PhD Defense

Data-Driven Visual Quality Estimation Using Machine Learning

Steve Göring

Audiovisual Technology Group,
Technische Universität Ilmenau, Germany;
Email: steve.goering@tu-ilmenau.de

May 18, 2022

TECHNISCHE UNIVERSITÄT
ILMENAU

Introduction – General Scope

TECHNISCHE UNIVERSITÄT
ILMENAU

- mobile phones, camera/screen technology, internet
- more content + consumption [45]
- higher resolutions (4K/8K screens, recordings)
- traditional 2D videos and images

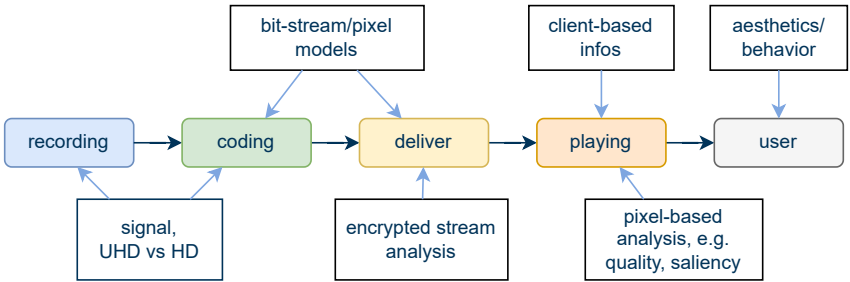
Introduction

- internet:
  - **higher demand for video streaming**, up to 80% video streaming [5]
  - **increase of uploaded images**, e.g., up to 95M per day for Instagram [1]
  - internet bandwidth not necessarily adapting to trends
- new image/video compression methods

image: `https://matrix.fandom.com/wiki/The_Architect`

Motivation

- 10 s UHD-1 raw video[1] $\approx$ 12 *GB* vs. Wikipedia-text-only-zip $\approx$ 14 *GB* [48]
- lossy image/video compression $\rightarrow$ quality evaluation required



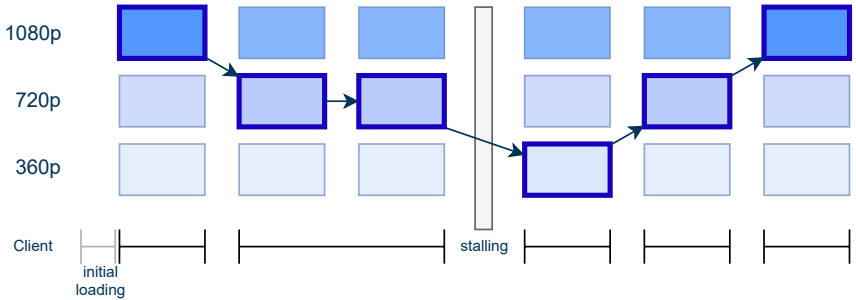
[1] (YUV, 4:2:2, 10bit, 60 fps)

Context – Quality of Experience

- system factors, context factors, human factors[2]
- quality perceived by users; overall media experience

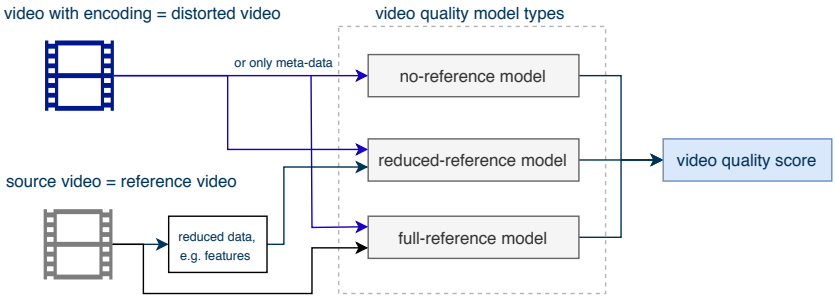
[2] *Le Callet et al.* 2012: "Qualinet white paper on definitions of quality of experience"

Context – HAS/DASH Streaming and Encoding



- long-term audio-visual quality, e.g, P.1203.3 [21]
- short-term "segment"-level visual quality

Context – Visual Quality Prediction Models



- parametric models: P.1203.1 P_v [21], P.1204.5 [23], ...
- machine learning/dnn models: P.1204.3 [22], VMAF [31], NIMA [44], ...
- common: overall mean opinion score VQ_{mos}, other: VQ_{class}, VQ_{prop}

Research Questions

TECHNISCHE UNIVERSITÄT
ILMENAU

- **RQ 1**: robust video quality models using machine learning
- **RQ 2**: speed up of state-of-the-art quality models
- **RQ 3**: perceivable difference for higher quality content
- **RQ 4**: predict more than mean opinion scores
- **RQ 5**: using video quality models/compression for images

