

› J0030: DESCRIPTION OF 360° VIDEO CODING TECHNOLOGY PROPOSAL BY TNO

Alexandre Gabriel
Emmanuel Thomas

TNO innovation
for life

CONTENT

- › Technique description – Spatial scalability with single-layer coding tools
 - › Description of polyphase decomposition
 - › Applicability for 360 videos
 - › Configuration and scalability features

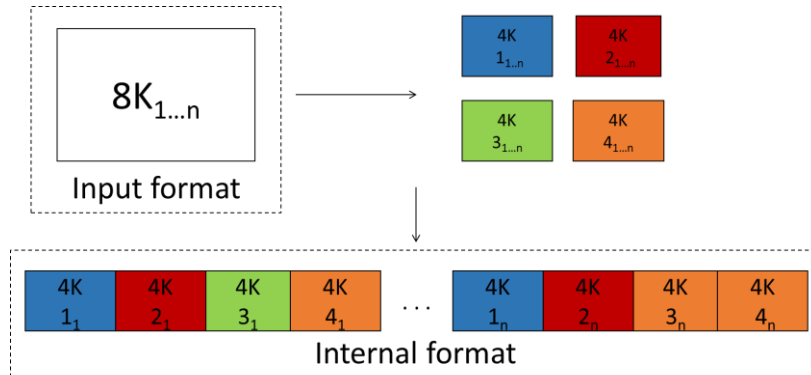
- › Results
 - › Objective quality
 - › Complexity analysis
 - › Scalability analysis
 - › Conclusion

› **TECHNIQUE DESCRIPTION**



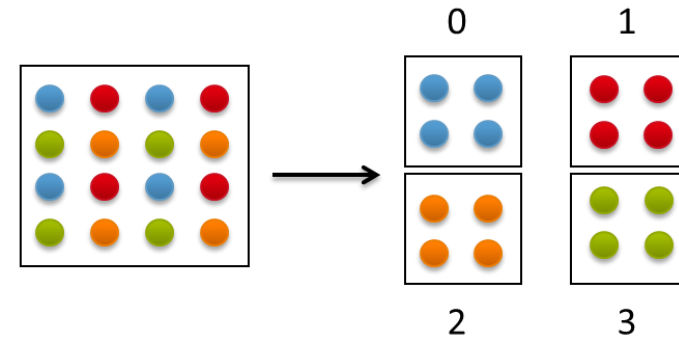
POLYPHASE DECOMPOSITION OVERVIEW

1) Each input frame is subsampled producing 4 different sub frames



Polyphase
decomposition

4 low-res
components



2) Each subframe is temporally multiplexed constituting the internal format with 4x more frames and $\frac{1}{4}$ less pixels than the input frame.

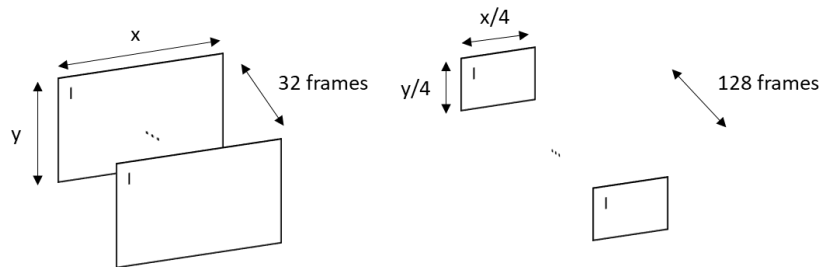
APPLICABILITY FOR 360 VIDEO

- › Would subsampling 8k 360 video cause aliasing?
 - › Subsampled 8K should not have aliasing artefacts as energy density is shifted towards low frequencies
 - › Deformation due to ERP projection means that there are less straight lines –which should also help with reducing risk of aliasing
- › In other technical domains, the increased pixel density makes obsolete the need of anti-aliasing filters. Examples include;
 - › Video game rendering (4K)
 - › Capture devices like camera

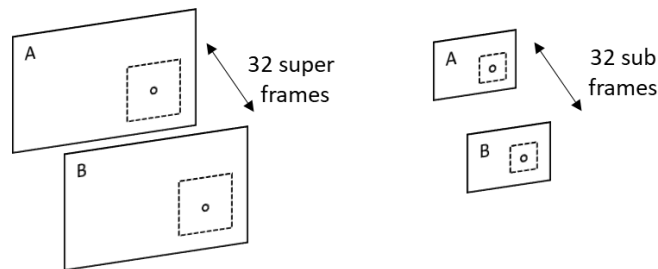


ENCODING CONFIGURATION

- random_access_main10.cfg was taken as base configuration scheme
- I interval was increased to every 128 frames as the frame rate was 4 times higher, matching the configuration of the anchor
- I frames were given a larger QP offset to compensate for the longer I interval
- Motion search range was set to 128 as opposed to 384 in the anchor. This is possible because object displacement is smaller between frames as the temporal distance is smaller.



