INFORMATION ON AND ANALYSIS OF THE VVC ENCODERS USED IN THE SDR UHD VERIFICATION TEST

Christian Helmrich
Fraunhofer HHI

Input document JVET-**T0103**

Questions Arising from Results of SDR UHD Verification Test

- VVC verification test quantifies the **subjective** Bjøntegaard delta-rate gain over HEVC [1]
- Two VVC compliant encoders were evaluated in the UHD SDR verification test:
  - **VTM**: VVC reference software, https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM
  - **VVenC**: Fraunhofer's alternative fast encoder, https://github.com/fraunhoferhhi/vvenc
- VTM 9.0 encoder is VVC draft 9 compliant, VVenC 0.1.0 encoder is draft 10 compliant [2]
- **Two questions** arise from the encoders tested and the results reported in JVET-T0097 [3]:
  - Is there a **difference in visual quality** between VTM 9 and 10 compliant bitstreams?
  - Which **objective metrics** can model the outcome of the subjective evaluation well?

[1]: JVET-S2009, [2]: JVET-S2001, [3]: JVET-T0097

Difference Between VTM 9 and 10 Compliant Bitstreams

- Two observations were made:
  - VTM encoder version 9 or 10 has **no influence** on visual quality, only on bitstream size
  - Only the choice of software, **VTM** or **VVenC**, affects the visual quality gain over HEVC

- Therefore, we can conclude:
  - In terms of YUVs presented to the viewing session participants, **VTM 9.0 = VTM 10.0**
  - Observed differences between the VTM encoder and VVenC are **not due** to the VVC draft text versioning ➔ direct comparison is possible

Objective Metrics to Model Subjective Encoder Ranking 1

- Subjective differences between VTM encoder and VVenC are observed and reported [3]:
  - MOS/rate curve for VVenC above MOS/rate curve for VTM enc. for ≈**90%** of the range
- VVenC uses perceptual optimization for improved visual quality based on **XPSNR metric**
  - Spatio-temporal enhancements of previous publications: H0047, K0206, L0181, M0091
  - Encoder-only deblocking optimization of M0428, GOP-based temporal filter of O0549
  - Speedups + multithreading → ≈**110 times faster than VTM** encoder on UHD content!
- A 5-point BD-rate evaluation, using several "visual distortion" measures, was conducted
  - XPSNR, SSIM, MS-SSIM and VMAF ("4K" and "neg" models); PSNR-YUV for reference

Objective Metrics to Model Subjective Encoder Ranking 2

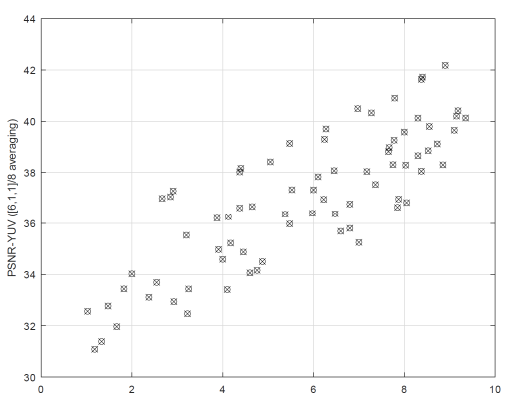
MOS	VTM 10.0		Over HEVC (HM 16.22)				
		PSNR YUV	XPSNR	VMAF-4K	VMAF-neg	MS-SSIM	SSIM
-42.75%	Overall	-35.66%	-35.22%	-41.02%	-39.23%	-33.64%	-32.39%

MOS	VVenC 0.1		Over HEVC (HM 16.22)				
		PSNR YUV	XPSNR	VMAF-4K	VMAF-neg	MS-SSIM	SSIM
-49.28%	Overall	-29.38%	-36.58%	-32.09%	-31.86%	-34.24%	-37.28%

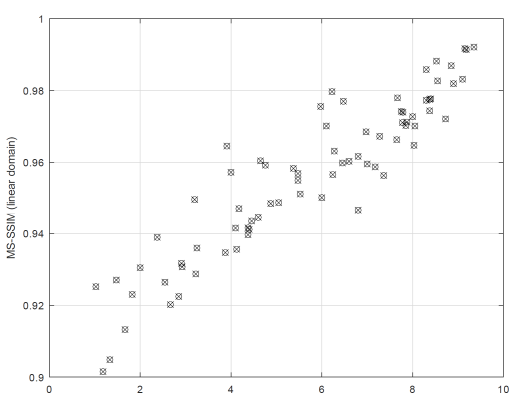
- SSIM, MS-SSIM and XPSNR **predict subjective ranking** between VTM encoder and VVenC
  - PSNR and VMAF (both models) don't! Overall, all ≈**10 – 15%** lower than MOS BD-R
  - Open-source XPSNR plug-in for FFmpeg at https://github.com/fraunhoferhhi/xpsnr

Objective Metrics to Model Subjective Encoder Ranking 3

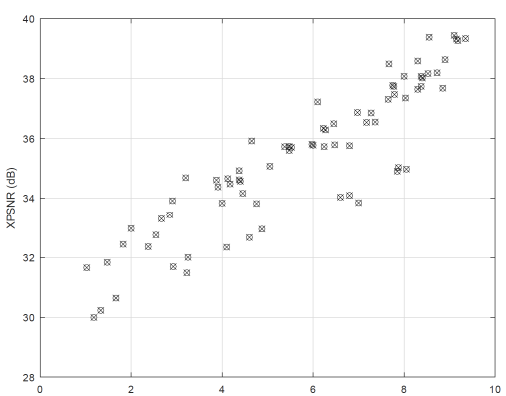
- Some more statistical analysis of verification test results (75 pairs) using VQA measures:
  - Pearson correlation with MOS values: **MS-SSIM 0.92**, XPSNR 0.90, PSNR 0.83, rest ≤ 0.8
  - Spearman correlation w/ MOS values: **MS-SSIM 0.92**, XPSNR 0.90, SSIM 0.81, rest < 0.81
- MS-SSIM or XPSNR seem to be useful as "subjective" alternatives to PSNR in JVET's work



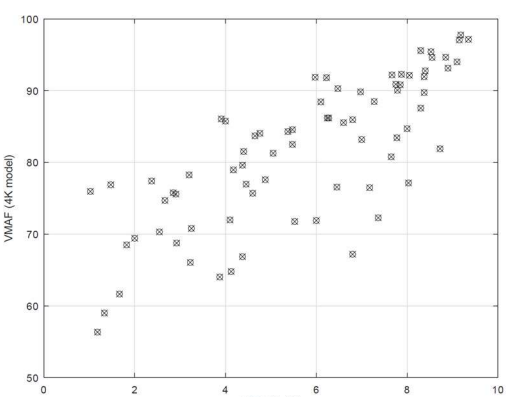
PSNR-YUV



MS-SSIM



XPSNR



VMAF-4K