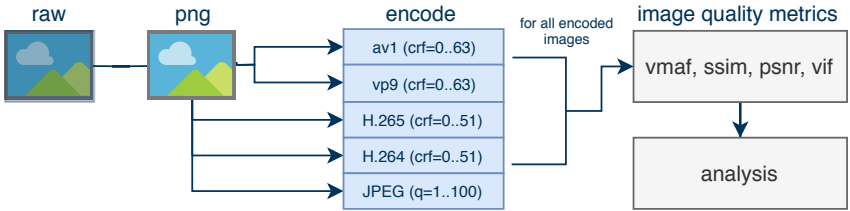
  $\rightarrow$ image quality, video quality, other applications

High Resolution Image Quality Evaluation

- image formats for web [6]: JPEG, PNG, GIF
- recent: AVIF [32], HEIF [26], VVC + DNN [28]: video based
- **open:** quality evaluation/models for
  - newer methods, higher resolutions,
  - large + diverse datasets, more than PSNR/SSIM

  $\rightarrow$ pre-analysis for video compression approaches

Objective Evaluation for Image Compression using Video Encoders

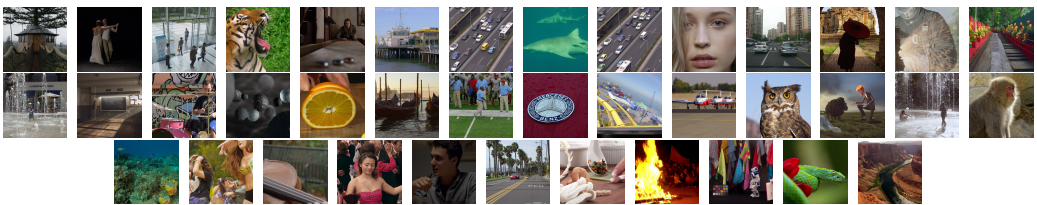


- AVT-Image-Database[3]: 1133 source images CC0 (wesaturate)[4]
- 380k encoded images + objective scores: AV1|VP9 > H.265 > JPEG

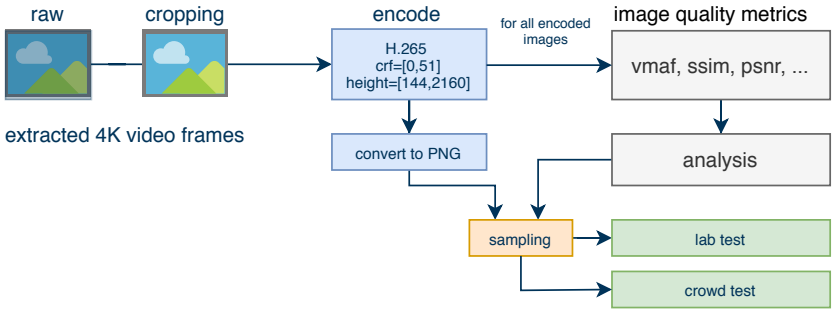
[3] *Göring et al.* 2019: "Evaluation of Intra-coding based image compression"
[4] code + data: `http://git.avt-imt.de/image_compression`

Quality Evaluation for HighRes Images

- VMAF suitable for image quality?→ lab or remote/crowd test required
- H.265 encoding, UHD-1/4K center cropped frames



Quality Evaluation for HighRes Images – Pipeline



- 39 source UHD-1/4K frames $\rightarrow$ 246k compressed images
- meta-data models: *IMG-h265-rf*, *IMG-h265-para*: VMAF predictions
- two rounds of sub-sampling $\rightarrow$ 371 stimuli

Quality Evaluation for HighRes Images – Tests

TECHNISCHE UNIVERSITÄT ILMENAU

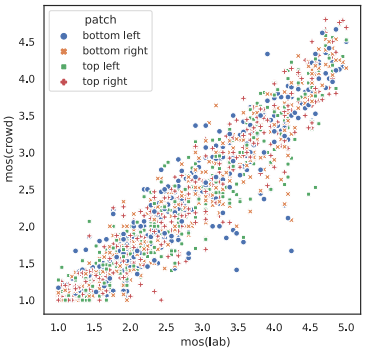
test	tool	environment	#images	to-be-rated	#participants
lab-test	AVRateNG [2]	standardized; 4K screen	full; 371	all	21
online-test	AvrateVoyager-Dev [17]	wide range	1080p-patches; 1488	150 random	238

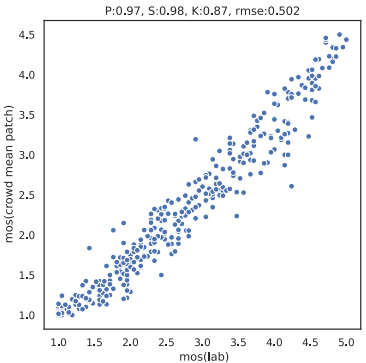
- lab-test: no outliers $\rightarrow$ VMAF best
- online-test: mostly: $720p$-$900p$ screens; pre-study, follow up [37]
- both: quality distribution + SOS-analysis [20] similar results

