

Project Hodor

Characterising objective metrics using a large scale database

Sky affiliates in collaboration with Ghent University, University College of Applied Sciences, AGH
University of Science and Technology, Politecnico Di Torino and Marcus Barkowsky

Contact: Florence Agboma

Project Hodor

A collaboration between Sky affiliates and academia

We are looking to

- Characterise different objective metrics using a large scale database
- Investigate systematic weaknesses in objective metrics and how to interpret the results
- Understand how different objective metrics perform based on content characteristics and encoding configurations
- Come up with a general methodology for testing different objective metrics, specifically where they agree/disagree
- Understand how to analyse failures without spending resources on subjective assessment
- Understand whether objective metrics behave alike based on the similarity of their algorithms and the category of their approach (i.e. signal based vs. modeling of the HVS)

Collaboration parties



POLITECNICO
DI TORINO



AGH UNIVERSITY OF SCIENCE
AND TECHNOLOGY



Note: Collaboration parties can consult VQEG members for their expertise

Example use case: Compare VQ during transcoder shootout

Example: Can you tell the difference in quality on the next few slides?

- Image has been removed

Example use case: Compare VQ during transcoder shootout

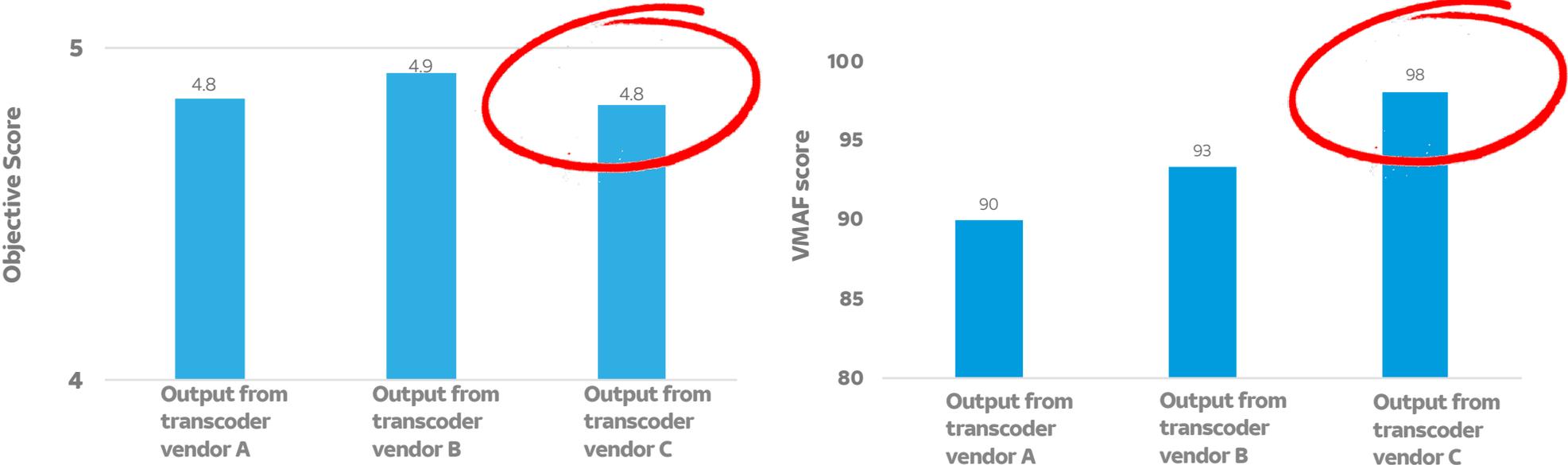
- Image has been removed

Example use case: Compare VQ during transcoder shootout

- Image has been removed

Example use case: Compare VQ during transcoder shootout

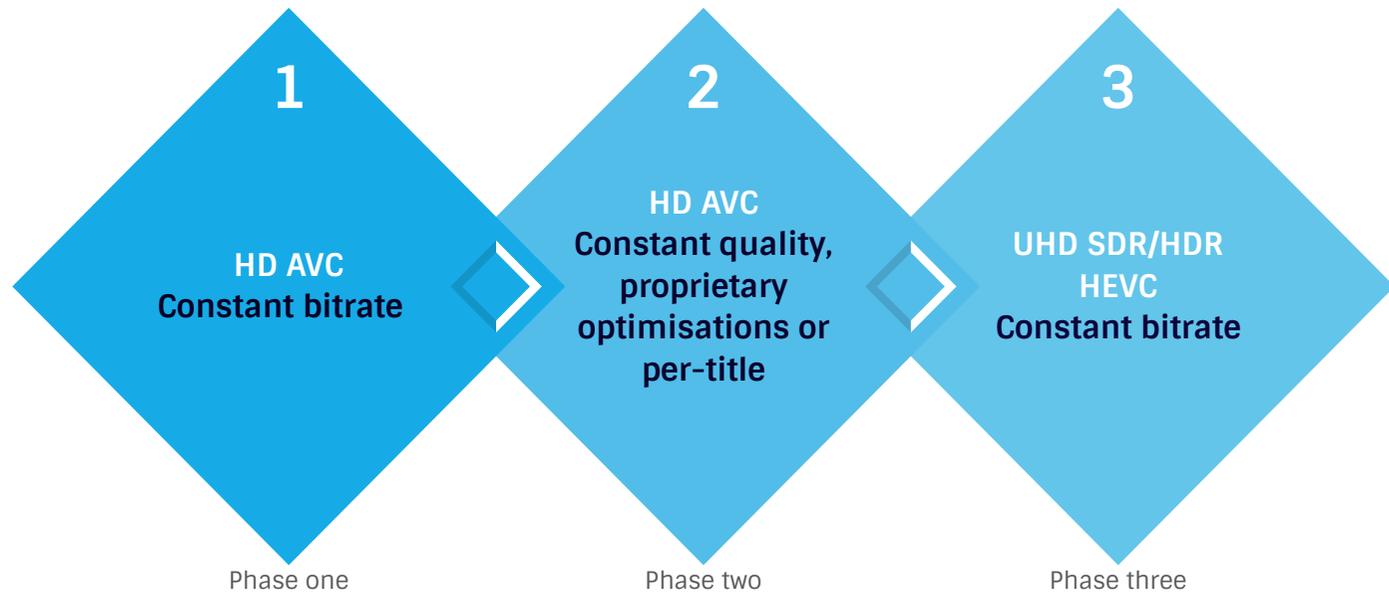
- Same images but inconsistent results from different VQ metrics



Our approach

Phases

- The agreement in place is valid until May 2021
- Project Hodor is broken down into phases
- Applied lessons learnt to re-prioritise milestones
- Phase one explores **HD AVC constant bitrates**