Maximum temporal distance

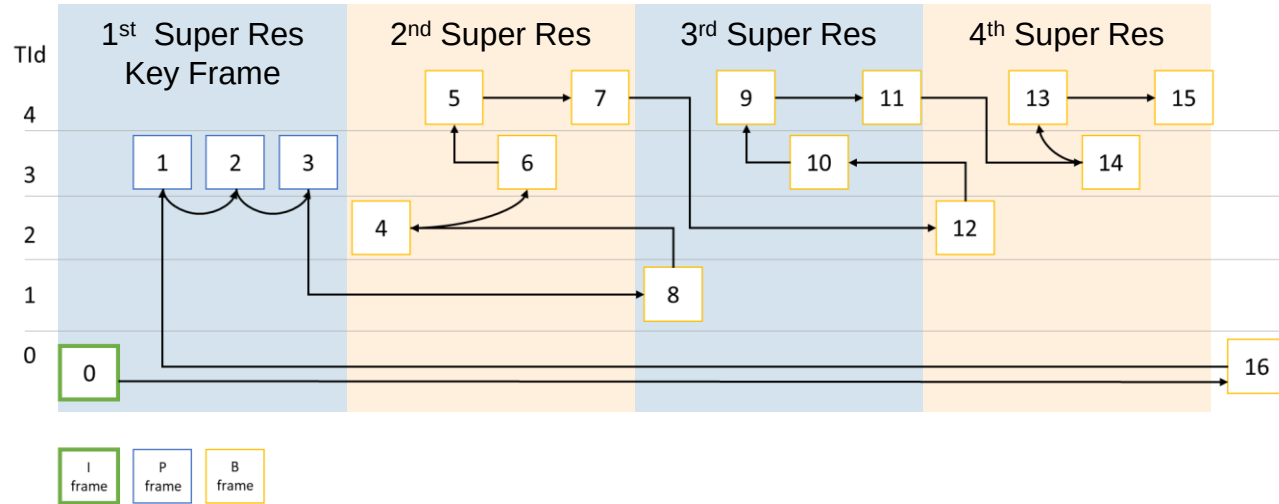


$$\frac{1}{30} * 32 = 1067 \text{ milliseconds}$$

$$\frac{1}{30 * 4} * 32 = 267 \text{ milliseconds}$$

GOP STRUCTURE AND SCALABILITY

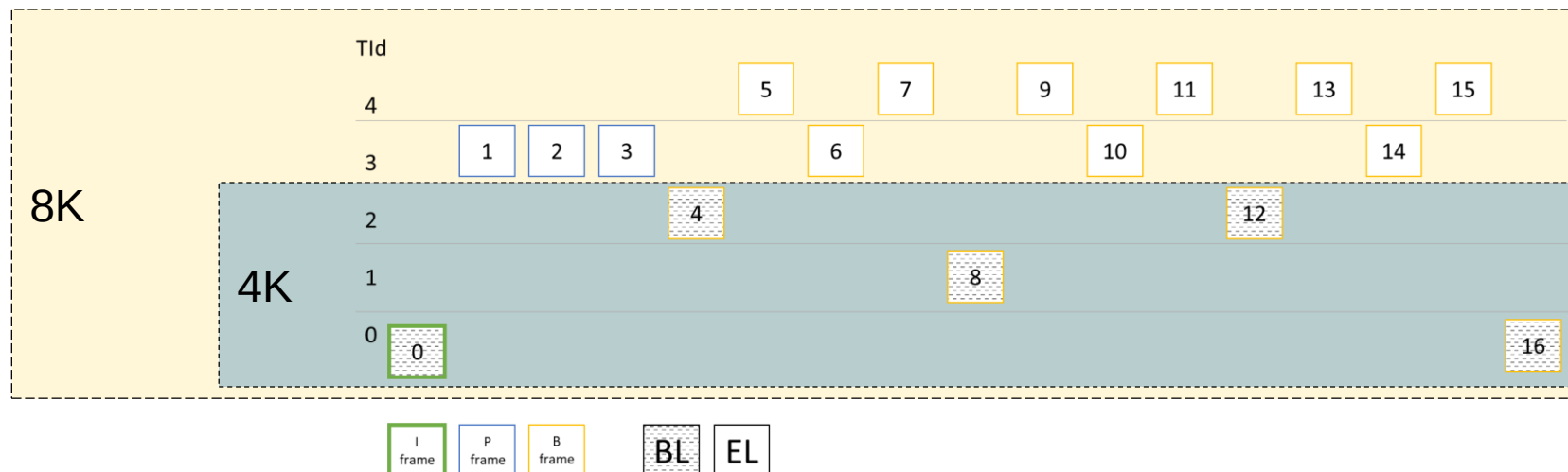
- › Frames 1, 2 and 3 were changed to P frames for decreasing temporal random access (only need to decode frame 16 and no longer 4 and 8)
- › Decode and encode order other than this is the same to normal encoding configuration