Quality Evaluation for HighRes Images – Results

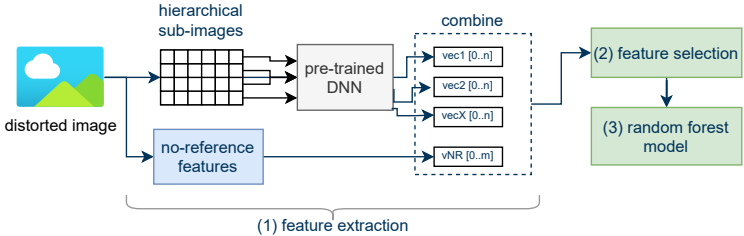
TECHNISCHE UNIVERSITÄT ILMENAU

- lab vs. online test: per patch, mean patch
- individual patches: high correlation to lab test





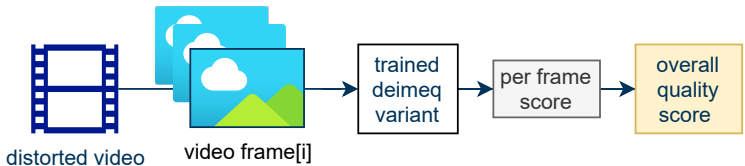
Pixel-based Image Quality Prediction – deimeq



- *deimeq*[5]: DNN-based image quality (mid-high resolution) model
  - feature extractor + RF model, hierarchical sub-images
  - cross dataset evaluation (TID 2013 [34], Live 2 [43])
  - similar to NIMA [44]

[5] *Göring et al.* 2018: "deimeq – A DNN Based Hybrid No-reference Image Quality Model"

Pixel-based Image Quality Prediction – deviq



- *deviq*[6]: extension of *deimeq* to UHD-1/4K video quality prediction
- **limits:** processing time; trained on VMAF per frame; motion aspects

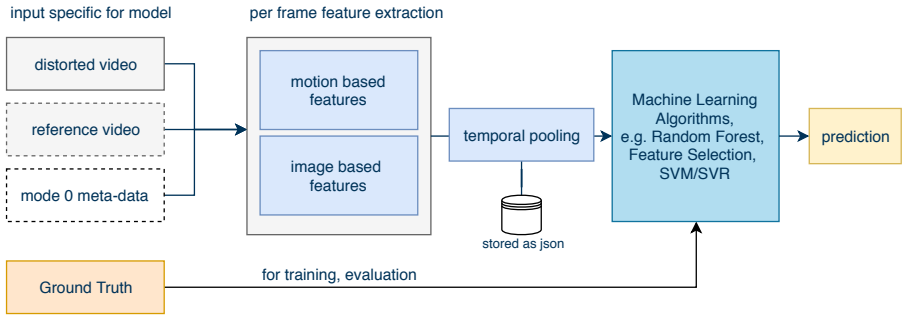
[6] *Göring et al.* 2018: "DeViQ – A deep no reference video quality model"

Models for Video Quality Prediction

- limits of VMAF: framerate: test #4 AVT-VQDB-UHD-1[7]
- feature-based models + machine learning
  - pixel-based features (no bitstream models)
  - context of P.NATS Phase 2
  - TUIL+DT pixel candidate models
  - framework[8], extensible, reproducible, open-source[9]
  - extensions to other applications

[7] *Rao et al.* 2019: "AVT-VQDB-UHD-1: A Large Scale Video Quality Database for UHD-1"
[8] *Göring et al.* 2021: "Modular Framework and Instances of Pixel-based VideoQM for UHD-1/4K"
[9] library: `http://git.avt-imt.de/quat` | models: `http://git.avt-imt.de/pixelmodels`

Models for Video Quality Prediction – Architecture



Models for Video Quality Prediction – Features

TECHNISCHE UNIVERSITÄT ILMENAU

Feature	Feature Type	Source	#Values
contrast, blur	img	own [15]	1/F
fft	img	[25]*	1/F
colorfulness	img	[19]*	1/F
tone, saturation	img	[3]*	1/F
scene_cuts	mov	own	1/F
movement, temporal	mov	own [15]	1/F
si, ti	img, mov	[24]	1/F
blockmotion	mov	own [12, 15]	3 /F
cub{row,col}.{0,0.3, 0.5, 0.6,1.0}	mov	own [12]	1/F
staticness	mov	own [12, 15]	1/F
uhdhdsim	img	own [15]	1/F
blockiness	img	own [12]	1/F
noise	img	[7]	1/F
PSNR, SSIM, VIF	img-fr	[47, 46, 42]	1/F
fps_est	mov-fr	own	1/F
framerate , bitrate, codec, resolution, bpp	bs		1/S
log(framerate, bitrate, resolution), norm(framerate, resolution)	bs		1/S
brisque	img-nofu	[30]	36/F