Phase one

HD Samples

JEG Hybrid database



Movies & Ents

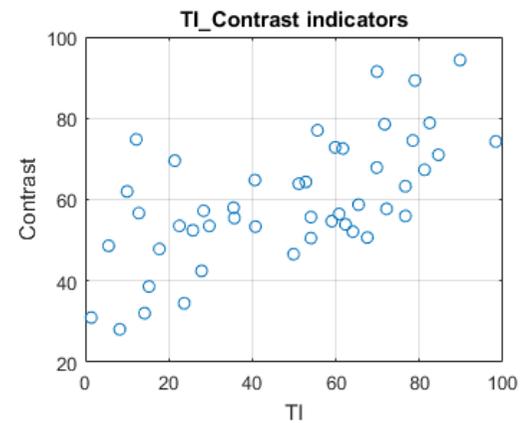
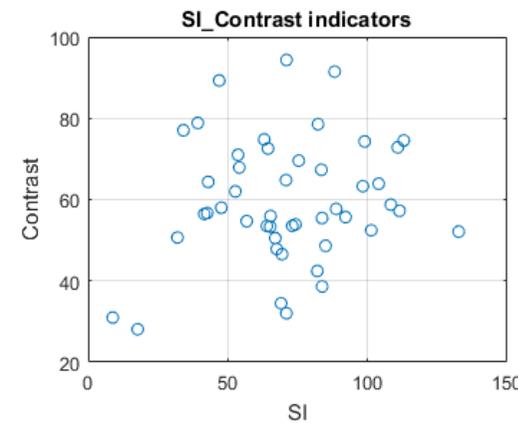
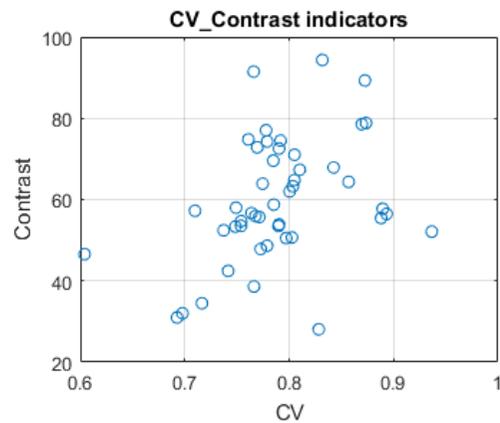
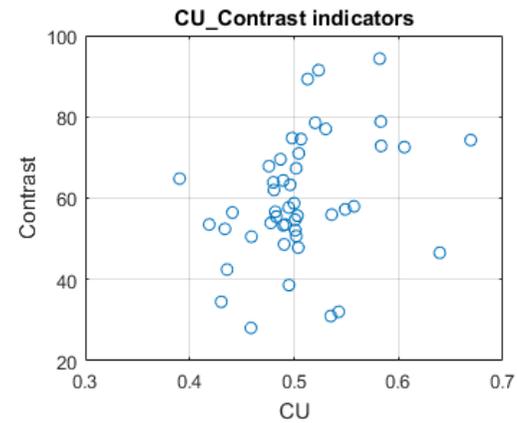
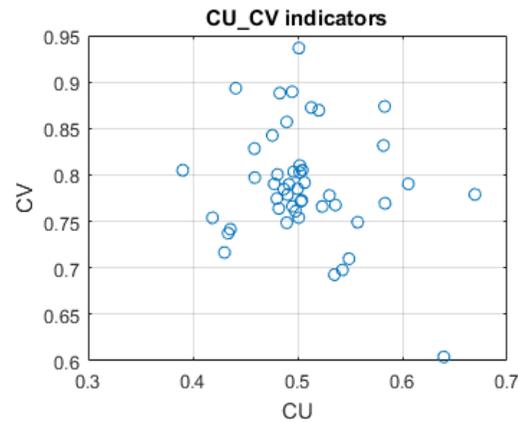
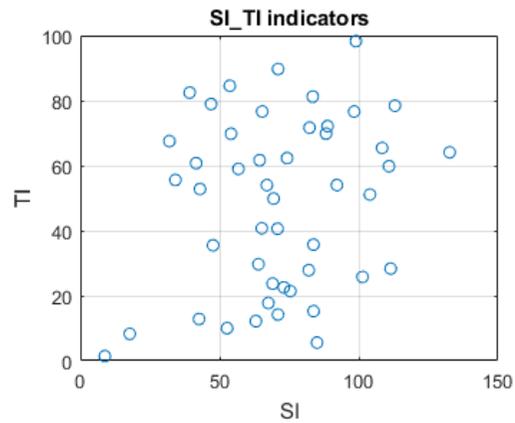
- Images have been taken out

Sports

- Images have been taken out

Phase one: content diversity of HD clips

University College of Applied Sciences (UCAS)