Random access interval calculation:

$$t = \frac{\text{frames to be decoded}}{\text{FPS}} = \frac{128 + f_{16} + f_1 + f_2 + f_3}{30 * 4} = \frac{132}{120} = 1.1 \text{ seconds}$$

GOP STRUCTURE AND SCALABILITY

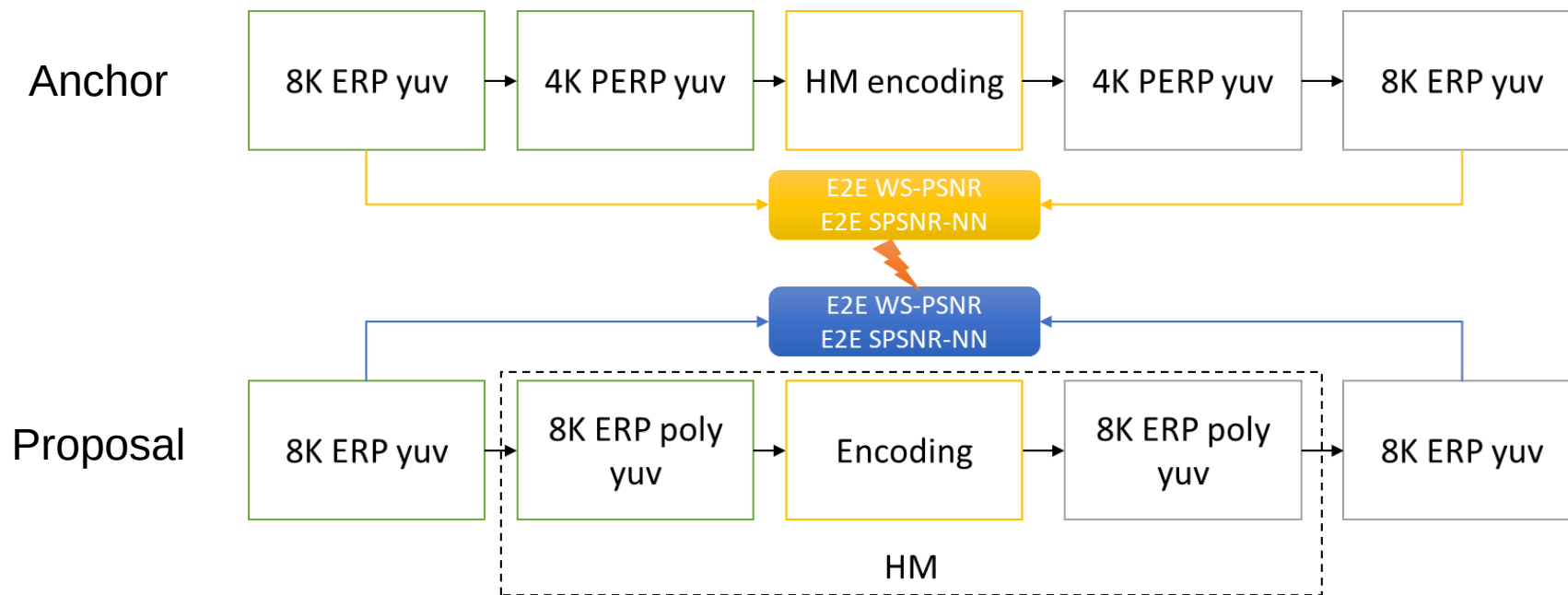


- › Frames with multiple of 4 constitute a 4K stream (i.e. 0, 4, 8, 12, 16 ...)
- › By decoding only the frames with TId < 3 a 4K bitstream is output
- › By decoding all frames, 8K output resolution can be achieved with temporal demultiplexing

