
Status update on the Content of the Large Scale Database New Content and Pooling Strategies

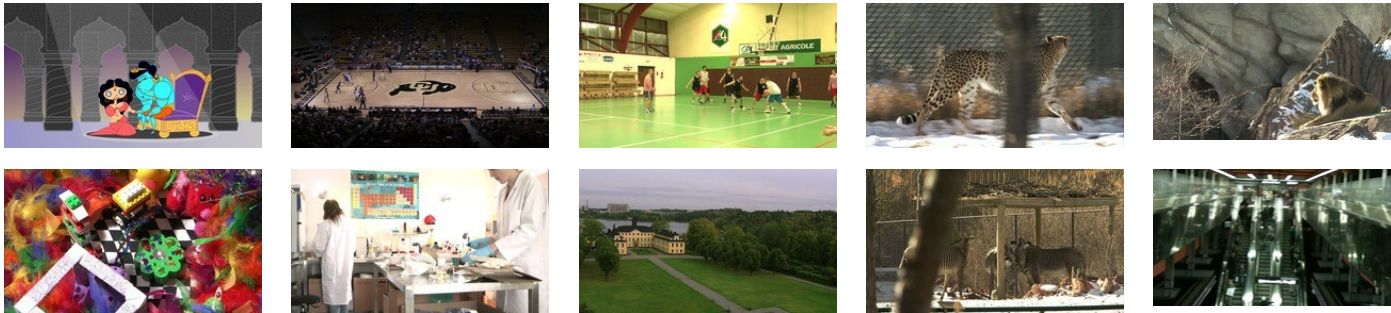
Enrico Masala
Politecnico di Torino, Italy
enrico.masala@polito.it

VQEG JEG-Hybrid session in Madrid, March 2018




Context

- “Base version” of JEG-Hybrid Large Scale Database
 - 59,520 HEVC-encoded video sequences (1,920 HRCs)
 - 10 sources, 250 frames each, 25 fps
 - 3 resolutions: 1920x1080, 1280x720, 960x544
(details in references, already presented in previous meetings)
 - Distortion due to encoding
 - Distortions due to encoding + data (packet) loss (~500,000 samples)



Recent Additions

- 3 new sources (sports content by Sky – from CDVL) 
- At 1920x1080 resolution, 25 fps (for homogeneity) – 1,920 HRCs
- 8 metrics available: PSNR, SSIM, VIF, VQM, PVQM, MS-SSIM, VMAF (model v. 0.6.0 and 0.6.1)
 - Also for the new content
 - $PSNR_{sf}$, $SSIM_{sf}$, $MS-SSIM_{sf}$, VIF_{sf} metrics also mapped to a 1-5 scale with sigmoidal fitting (parameters fitted on VQEG HDTV database and its MOS scores, thanks to Marcus Barkowsky)



New Temporal Pooling Methods

- For the “frame based” metrics PSNR, SSIM, MS-SSIM, VIF, VMAF:
 - Different pooling methods
 - Harmonic mean (as proposed in the VMAF software)
 - L(n) norm
 - 10,25,75,90-th percentile
 - Others
 - Median
 - Geometric mean
 - Indicators of variation over time
 - Average of consecutive absolute differences (Total_variation, as proposed in the VMAF software)
 - Standard deviation

Results

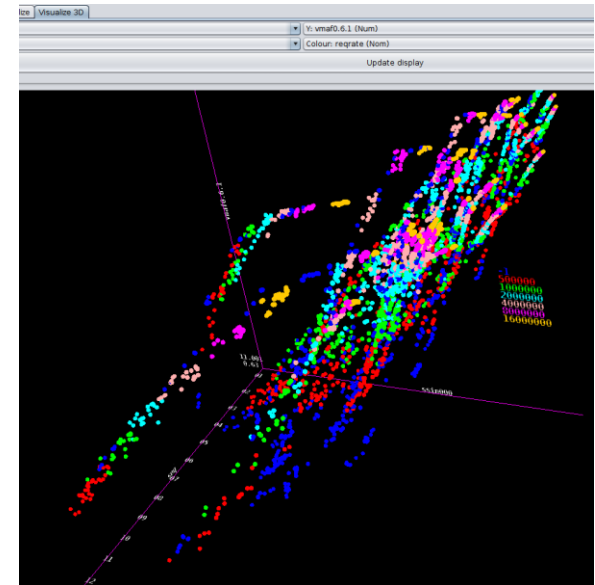
■ New content

- Sport content behavior (soccer in particular) does not seem to deviate too much from other content in the database
- The “confetti” sequence deviates significantly, probably due to the encoding difficulty for the small and fast moving parts
 - Fixed QP produces almost double bitrate w.r.t. most difficult content in the rest of the DB