Models for Video Quality Prediction – Instances

TECHNISCHE UNIVERSITÄT ILMENAU

- instances:
  - **nofu**: no-reference: img + mov + img-nofu features
  - **hyfu**: hybrid no-reference (mode 0): img + mov + bs features
  - **fume**: full-reference: img/mov(dis, ref), img-fr, mov-fr
  - **hyfr**: hybrid-full-reference: fume + bs features
- all 360p center crop, temporal pooling, RF-based with feature selection
- trained+validated for VQ_{mos} , VQ_{class}, VQ_{prob}

Models for Video Quality Prediction – Datasets

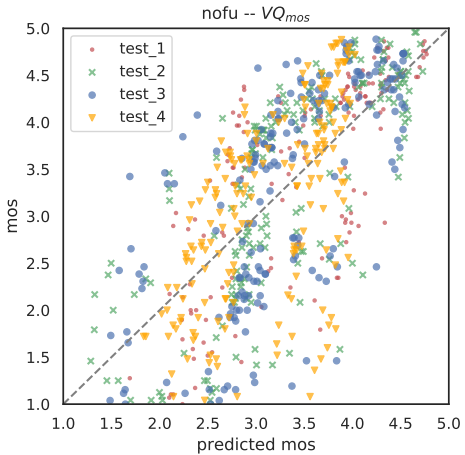
TECHNISCHE UNIVERSITÄT ILMENAU

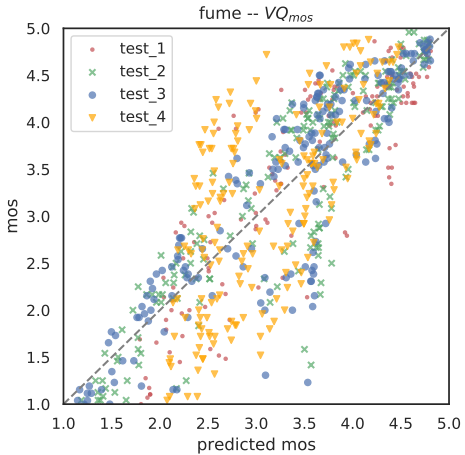
dataset	# tests	# srcs	fps	bitrates	codecs	resolutions
AVT-PNATS-UHD-1	4	>50	15 - 60	100 kbps to 50 mbps	VP9, H.264, H.265	360p - 2160p
AVT-VQDB-UHD-1	4	17	15 - 60 (test #4)	200 kbps to 50 mbps	VP9, H.264, H.265	360p - 2160p

- training: AVT-PNATS-UHD-1: subset of P.NATS Ph2 data [36]
- validation: AVT-VQDB-UHD-1[10]: publicly available
- no common sequences in train/validation

[10] *Rao et al.* 2019: "AVT-VQDB-UHD-1: A Large Scale Video Quality Database for UHD-1"

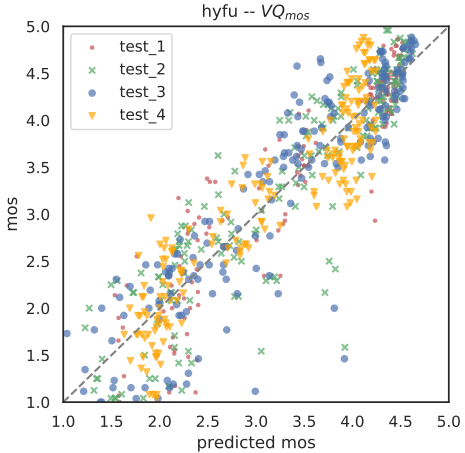
Models for Video Quality Prediction– Evaluation (1)

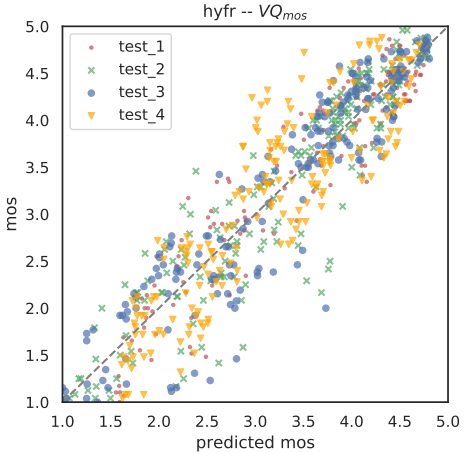




- **nofu** (P=0.701) < VMAF (P=0.816) < **fume** (P=0.835)