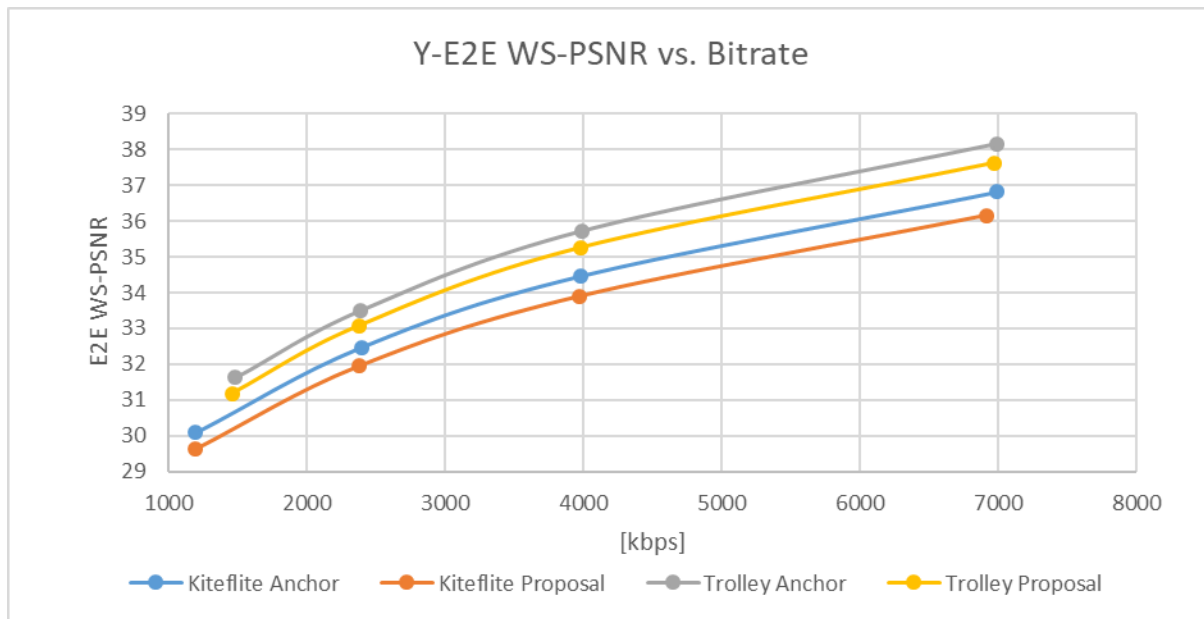
› RESULTS



ANCHOR VS PROPOSAL



OBJECTIVE RESULTS



OBJECTIVE RESULTS - TABLE

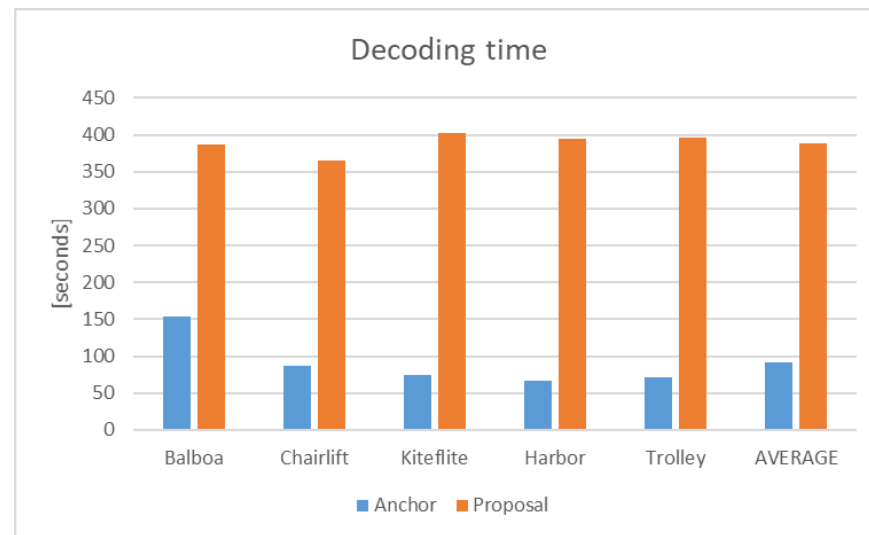
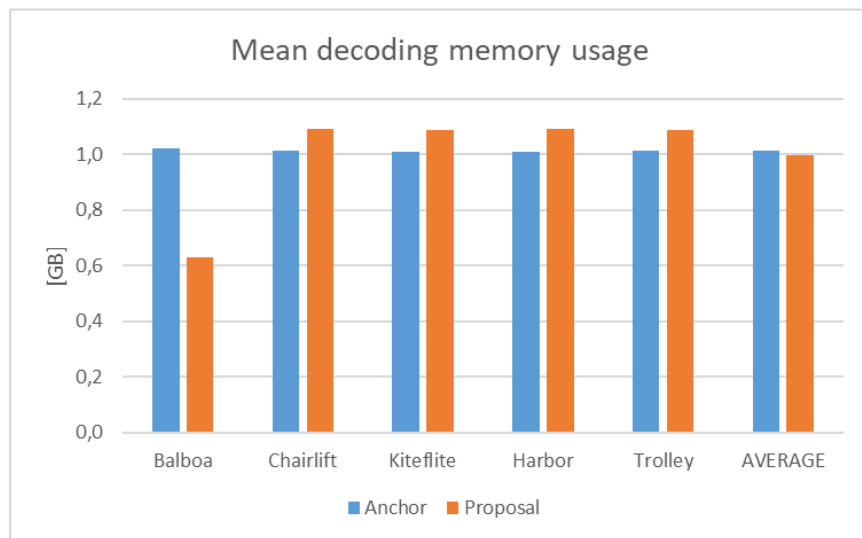
› Comparison with HM

- › Proposal is on average 0,52 dB lower in Y component and on average 0,26 lower in U/V components

	E2E WS-PSNR			E2E SPSNR-NN		
	BD-rate(Y)	BD-rate(U)	BD-rate(V)	BD-rate(Y)	BD-rate(U)	BD-rate(V)
Balboa	26.0%	14.1%	7.5%	26.4%	14.5%	7.8%
Chairlift	16.2%	11.4%	8.5%	16.3%	11.4%	8.6%
KiteFlite	15.1%	11.4%	12.4%	15.1%	11.5%	12.7%
Harbor	14.5%	12.9%	12.3%	14.4%	13.0%	12.3%
Trolley	11.4%	11.5%	12.9%	11.4%	11.5%	12.8%
Average 360	16.6%	12.3%	10.7%	16.7%	12.4%	10.8%

