder complexity

Temporal skip motion vectors

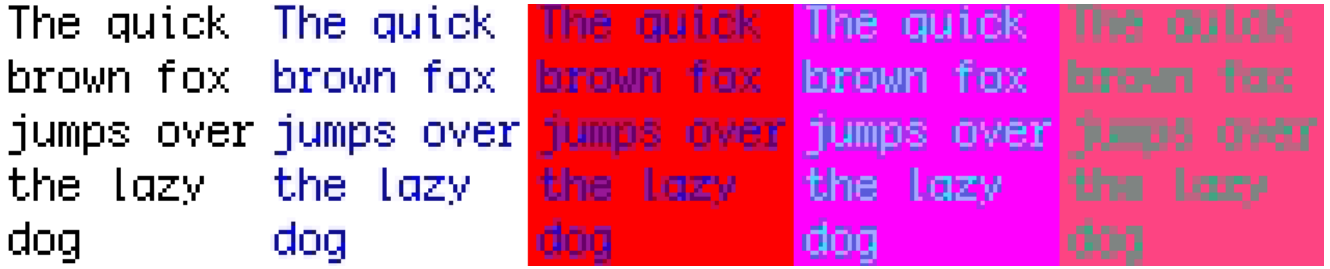
Sequence	Interpolated ref. frames	Temporal skip mv
Kimono	-4.6%	-0.9%
ParkScene	-3.1%	-0.9%
Cactus	-5.1%	-4.5%
BasketballDrive	-2.7%	0.6%
BQTerrace	-2.1%	-1.0%
ChangeSeats	-8.5%	-2.6%
HeadAndShoulder	-8.9%	-4.6%
TelePresence	-8.0%	-2.5%
WhiteBoard	-10.1%	-4.2%
FourPeople	-7.5%	-5.3%
Johnny	-7.7%	-3.2%
KristenAndSara	-8.2%	-4.6%
Average	**-6.4%**	**-2.8%**

Support for YUV 4:4:4

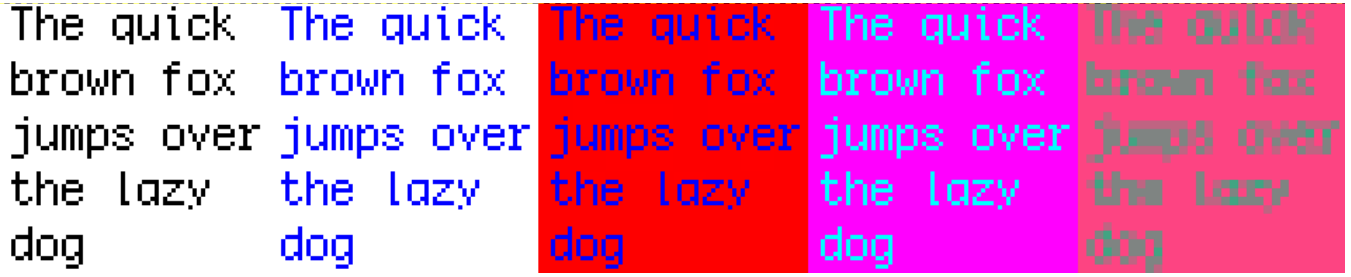
- No subsampling of chroma, and luma and chroma will use the same interpolation filter
- Significantly increases the number of bits spent for chroma (increases the total bitrate by ~5-15%)
- An experimental mode was added:
  - Support for optional subsampling of the residual at SB level
  - Adaptive upsampling in the reconstruction using a linear model to reconstruct chroma samples from luma samples
  - Significantly better than regular upsampling, but only regains about 25% of the subsampling loss objectively
- Abandoned in favour of a new chroma prediction