“Confetti” sequence

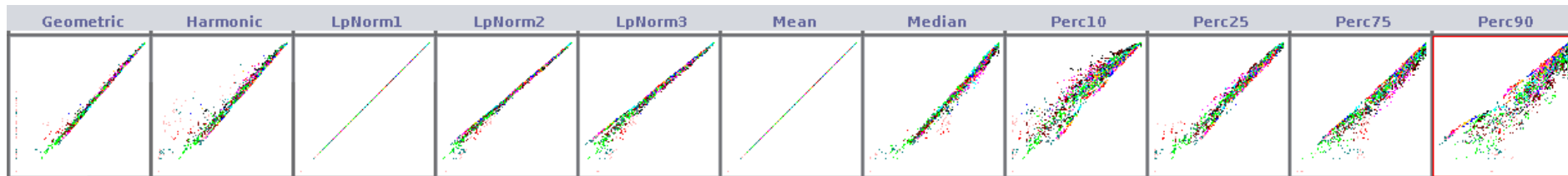
- VMAF vs SSIM: VMAF seems not to consider “small” artifacts so important (in accordance with src08)
- vs MS-SSIM: less evident effect
- vs VIF: some effect, not so strong as for src08



Using Weka for visualization

Different Pooling Methods

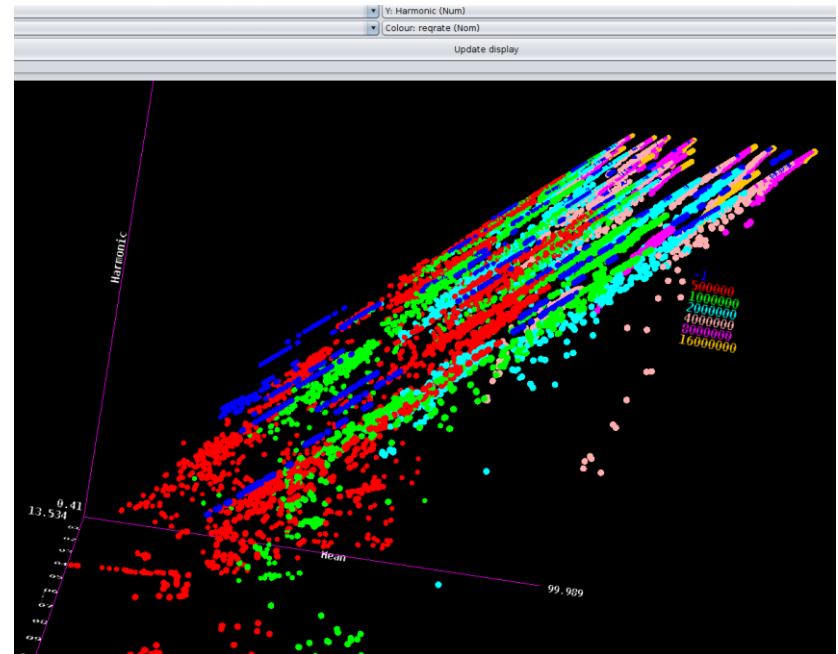
■ VMAF 0.6.1 X: different methods Y: average



- Most differences at low bitrate (500Kbit/s @1080p)
- More frequent on more complex sequences (src05 + sport ones)
- NO effect in case of fixed QP
 - The different results after pooling capture the change of metric values over time due to the rate-control

Different Pooling Methods

- Better seen in 3D to distinguish source peculiarities
- Harmonic mean: can be seen easily (best of all)
- Geometric mean, Median: much less visibility
- 10% Percentile value: low dependency on content