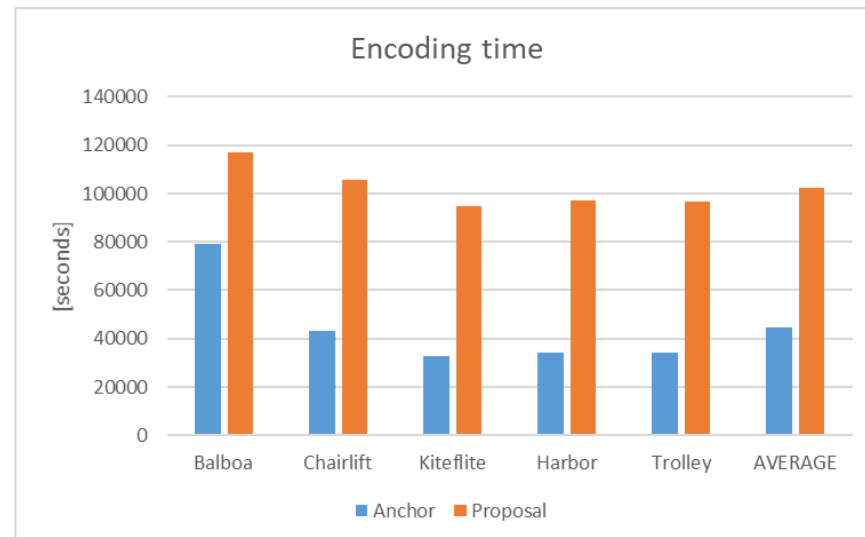
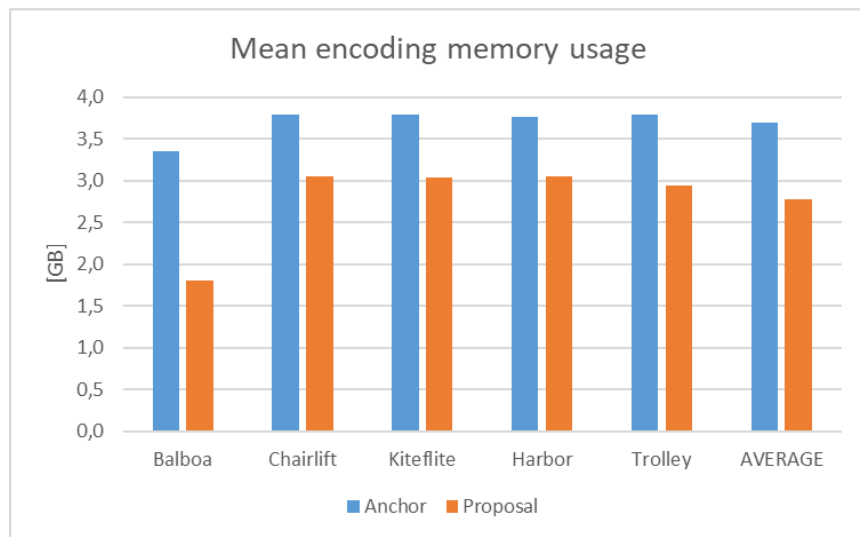
COMPLEXITY ANALYSIS – DECODING

- › Memory usage is on par with anchors as frame resolution is the same – Balboa is an exception because the resolution is lower than the anchor's
- › Decoding time is higher as 4 times the frames are decoded in each sequence