Models for Video Quality Prediction– Evaluation (2)





- **fume** (P=0.835) < **hyfu** (P=0.910) < **hyfr** (P=0.922)

Models for Video Quality Prediction– Evaluation (3)

- VQ_{class}, VQ_{prob}; other performance metrics; ML-algos: similar results
- **fume**: problems with test #4 (training: lesser low fps cond.)
- **nofu**: similar to PSNR/SSIM
  - better for specialized use-cases
  - larger center crop: small improvement
  - pure no-reference models: still challenging
- switching train/validation: better results; no reproducibility of predictions

Other Applications of the Model Pipeline

TECHNISCHE UNIVERSITÄT ILMENAU

- proposed architecture (features, pooling, ml-aglos) usable for:
  - source video classification for UHD-1/4K [15]
  - gaming video quality [12] and genre prediction [16]
  - encoding parameter estimation [13]
  - speed up approaches [8]

Gaming Video Quality Prediction

TECHNISCHE UNIVERSITÄT ILMENAU

- specialized **nofu** model= **nofu-gaming**[11]
  - features: *fft*, *ti*, *si*, *block{iness, motion}*, *staticness*, *cub{col,row}-{0,1.0}*
  - pooling: mean value, std, first value, $n = 3$ temporal groups: mean + std
- dataset: GamingVideoSET [4]
- 10-fold cross-validation: P=0.91 > re-trained brisque+niqe > vmaf
- src-video fold: nofu-gaming > brisque+niqe

[11] *Göring et al.* 2019: "nofu - A Lightweight No-Reference Pixel Based VideoQM for Gaming Content"

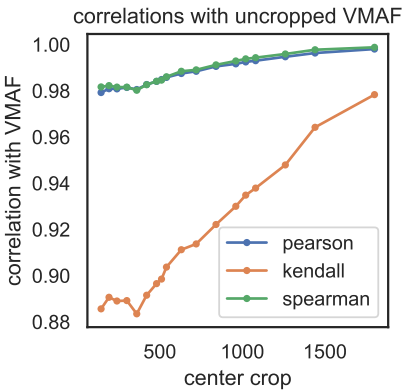


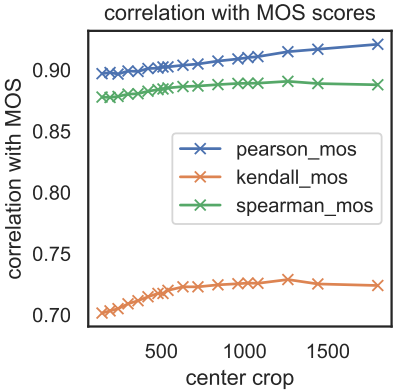



Speed-up Approaches

TECHNISCHE UNIVERSITÄT ILMENAU

- **cencro**[12]: reduce processing time of SoA FR-models; per frame reduction
- center crop $144p$ to $1800p$: $360p$ (45 min to 3 min)





[12] *Göring et al.* 2019: "cencro – Speedup of Video Quality Calculation using Center Cropping"

Contributions to SoA

- publications
  - pixel-based image/video quality models, features, or analysis [18, 12, 15]
  - P.1203 [21, 40], P.1204 [39, 36], or bitstream-related [11, 35, 41], ...
- open source software for video quality
  - AVrateNG [2], AVrateVoyager [37, 17]
  - P.1203 open source implementation [40], ITU-P.1204.3 reference software [39]
  - cencro [8], quat + pixelmodels [18],
  - AVT-VQDB-UHD-1 [38], ...
- presented work= summary of 5 years

Conclusion

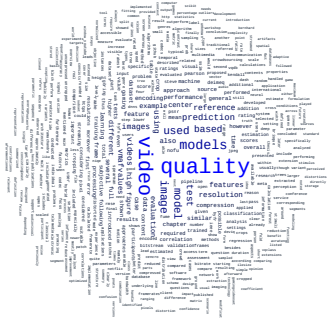
- high resolution image quality + compression:
  - using video codecs > JPEG
  - lab and crowd/remote/online tests for image quality
  - image quality prediction models
- video quality models for UHD-1/4K
  - general architecture, various model instances
  - other applications of the provided architecture

Future Work

- integration of video quality models in long-term quality prediction
- evaluation of new distortions: bending, compression artifacts (DNNs)
- UHD-2/8K, HDR, HFR, UGC, 360°, Point Clouds, light field
- end-to-end chain: encrypted streams; cameras; liking
- immersive media technology: room lighting, haptic feedback ...

Thank you for your attention

TECHNISCHE UNIVERSITÄT
ILMENAU



..... are there any questions?

Thanks to all ..

TECHNISCHE UNIVERSITÄT
ILMENAU

colleagues, collaborators, friends, family, . . .

