

JEG-Hybrid

Joint Effort Group

Plan of Experiments for the IGVQM Project

Enrico Masala

Politecnico di Torino University, Turin, Italy

VQEG Nov 2024 Meeting

Implementer's Guide to Video Quality Metrics

- More info on JEG-Hybrid website: <https://vqegjeg.github.io/jeg-hybrid/igvqm>
- Main idea: use a large set of subjectively-annotated datasets to offer standard guidelines in correctly using VQMs
 - Based on open data
 - Reproducible experiments

Database name	Contributor	Year	# of sources	# of PVS	# of opinion scores	Methodology	Scale (discrete/continuous)	Scoring
FOO	BAR	2008	30	300	9000	ACR	Discrete	
VQEG HD 1 to 5	VQEG	2010				ACR-HR	Discrete	
ITS4S	NTIA/ITS	2018				ACR	Discrete	
ITS4S2	NTIA/ITS	2019				ACR	Discrete	
ITS4S3	NTIA/ITS	2019				ACR	Discrete	
ITS4S4	NTIA/ITS	2020				ACR	Discrete	
CCRIQ	Intel, ITS, U. Ghent	2016				ACR	Discrete	
Public Safety #1	NTIA/ITS	2005				ACR-HR	Discrete	
Public Safety #2	NTIA/ITS	2006				ACR-HR	Discrete	
T1A1	ATIS	1996				DSIS	Discrete	
GamingVideoSET	KU	2018	6	90	2250	ACR	Discrete	
KUGVD	KU	2018	6	90	1530	ACR	Discrete	
AVT-VQDB-UHD-1	TU Ilmenau	2019	12	756	around 18,000 (t	ACR	Discrete	
AGH/NTIA/Dolby	AGH, NTIA, Dolby	2018				ACR-HR	Discrete	

https://docs.google.com/spreadsheets/d/1wYVjaqKrSB0z7D2eJlAd1g6y_831IF7TH6GhAKdKxA/edit

IGVQM: Plan of Experiments

- Selection of datasets
- Objective metrics
- Reproducibility
- Investigation of results / expected output

Selected Datasets

Dataset	PVS Resolution (max → min)	# PVS	# SRC	Codecs	Method	Notes
ITS4S	720p	514	514	?	ACR	
KUGVD	1080p → 480p	90	6	AVC	ACR	
GamingVideoSET	1080p → 480p	90	6	AVC	ACR	
AGH-NTIA-Dolby	1080p	224	84	AVC, MPEG2	ACR-HR	Discarded 6 SRC-PVS pairs due to temporal misalignment
AVT-VQDB-UHD-1	2160p → 360p 60 → 30 → 15 <i>fps</i>	756	20	AVC, HEVC, VP9	ACR	Some seq are the same but shorter version
TOTAL		1674	630			

- Manual conversion, for each dataset, of all subjective scores into a common JSON format suitable for automatic processing

Metrics

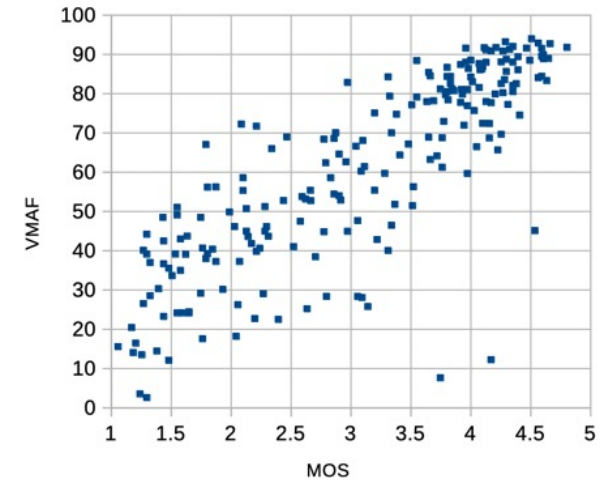
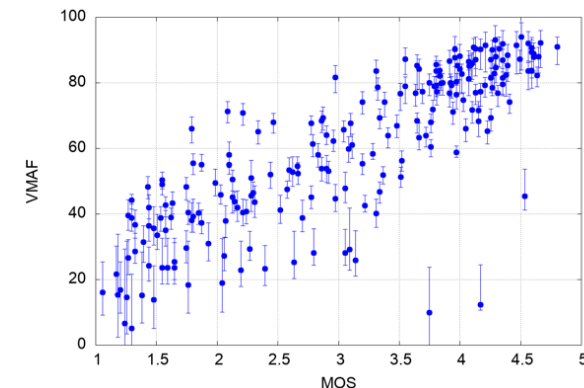
- Full reference metrics (from LibVMAF v. 3.0.0)
 - VMAF, computed using all 9 available models:
v0.6.1 regular+neg, float_v0.6.1 regular+neg, bootstrapping, float_bootstrapping, 4k, float_4k regular+neg
 - CAMBI
 - PSNR (Y,Cb,Cr), SSIM, MS-SSIM, CIEDE2000, PSNR-HVS
- No-reference metrics (still to be decided, e.g. python scikit-video(?))
 - Any suggestion for frameworks/libraries? (<https://github.com/NTIA/NRMetricFramework> ?)
 - Does using Matlab code impact on reproducibility?

Reproducibility and Software

- **Docker**-style solution
 - Can (also) be directly translated in “Singularity” format suitable for HPC (High-Performance Computing)
- VMAF v. 3.0.0: to compute all metrics
- FFmpeg v. 7.0.2 to perform:
 - PVS decoding (e.g. from mp4)
 - Format conversion (yuv, y4m)
 - Color handling (yuv422 → yuv420)
 - Upscaling (with best lanczos filter settings)
- Storing hashes of the intermediate results (e.g., decoded content) to allow checking reproducibility correctness
- Storing full results from metric computation (i.e., libVMAF detailed JSON output)

Investigation of Results / Expected Output

- **Investigation**
 - Metrics vs MOS relation (e.g. scatter plots)
 - Including different temporal pooling methods
 - Influence of the various VMAF models on the result
 - Including using wrong models (e.g., fullHD model with 4K content)
 - VMAF bootstrapping models vs MOS confidence intervals
- **Expected output**
 - Publication of all software to reproduce experiments
 - Drafting recommendation (white paper / publication?)



Comments / Proposals / Ideas

- Feel free to do it **now!**



- Join our biweekly conference call (next: **Nov 25, 2024, 4pm CET/CEST**)
- References:
 - **JEG-Hybrid website** <https://vqegjeg.github.io/jeg-hybrid/>
 - Including an up-to-date publication section with immediate download link for each one
 - **IGVQM** subsection: <https://vqegjeg.github.io/jeg-hybrid/igvqm>
 - VQEG-JEG mailing list: jeg@vqeg.org (<https://vqeg.org/email-reflectors.aspx>)