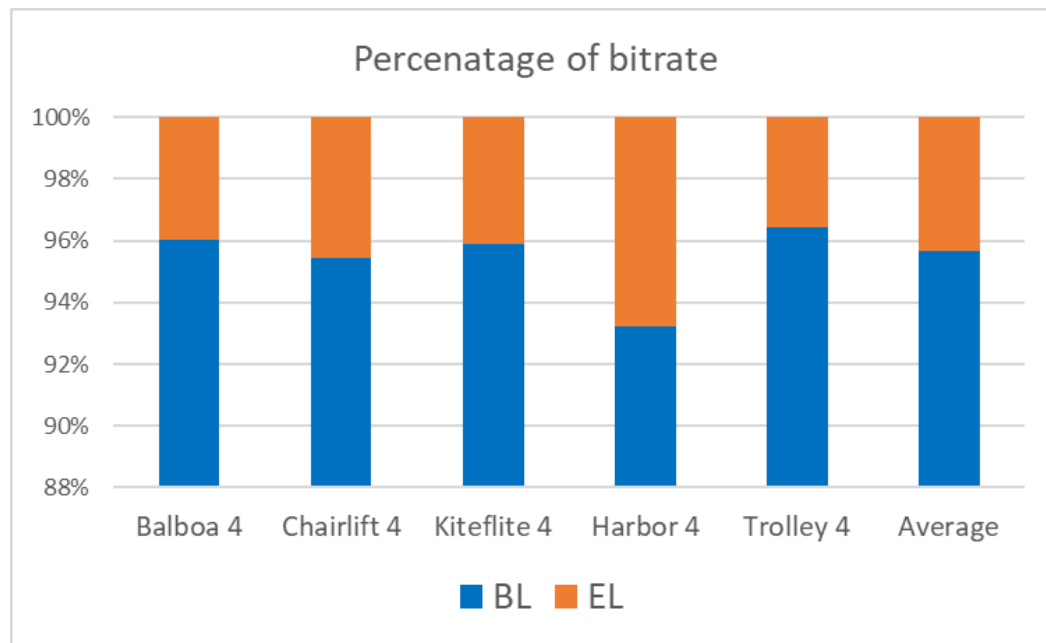
COMPLEXITY ANALYSIS – ENCODING

- › Memory usage is slightly lower
- › Encoding time is roughly 2,5 times higher – though not 4 fold as the motion search range is also smaller



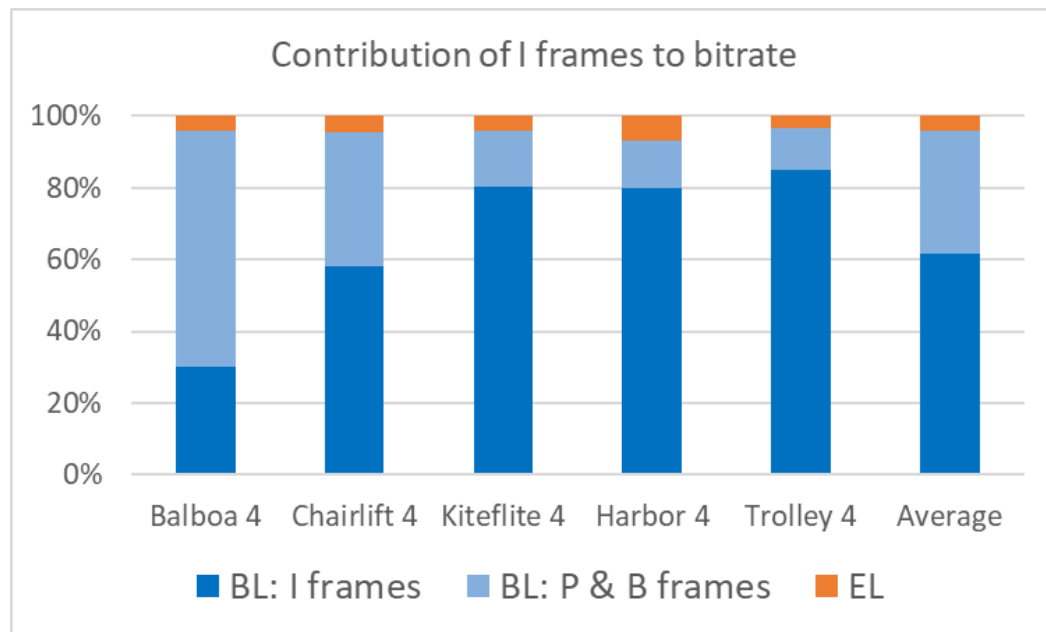
SCALABILITY ANALYSIS

- › This is the current bitrate distribution
 - › Blue – EL % of bitrate
 - › Orange – EL % of bitrate
 - › On average EL is around 5%
- › Focus on video compression is reason for disbalanced ratio