Alexander Raake Peter List Robert Steger Christiane Wisser Saman Zadtootaghaj Matthias Döring Nabajeet Barman Dominik Keller Konstantin Brand Silvio Borer Detlef Beyer Simon Broom Martin Meyer Jochen Seitz Lukas Treybig Ashutosh Singla Lea Skorin-Kapov Simon Wedel Thomas Wöhner Patrick Vogel Erik Hofman Marie-Neige Garcia Werner Robitza Janto Skowronek Frank Hofmeyer Monique Rodegast Bernhard Feiten Ulf Wüstenhagen Julian Zebelein Claudia Stirnat Alexander Dethof Rakesh Rao Ramachandra Rao Madlen Langbein Thaden Cohrs Gunnar Heikkilä Andrea Reis Patrick Le Callet Stephan Fremerey Bernd Hildenbrandt André Siegel Jörgen Gustafsson Juan Jose Villamar Villarreal Kazuhisa Yamagishi Tim Gubner Nicolas Pachatz Torsten Demmler Petra Göring David Lindero Annika Neidhardt Christopher Krämmer Shahid Satti Susann Kohout Stephan Werner Eckhardt Schön . . .

References I

[1] 99firms. *Instagram Marketing Statistics*. 2020. URL: https://99firms.com/blog/instagram-marketing-statistics/%5C#gref (visited on 03/07/2020).

[2] AVRateNG. *AVRateNG – github project*. URL: https://github.com/Telecommunication-Telemedia-Assessment/avrateNG (visited on 03/07/2020).

[3] Tunç Ozan Aydın, Aljoscha Smolic, and Markus Gross. "Automated aesthetic analysis of photographic images". In: *IEEE transactions on visualization and computer graphics* 21.1 (2015), pp. 31–42.

[4] Nabajeet Barman et al. "GamingVideoSET: a dataset for gaming video streaming applications". In: *2018 16th Annual Workshop on Network and Systems Support for Games (NetGames)*. IEEE. 2018, pp. 1–6.

References II

[5] Cisco. *Cisco visual networking index: Forecast and trends, 2017–2022 (White Paper)*. 2018.

[6] Anton Garcia Diaz. *Picture this: the best image format for the web in 2019*. 2018. URL: https://www.freecodecamp.org/news/best-image-format-for-web-in-2019-jpeg-webp-heic-avif-41ba0c1b2789/ (visited on 03/07/2020).

[7] David L Donoho and Jain M Johnstone. "Ideal spatial adaptation by wavelet shrinkage". In: *biometrika* 81.3 (1994), pp. 425–455.

References III

[8] Steve Göring, Christopher Krämmer, and Alexander Raake. “cencro – Speedup of Video Quality Calculation using Center Cropping”. In: *21st IEEE International Symposium on Multimedia (IEEE ISM)*. Dec. 2019, pp. 1–8.

[9] Steve Göring and Alexander Raake. “deimeq – A Deep Neural Network Based Hybrid No-reference Image Quality Model”. In: *7th European Workshop on Visual Information Processing (EUVIP)*. IEEE. Nov. 2018, pp. 1–6.

[10] Steve Göring and Alexander Raake. “Evaluation of Intra-coding based image compression”. In: *8th European Workshop on Visual Information Processing (EUVIP)*. IEEE. 2019, pp. 1–6.

References IV

[11] Steve Göring, Alexander Raake, and Bernhard Feiten. “A framework for QoE analysis of encrypted video streams”. In: *Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. May 2017, pp. 1–3.

[12] Steve Göring, Rakesh Rao Ramachandra Rao, and Alexander Raake. “nofu - A Lightweight No-Reference Pixel Based Video Quality Model for Gaming Content”. In: *Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. Berlin, Germany, June 2019.

[13] Steve Göring, Rakesh Rao Ramachandra Rao, and Alexander Raake. “Prenc – Predict Number Of Video Encoding Passes With Machine Learning”. In: *Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland, May 2020.

References V

[14] Steve Göring, Janto Skowronek, and Alexander Raake. "DeViQ – A deep no reference video quality model". In: *Electronic Imaging, Human Vision Electronic Imaging* 2018.14 (2018), pp. 1–6.

[15] Steve Göring et al. "Analyze And Predict the Perceptibility of UHD Video Contents". In: *Electronic Imaging, Human Vision Electronic Imaging* 2019.12 (2019).

[16] Steve Göring et al. "Automated Genre Classification for Gaming Videos". In: *2020 IEEE 22st International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. 2020, pp. 1–6.

[17] Steve Göring et al. "AVRate Voyager: an open source online testing platform". In: *2021 IEEE 23st International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. 2021, pp. 1–6.

References VI

[18] Steve Göring et al. “Modular Framework and Instances of Pixel-based Video Quality Models for UHD-1/4K”. In: *IEEE Access* 9 (2021), pp. 31842–31864.