Phase one: content diversity of HD clips

Politecnico Di Torino

PCA: 93% of variability

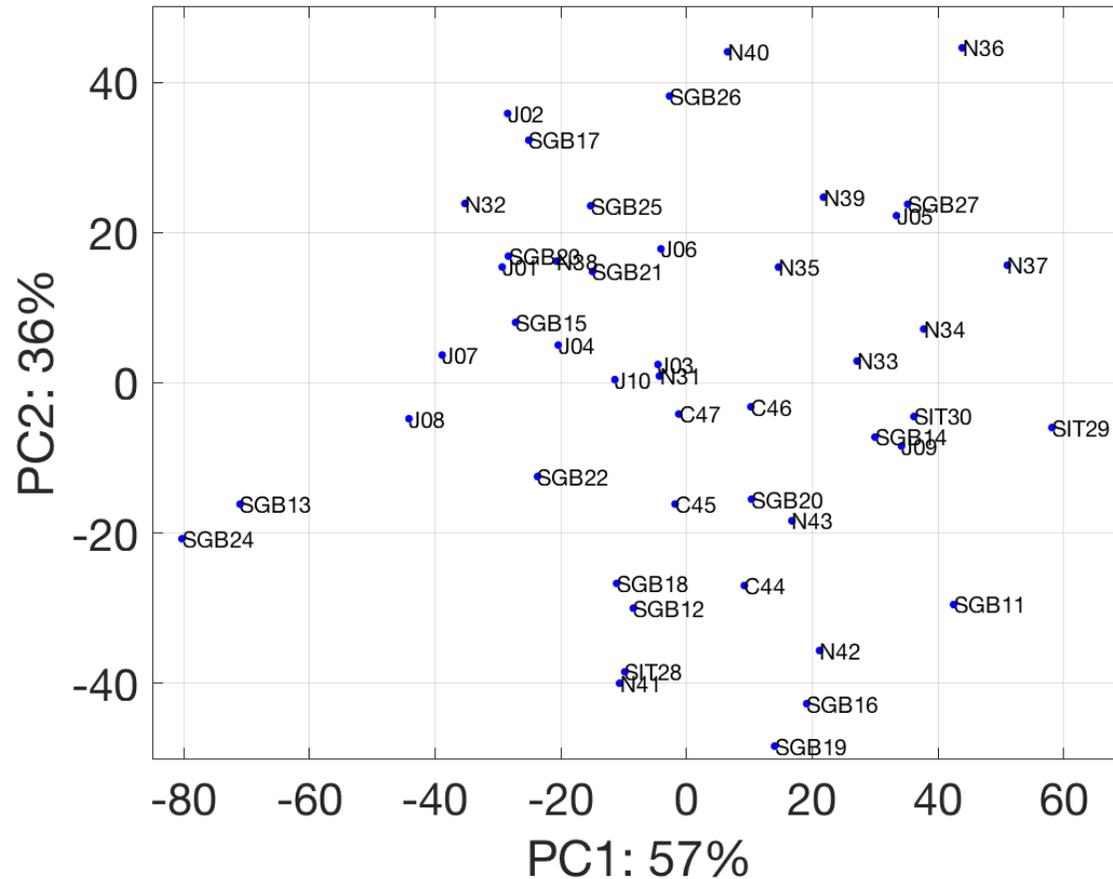


Table 1: Legend

Source	Label
JEG	J
SKY UK	SGB
SKY IT	SIT
NBC	N
CONCAST	C

Phase one: content diversity of HD clips

Politecnico Di Torino

PCA: 100% of variability

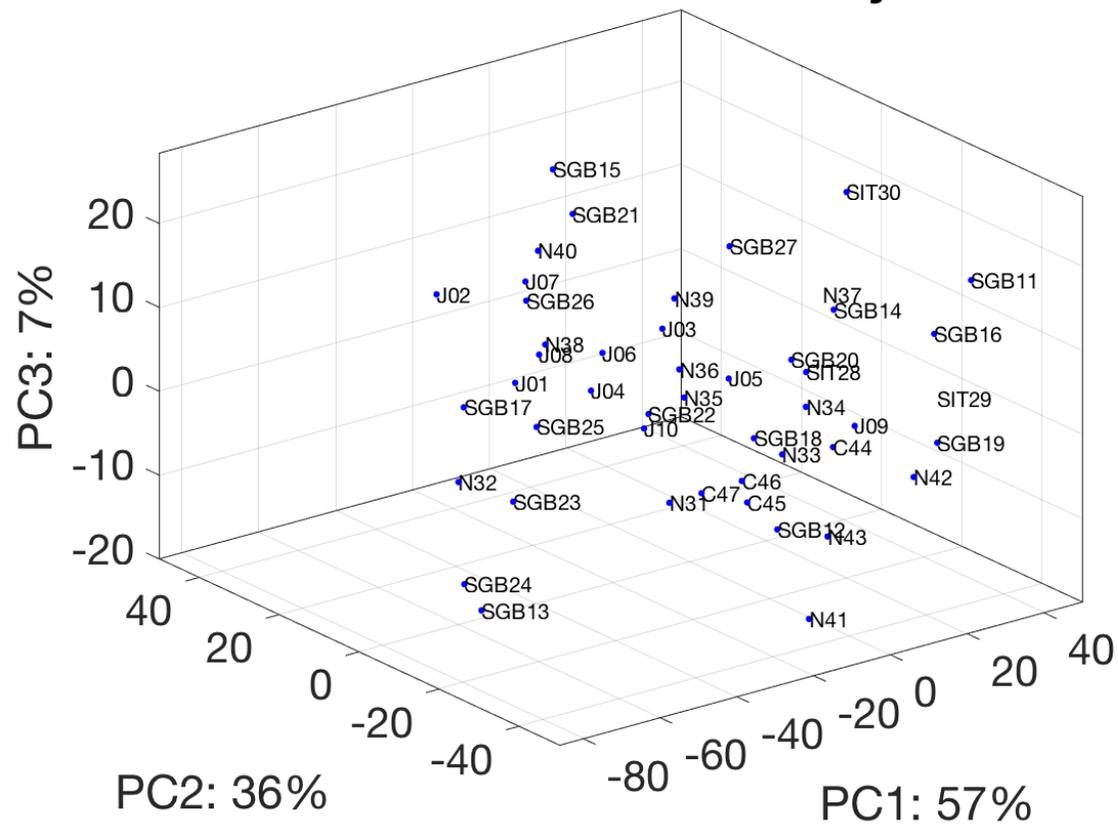


Table 1: Legend

Source	Label
JEG	J
SKY UK	SGB
SKY IT	SIT
NBC	N
CONCAST	C

Phase one: test matrix

Test matrix

- Selection of source inputs based on [1]

Characteristics	Name of asset HD/UHD
Animation, graphic overlays, and scrolling text	...
Repetitious or indistinguishable fine detail (e.g. gravel, grass, hair, rug, pinstripes)	...
Sharp black/white edges	...
Blurred background, with an in-focus foreground	...
Ramped color (e.g. sunset)	...
Water, fire, or smoke (for unusual shades and shifting patterns)	...
Flashing lights or extremely fast events	...
Action in a small portion of the total picture	...
Colourful scene	...
Multiple objects moving in a random, unpredictable manner	...
Visually simple imagery	...
Very saturated colors	...
Rotational movement (e.g. a carousel or merry-go round)	...