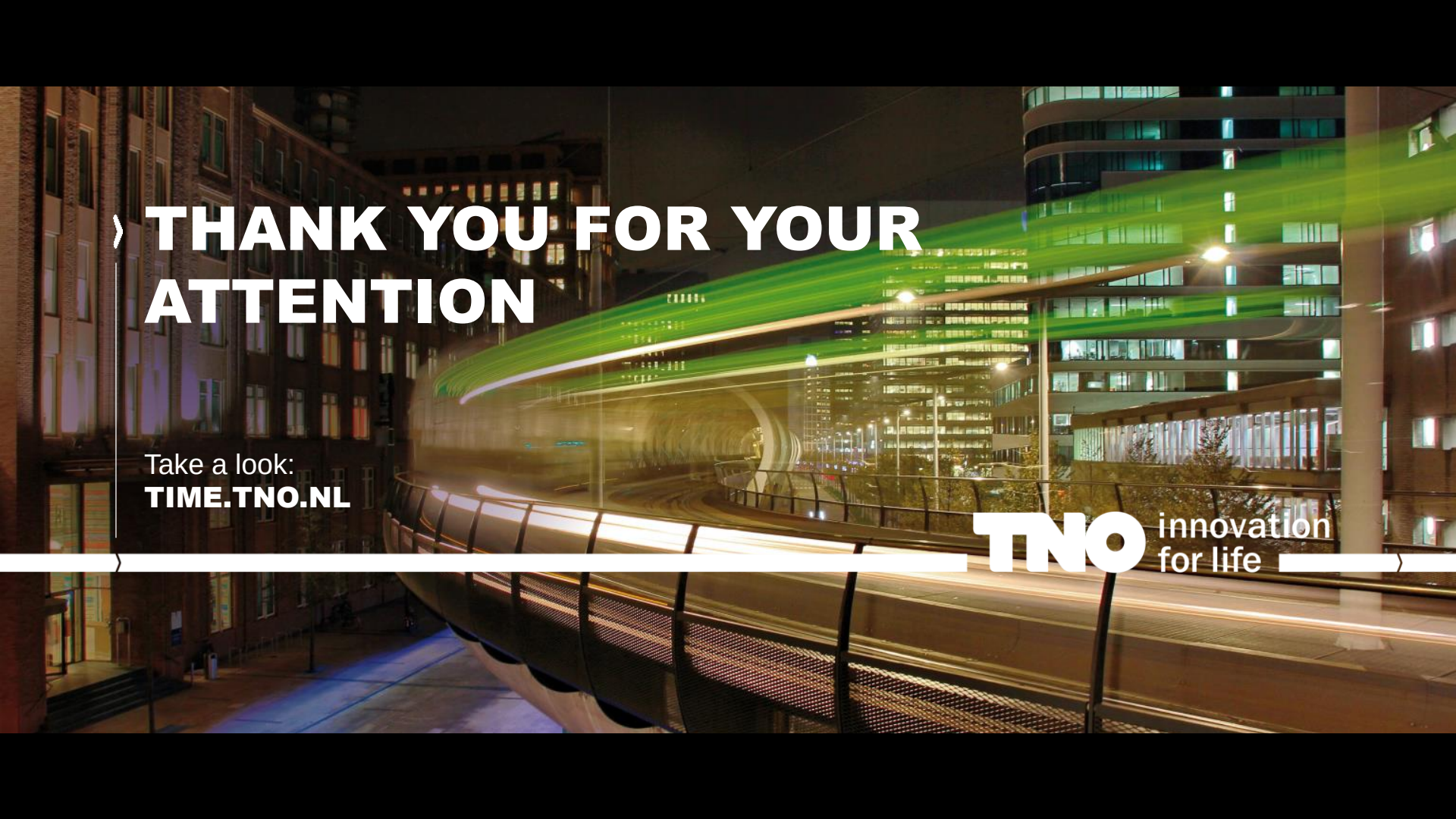
SCALABILITY ANALYSIS

- › I frames are reason for large discrepancy – they take up most bitrate in static sequences
 - › Dark Blue – I frames from BL
 - › Light Blue – P & B frames from BL
 - › Orange – Enhancement Layer
- › Less focus on video compression should lead to different distribution, e.g. with different QPs
- › I frames take up a huge part of the bitstream
 - › Higher QP makes this effect more noticeable than usual



CONCLUSION

- › What is proposed?
 - › Spatial scalability design at a cost of 16% BD rate increase for the luminance in the 360 video category.
 - › Complexity matching encoding/decoding single-layer complexity, reusing single-layer coding tools
- › How?
 - › By polyphase decomposing the original input sequences and encoding all input samples (e.g. 8K) in temporally multiplexed sub frames
- › Why the 360 category?
 - › 8K input resolution ensures the absence of aliasing artifact in sub frames, no dramatic bitrate increase, nor visual artifacts (compared to some 4K test sequences, see CfE input JVET-G0026)