Blue color = fixed QP

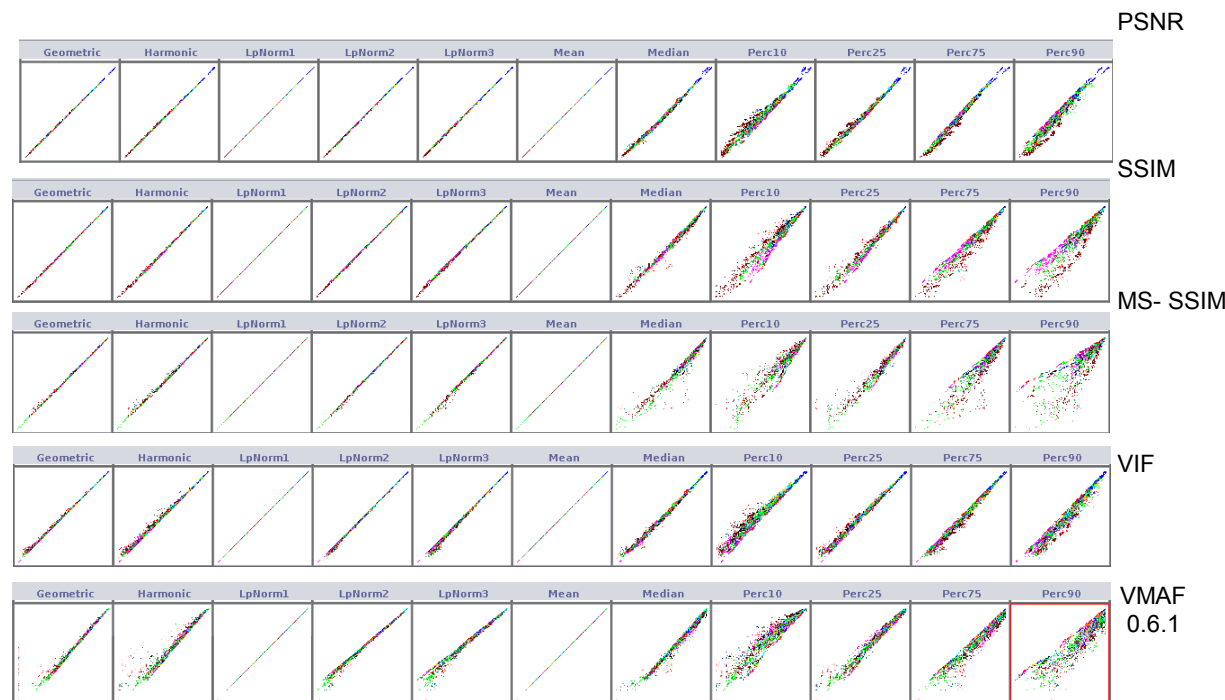
Different Pooling Methods: Metrics Comparison

□ PSNR, SSIM, MS-SSIM

- No significant differences
- Maybe an indication that some metrics are not sensitive enough to quality variations?