[1] Margaret Pinson, Marcus Barkowsky, and Patrick le Callet. Selecting scenes for 2D and 3D subjective video quality tests. EURASIP J. Image and Video Processing 2013: 50 (2013)

Phase one: test matrix

Test matrix

- SRC: 47 sources
- Duration : 10 seconds each
- HRC: Constant bitrates using Apple's HLS recommendations
- PVS: 376 processed video sequences
- Resolution: Up to 1080p
- H264 AVC

Profile	1	2	3	4	5	6	7	8
Resolution (w x h)	512 x 288	768 x 432	768 x 432	960 x 540	1280 x 720	1280 x 720	1920 x 1080	1920 x 1080
Video bitrate (kbps)	365	730	1100	2000	3000	4500	6000	7800

Current status of phase one

Create HD videos repository

- Minimum 40 different HD videos
- Duration of videos
- Content storage
- Generate content characteristics
- Apple's encoding rec

Generate VQ metrics

- VQ metric 1
- VQ metric 2
- SSIM, MS-SSIM, VIF, PSNR, and VMAF
- E-VMAF & ST-VMAF if available

Generate transcoded ABR ladder

- Apple's encoding rec
- Produce transcoded outputs
- Generate bitstream characteristics
- Prepare transcoded clips

Data analysis

- Statistical inference
- Correlation of metrics such as signal-based, HVS & neural networks
- PLCC, SRCC & RMSE
- Apply machine learning to analyse the objective scores

Report / End of phase one

We are here

Thank you