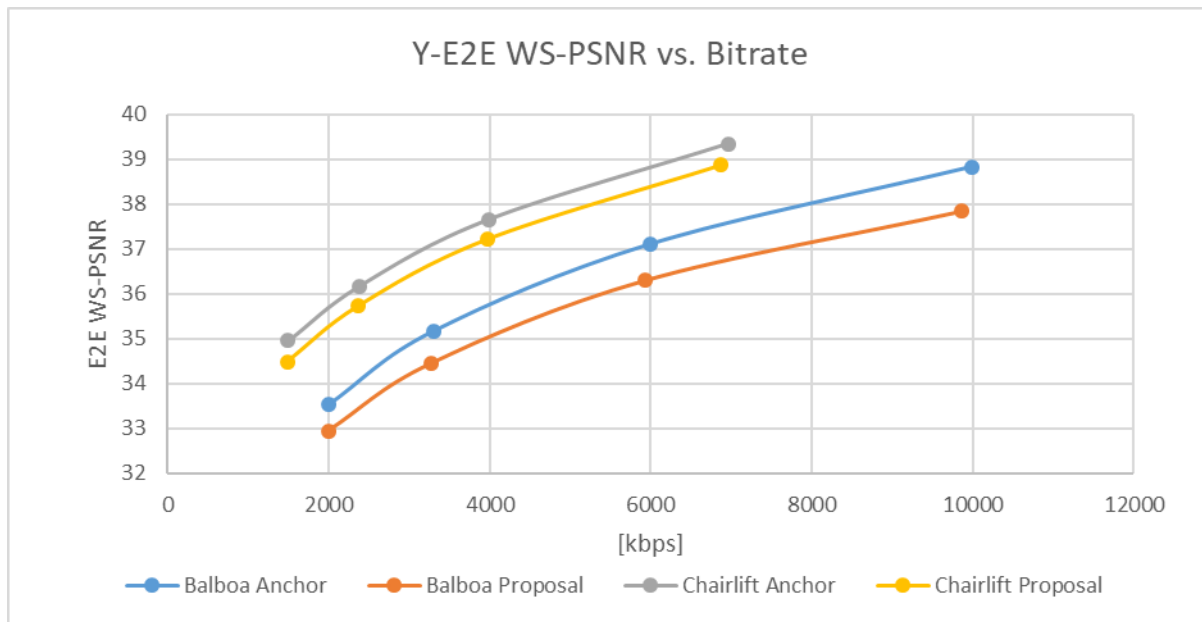


› **THANK YOU FOR YOUR
ATTENTION**

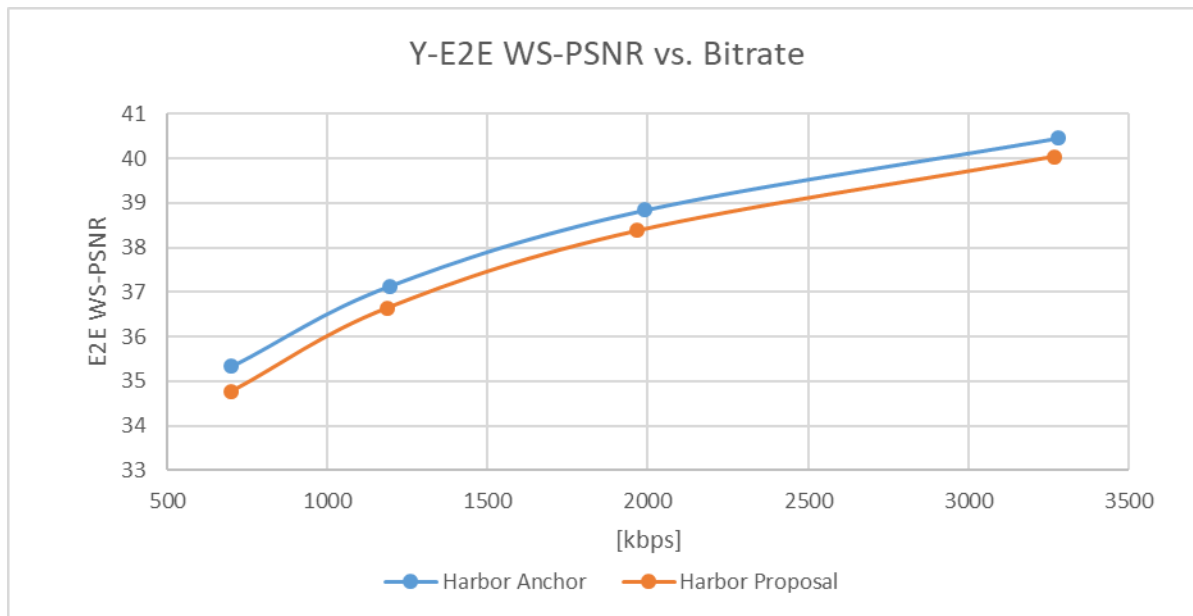
Take a look:
TIME.TNO.NL

TNO innovation
for life

OBJECTIVE RESULTS



OBJECTIVE RESULTS



EVOLUTION OF ALIASING IN OTHER DOMAINS

- › **Camera manufacturing** – Disparition of low-pass filtering in front of sensors
 - › “the pixel density of sensors has become so high that there are fewer and fewer patterns that we are likely to photograph that have a high enough frequency to cause moiré patterning” (<http://www.techradar.com/news/photography-video-capture/cameras/non-low-pass-filters-explained-goodbye-to-anti-aliasing-1160397>)
 - › So far cameras had a low pass filtering in front of the sensor to avoid moiré effect.
 - › But many 4K cameras no longer have such a low-pass filtering.



EVOLUTION OF ALIASING IN OTHER DOMAINS

- › **Video games** – Anti-aliasing for 4K and higher resolutions yields limited improvements
 - › “The need for antialiasing is diminished at ultra-high resolutions like 4K, especially on smaller displays where the pixels are more densely packed, but it will still improve image quality.”
(<http://www.pcworld.com/article/2990895/hardware/the-state-of-4k-gaming-what-you-need-to-know-from-pricing-to-performance-needs.html>)
 - › Anti-aliasing techniques have been essential when rendering video games.
 - › But with higher output resolution and higher pixel density, artefacts due to aliasing occurs less and if present are averaged by the human eye at a certain viewing distance.