Support for YUV 4:4:4

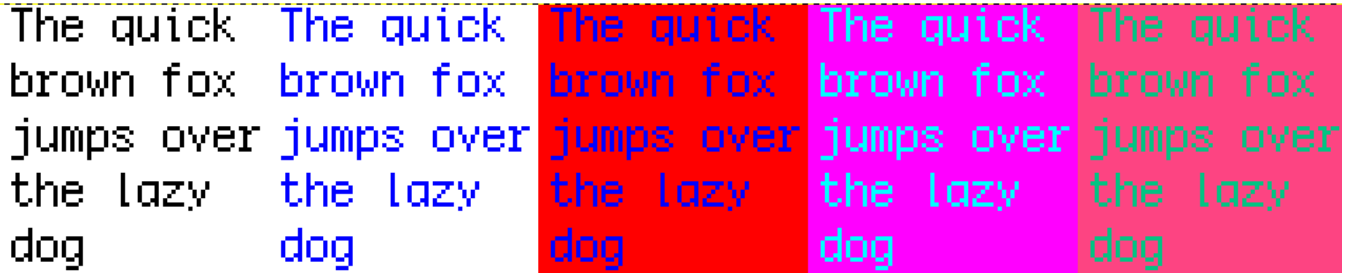
Regular 4:2:0:



Upsampling using reconstruction from luma:

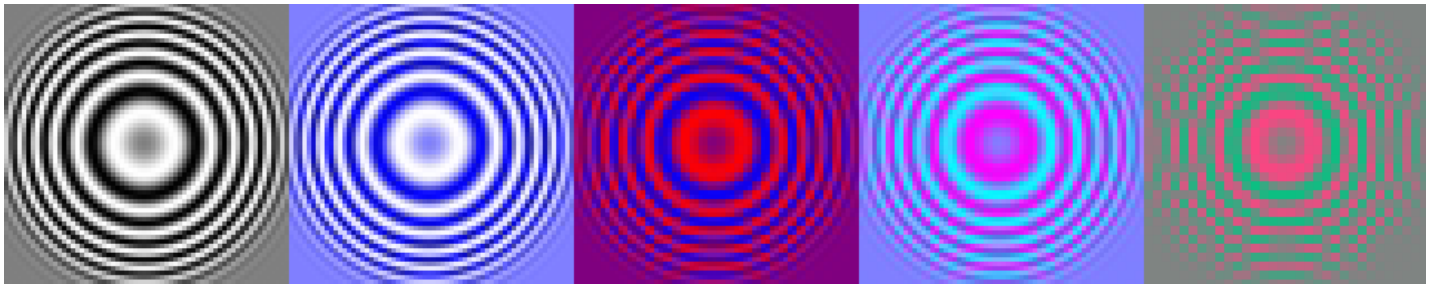


Original 4:4:4:

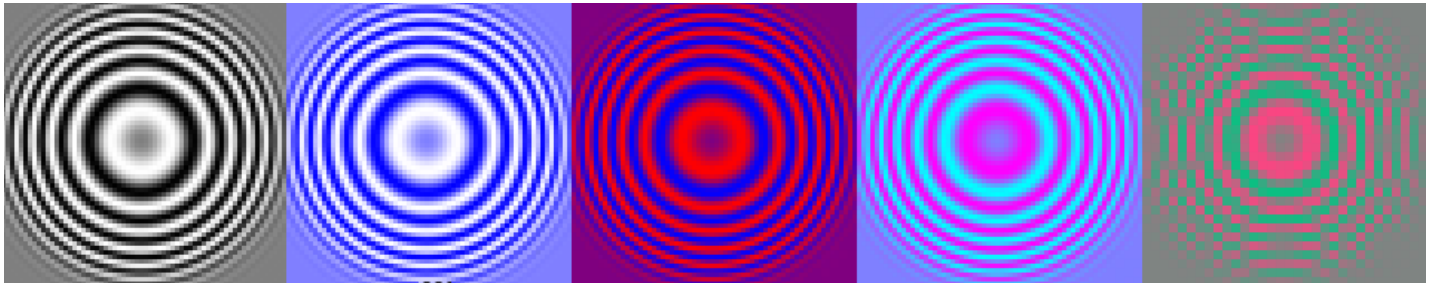


Support for YUV 4:4:4

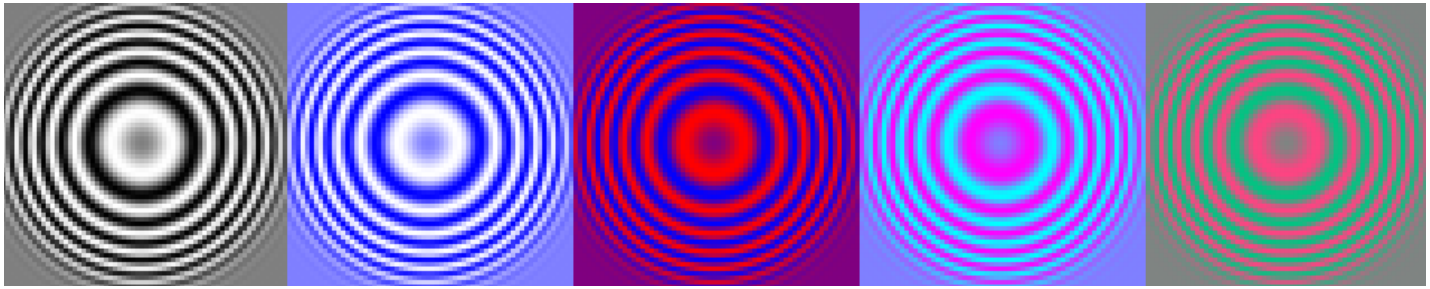
Regular 4:2:0:



Upsampling using reconstruction from luma:

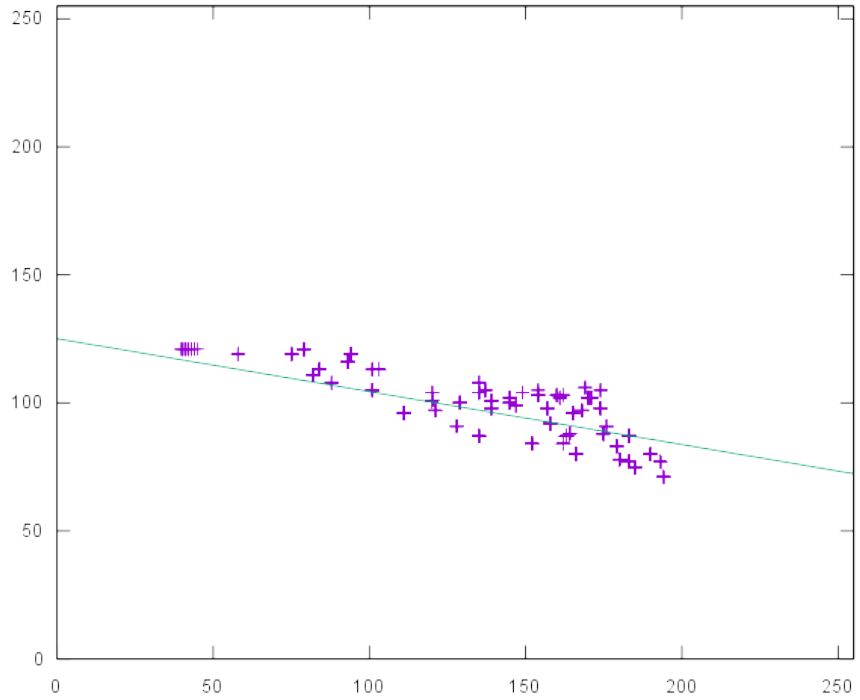


Original 4:4:4:



Improved chroma prediction

- As seen in the previous slides, chroma can be highly correlated with luma.
- The relationship is often nearly linear



Improved chroma prediction

- Finds a linear mapping from the predicted luma block to the initially predicted chroma block
- If certain criteria are satisfied, we'll use this mapping to get a new chroma prediction from the reconstructed luma block
- Informally: "Copy the reconstructed luma block and paint it with the colours of the initial prediction"
- Assumptions:
  - Colours can be identified by their luminosities
  - A reasonably strong linear mapping exists
  - Same luma-chroma correlation in the original block as in the initial prediction.