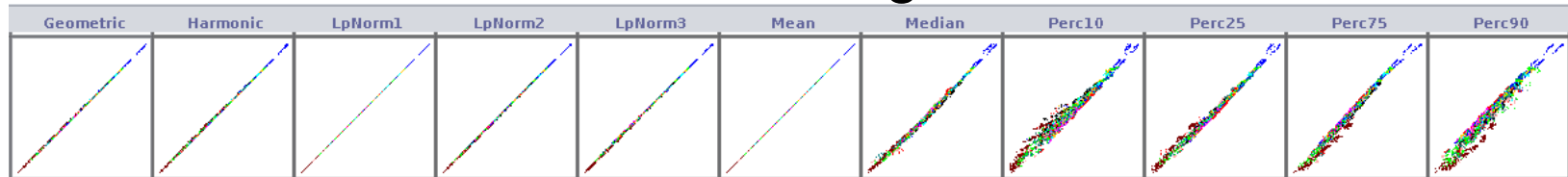
□ VIF

- Some variations as VMAF

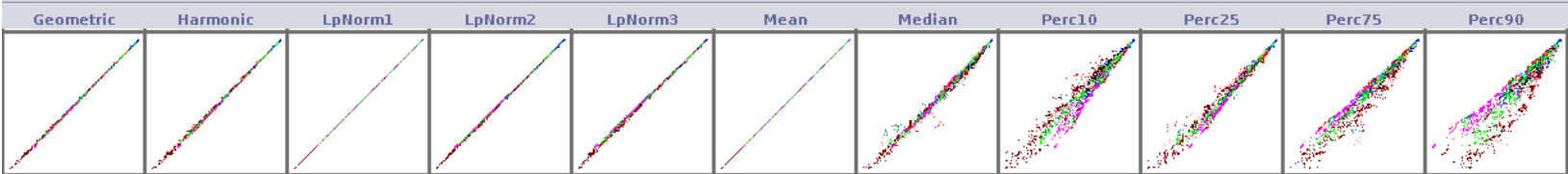


Different Pooling Methods: Metrics Comparison

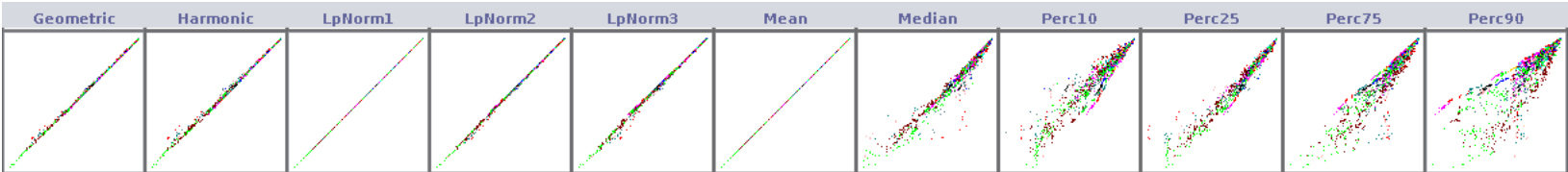