[19] David Hasler and Sabine E Suesstrunk. “Measuring colorfulness in natural images”. In: *Human vision and electronic imaging VIII*. Vol. 5007. International Society for Optics and Photonics. 2003, pp. 87–96.

[20] Tobias Hoßfeld, Raimund Schatz, and Sebastian Egger. “SOS: The MOS is not enough!” In: *2011 third international workshop on quality of multimedia experience*. IEEE. 2011, pp. 131–136.

[21] ITU-T. *ITU-T Rec. P.1203 (10/17)*. Tech. rep. Int. Telecommunication Union, 2017.

References VII

[22] ITU-T. *Recommendation P.1204.3 - Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full bitstream information*. Tech. rep. International Telecommunication Union, 2019.

[23] ITU-T. *Recommendation P.1204.5 : Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to transport and received pixel information*. Tech. rep. International Telecommunication Union, 2019.

[24] ITU-T. *Subjective video quality assessment methods for multimedia applications*. Serie P: Telephone Transmission Quality, Telephone Installations, Local Line Networks. Vol. Recommendation ITU-T P.910. International Telecommunication Union. Geneva, 2008.

References VIII

[25] Ioannis Katsavounidis, Anne Aaron, and David Ronca. "Native resolution detection of video sequences". In: *Annual Technical Conference and Exhibition, SMPTE 2015*. SMPTE. 2015, pp. 1–20.

[26] Jani Lainema et al. "HEVC still image coding and high efficiency image file format". In: *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2016, pp. 71–75.

[27] Patrick Le Callet, Sebastian Möller, Andrew Perkis, et al. "Qualinet white paper on definitions of quality of experience". In: *European network on quality of experience in multimedia systems and services (COST Action IC 1003)* 3.2012 (2012).

References IX

[28] Wei-Cheng Lee et al. "A Hybrid Layered Image Compressor with Deep-Learning Technique". In: *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. 2020, pp. 1–6.

[29] Hui Men, Hanhe Lin, and Dietmar Saupe. "Spatiotemporal Feature Combination Model for No-Reference Video Quality Assessment". In: *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2018, pp. 1–3.

[30] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain". In: *IEEE Transactions on Image Processing* 21.12 (2012), pp. 4695–4708.

[31] Netflix. *Netflix VMAF, https://github.com/Netflix/vmaf; 2019-07-03*.

References X

[32] The Alliance for Open Media. *AV1 Image File Format (AVIF)*. 2019. URL: https://aomediacodec.github.io/av1-avif/ (visited on 03/07/2020).

[33] Kiyotaka Otsuji and Yoshinobu Tonomura. "Projection-detecting filter for video cut detection". In: *Multimedia Systems* 1.5 (1994), pp. 205–210.

[34] Nikolay Ponomarenko et al. "Image database TID2013: Peculiarities, results and perspectives". In: *Signal Processing: Image Communication* 30 (Jan. 2015), pp. 57–77.

References XI

[35] Alexander Raake et al. “A bitstream-based, scalable video-quality model for HTTP adaptive streaming: ITU-T P.1203.1”. In: *Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. May 2017, pp. 1–6.

[36] Alexander Raake et al. “Multi-model standard for bitstream-, pixel-based and hybrid video quality assessment of UHD/4K: ITU-T P.1204”. In: *IEEE Access* 8 (2020), pp. 193020–193049.

[37] Rakesh Rao Ramachandra Rao, Steve Göring, and Alexander Raake. “Towards High Resolution Video Quality Assessment in the Crowd”. In: *13th International Conference on Quality of Multimedia Experience (QoMEX)*. 2021.

References XII

[38] Rakesh Rao Ramachandra Rao et al. “AVT-VQDB-UHD-1: A Large Scale Video Quality Database for UHD-1”. In: *21st IEEE International Symposium on Multimedia (IEEE ISM)*. Dec. 2019, pp. 1–8.

[39] Rakesh Rao Ramachandra Rao et al. “Bitstream-based Model Standard for 4K/UHD: ITU-T P.1204.3 – Model Details, Evaluation, Analysis and Open Source Implementation”. In: *Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland, May 2020.

[40] Werner Robitza et al. “HTTP Adaptive Streaming QoE Estimation with ITU-T Rec. P.1203 – Open Databases and Software”. In: *9th ACM Multimedia Systems Conference*. Amsterdam, 2018.

References XIII

[41] Werner Robitza et al. “Measuring YouTube QoE with ITU-T P. 1203 under Constrained Bandwidth Conditions”. In: *Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2018, pp. 1–6.

[42] Hamid R Sheikh and Alan C Bovik. “Image information and visual quality”. In: *IEEE Trans. Image Process.* 15.2 (2006), pp. 430–444.