Improved chroma prediction

- Use linear regression to find *a* and *b*, so that we can predict chroma from using: $c_p = a*y_r + b$
- $y_{sum} = \sum_{i=1}^{N}\sum_{j=1}^{N} y_p$ $c_{sum} = \sum_{i=1}^{N}\sum_{j=1}^{N} c_p$ $yy_{sum} = \sum_{i=1}^{N}\sum_{j=1}^{N} y_p^2$ $cc_{sum} = \sum_{i=1}^{N}\sum_{j=1}^{N} c_p^2$ $yc_{sum} = \sum_{i=1}^{N}\sum_{j=1}^{N} y_p c_p$

  $ss_{yy} = yy_{sum} - (y_{sum}^2 >> 2*\log_2(N))$
  $ss_{cc} = cc_{sum} - (c_{sum}^2 >> 2*\log_2(N))$
  $ss_{yc} = yc_{sum} - (y_{sum}c_{sum} >> 2*\log_2(N))$
  $a = (ss_{yc} << 16) / ss_{yy}$
  $b = ((c_{sum} << 16) - a \times y_{sum}) >> 2*\log_2(N)$

- 64 bit arithmetic in intermediate stages, but *a* is clipped to $[-2^{23}, 2^{23}]$ and *b* to $[-2^{31}, 2^{31}-1]$, lower 16 bits as the fractional part.
- If 4:2:0, subsample predicted luma block and the resulting new chroma prediction