PSNR X: different methods Y: average



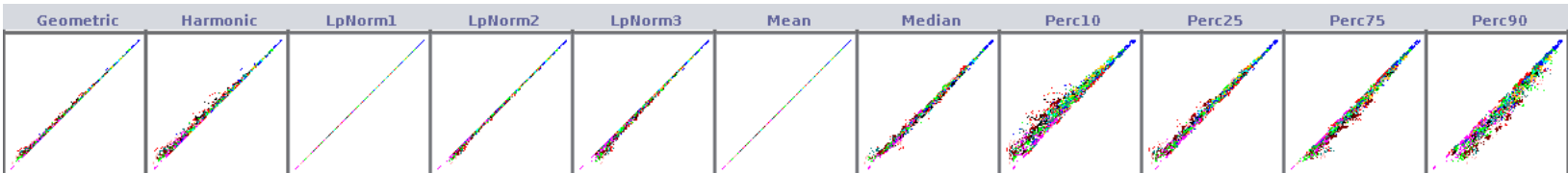
SSIM



MS-SSIM

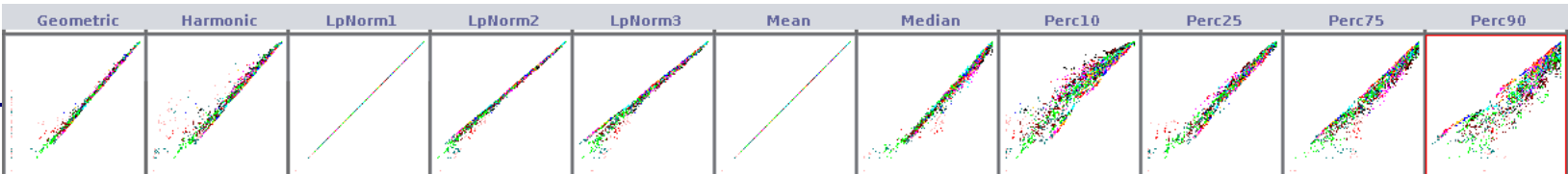


VIF



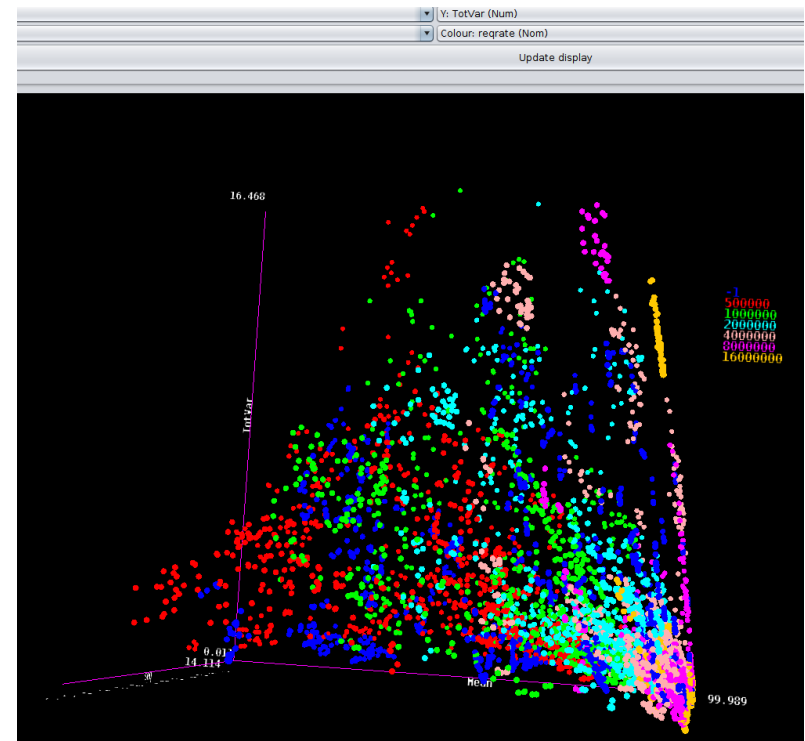
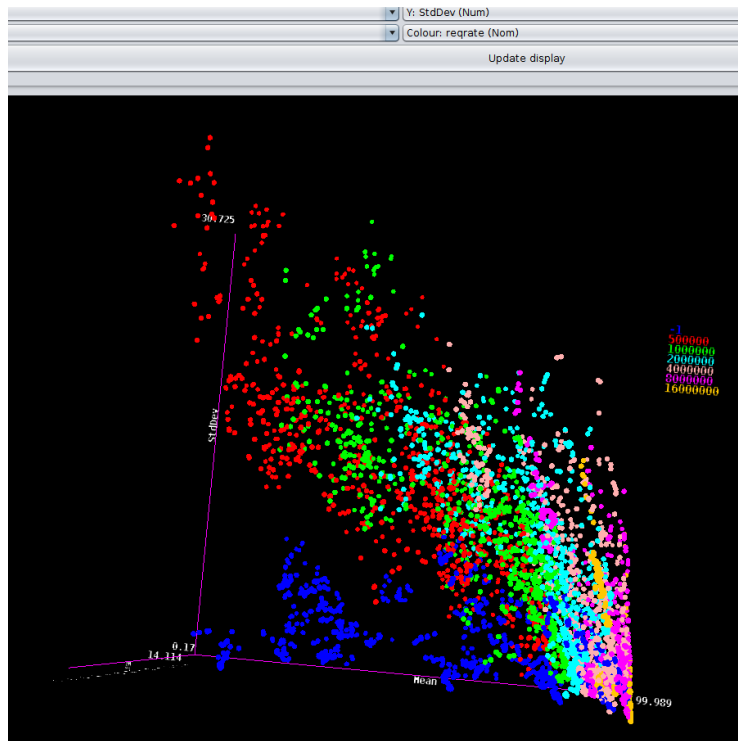
VMAF