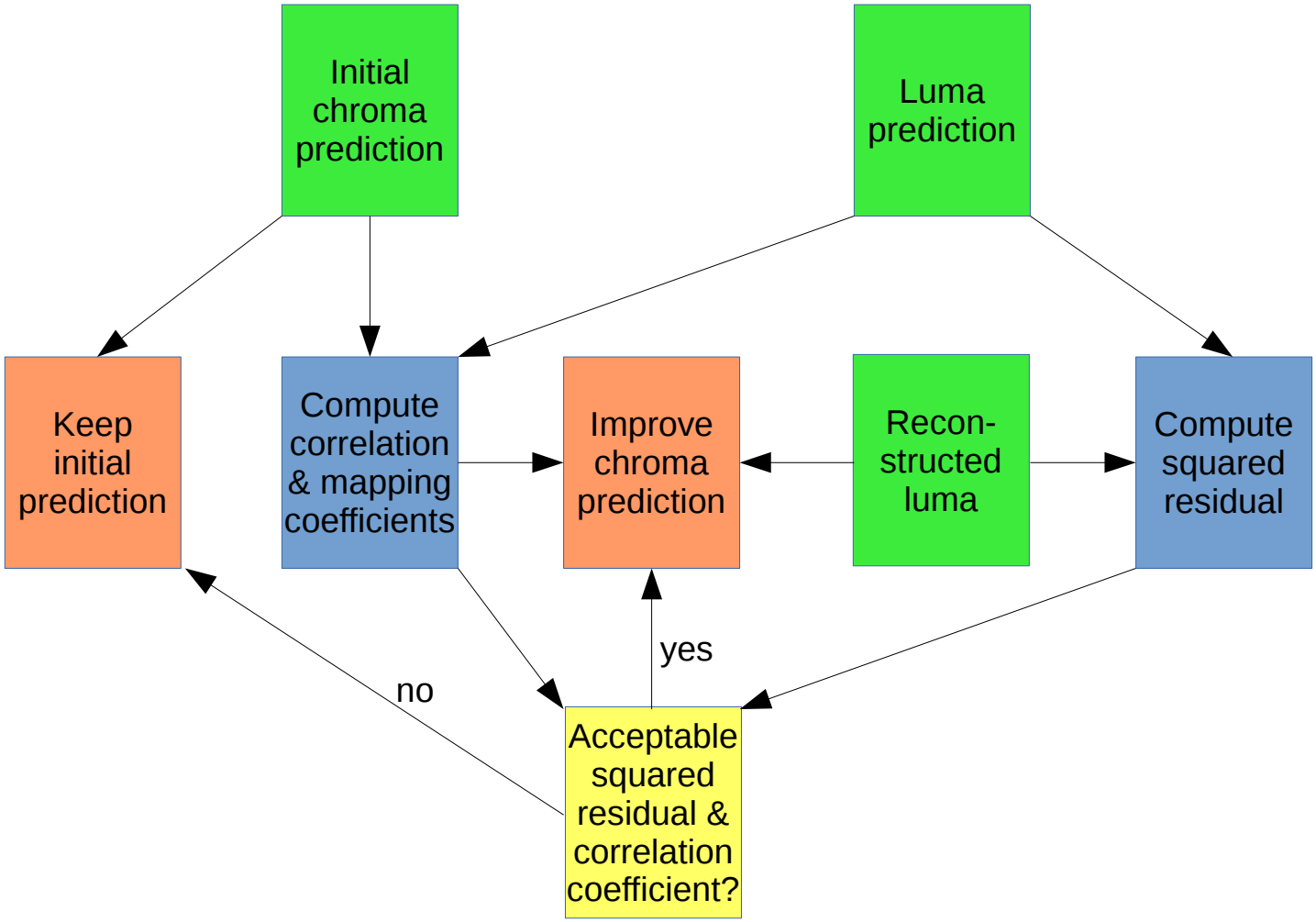
Improved chroma prediction

- Criteria for using the new prediction:
  - The luma prediction is poor. Use the average squared difference between the reconstructed and predicted luma samples (i.e. the residual) as a measure of quality where above 64 means low quality:

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(y_r(i,j)-y_p(i,j))^2}{N^2}>64$$

  - Sufficient luma-chroma correlation in the initial prediction. Require that the correlation coefficient is above 0.5: $ss_{yy} > 0$ and $2*ss_{yc}^2 > ss_{yy}ss_{cc}$.
- No signalling to select initial or improved prediction

Improved chroma prediction



Improved chroma prediction

Original U

Improved chroma prediction

Initial U prediction

Improved chroma prediction

Improved U prediction

Improved chroma prediction

(inter+intra)		4:4:4			4:2:0	
sequence	**Y**	**U**	**V**	**Y**	**U**	**V**
cad_waveform	-20.0%	-24.7%	-22.4%	-4.1%	-5.7%	-5.6%
pcb_layout	-7.3%	-11.1%	-10.1%	-1.1%	-1.8%	-1.6%
ppt_doc_xls	-19.6%	-8.9%	-9.0%	-0.3%	-1.2%	-0.8%
vc_doc_sharing	-3.2%	-6.5%	-10.1%	0.2%	-0.0%	-0.5%
web_browsing	-0.5%	-0.3%	-0.5%	-0.8%	-3.7%	-2.5%
wordEditing	-9.3%	-14.1%	-13.9%	-0.1%	-1.0%	-0.6%
park_joy	-1.4%	-7.4%	-1.2%	-0.8%	-9.9%	-1.4%
old_town_cross	-0.2%	-1.4%	-0.5%	-0.0%	-4.3%	-1.7%
Average	**-7.7%**	**-9.3%**	**-8.5%**	**-0.9%**	**-3.4%**	**-1.8%**

Improved chroma prediction

(intra coding)		4:4:4			4:2:0	
sequence	**Y**	**U**	**V**	**Y**	**U**	**V**
cad_waveform	-22.6%	-27.9%	-25.9%	-2.8%	-3.9%	-3.7%
pcb_layout	-18.9%	-27.1%	-20.5%	-1.1%	-1.8%	-1.6%
ppt_doc_xls	-6.4%	-12.4%	-13.5%	-0.4%	-0.2%	-0.8%
vc_doc_sharing	-5.7%	-11.9%	-11.9%	-0.1%	-2.9%	-0.6%
web_browsing	-1.4%	-1.8%	-1.8%	-0.6%	-1.0%	-1.2%
wordEditing	-12.9%	-16.3%	-13.5%	-0.3%	-5.4%	-1.2%
park_joy	-5.7%	-7.3%	-6.9%	-1.3%	-3.0%	-1.9%
old_town_cross	-1.9%	-2.4%	-2.4%	-0.2%	-4.9%	-1.7%
Average	**-9.4%**	**-13.4%**	**-12.1%**	**-0.8%**	**-2.8%**	**-1.7%**