0.6.1



Results: Variation over Time

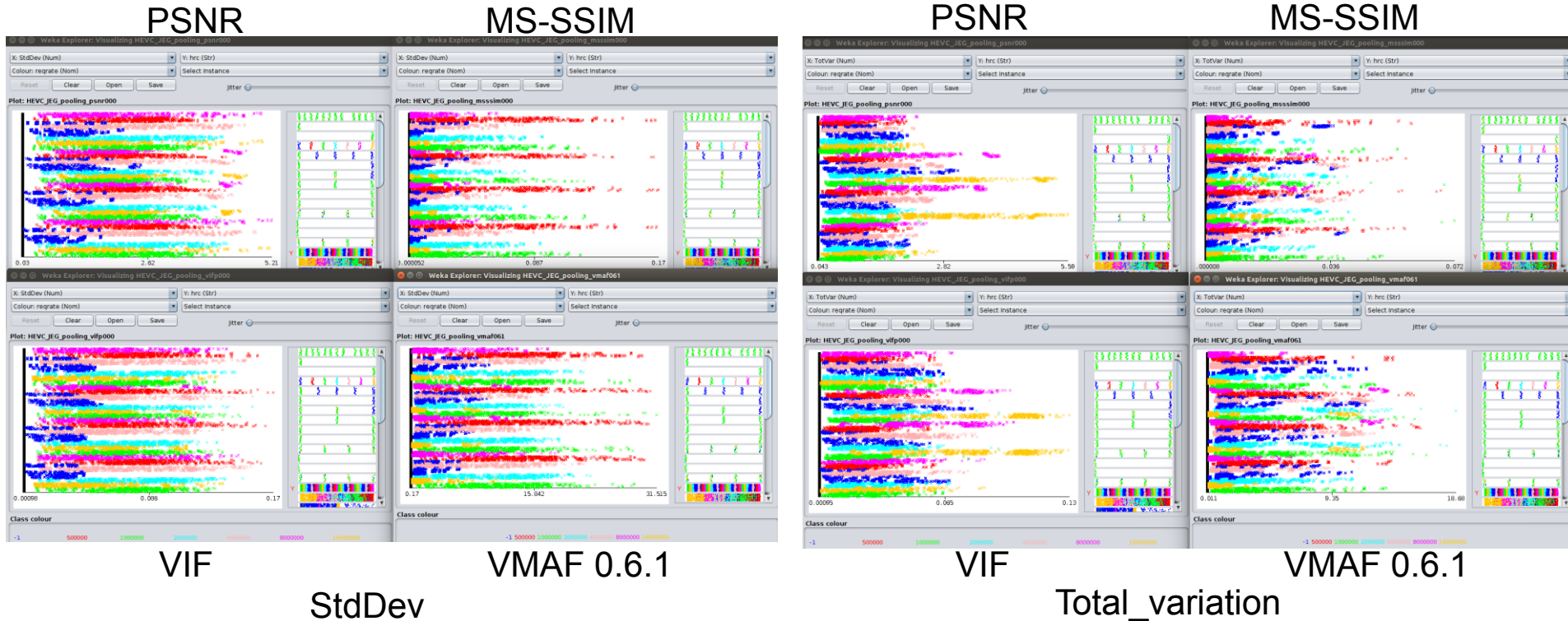
- No clear indications apart from less variability with fixed QP (blue color) especially with stddev (left: Y=stddev, right: Y=total_variation) vs X=mean Z=seq



Blue = fixed QP

Results: Variation over Time

- PSNR, VIF seem more affected by the method used to estimate quality variation over time
 - StdDev vs Total_variation



Future Plans

- Correlate single features of the metric (e.g., VMAF, VQM) with encoding parameters
- Compute other available metrics if feasible
 - e.g. OPVQ, derived from J.247
(<https://www.duo.uio.no/handle/10852/44706>)
- Experimenting with available clustering/classification algorithms

References

- M. Barkowsky, E. Masala, G. Van Wallendael, K. Brunnstrom, N. Staelens, P. Le Callet, Objective Video Quality Assessment – Towards large scale video database enhanced model development, IEICE Transactions on Communications, vol. E98-B, n. 1, pp. 2-11, Jan 2015
- ftp://vqeg.its.bldrdoc.gov/Documents/VQEG_Stockholm_Jul14/MeetingFiles/VQEG_JEG-Hybrid_2014_126_robust_decoder_Stockholm2014_EnricoMasala.pdf
- Van Wallendael, Glenn, Nicolas Staelens, Enrico Masala, and Marcus Barkowsky. "Full-HD HEVC-encoded video quality assessment database." In Ninth International Workshop on Video Processing and Quality Metrics (VPQM). 2015.
- ftp://ftp.ivc.polytech.univ-nantes.fr/VQEG/JEG/HYBRID/hevc_database/
- <http://media.polito.it/downloads/jeg/>
- Weka 3: <https://www.cs.waikato.ac.nz/ml/weka/>