[43] Hamid R Sheikh, Muhammad F Sabir, and Alan C Bovik. “A statistical evaluation of recent full reference image quality assessment algorithms”. In: *IEEE Transactions on image processing* 15.11 (2006), pp. 3440–3451.

[44] Hossein Talebi and Peyman Milanfar. “NIMA: Neural image assessment”. In: *IEEE Trans. on Image Processing* 27.8 (2018), pp. 3998–4011.

References XIV

[45] vau.net. *Media Usage 2018: Germans are using audiovisual media for more than nine hours each day*. 2019. URL: https://www.vau.net/pressemitteilungen/content/media-usage-2018-germans-using-audiovisual-media-than-nine-hours-day (visited on 03/07/2020).

[46] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. "Multiscale structural similarity for image quality assessment". In: *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*. Vol. 2. IEEE. 2003, pp. 1398–1402.

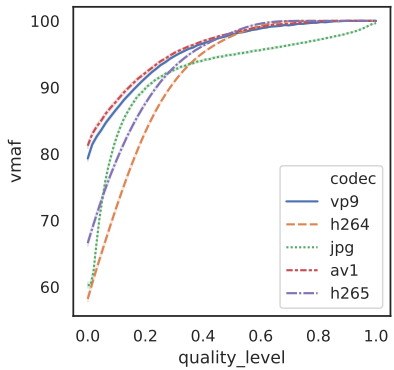
[47] Zhou Wang et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE Trans. Image Process.* 13.4 (2004), pp. 600–612.

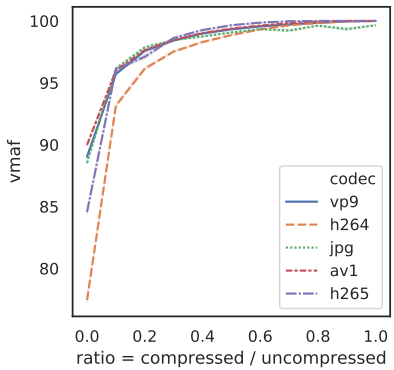
References XV

[48] Wikipedia. *Wikipedia Database Download*. 2020. URL: https://en.wikipedia.org/wiki/Wikipedia:Database%5C_download (visited on 12/25/2019).

[49] Ramin Zabih, Justin Miller, and Kevin Mai. *Feature-based algorithms for detecting and classifying scene breaks*. Tech. rep. Cornell University, 1995.

Image Compression using Video Encoders – Results





- quality-level= unified "setting" for compression
- similar results for other metrics: SSIM, PSNR, VIF
- higher quality range: video methods better (AV1|VP9; H.265)

Quality Evaluation for HighRes Images– Results Lab vs. obj. metrics

- lab test vs. objective quality metrics

metric	pearson	kendall	spearman
vmaf	0.919	0.757	0.925
adm2	0.868	0.722	0.901
vif scale2	0.861	0.740	0.911
vif scale3	0.852	0.786	0.941
vif scale1	0.846	0.674	0.859
ms ssim	0.701	0.658	0.851
psnr	0.698	0.524	0.719
ssim	0.658	0.802	0.948
vif scale0	0.619	0.472	0.643

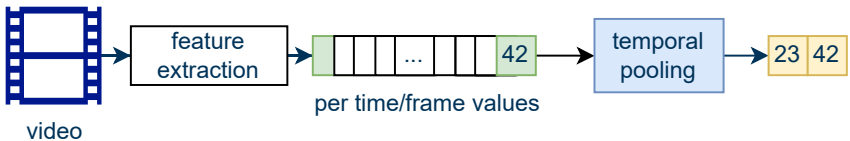
Models for Video Quality Prediction– Features – Examples (1)

- staticness: mean frame on all currently played frames + SI measure
- noise: wavelet-based estimator for noise [7] (skimage)
- scene_cuts: 0,1: resized $360p$ view + threshold-based detection, similar to [33, 49] (skvideo)
- cube{col,row}: similar to [29]: a sliding window=60 frames as cuboid, slicing planes + SI measure

Models for Video Quality Prediction– Features – Examples (2)

- movement: foreground-background % of moving foreground (open-cv)
- blockmotion: SE3SS algo, 10% of height as blocksize (may be suboptiomal, speed tradeoff), counting block movement directions
- temporal: $RMSE(f_i, f_{i+1})$, for RGB
- blockiness: "guessing" blocksize, canny edge detector, mean on blocklines horizontal, vertical, weighted
- uhdhdsim: PSNR to rescaled versions (Full-HD – 4K)

Models for Video Quality Prediction– Pooling



- most of the features produce per-frame scores
- time-independent features for ml-model required: temporal pooling
- for a given feature value vector $f \rightarrow$ total 25 statistical values:
  - mean value, std, skewness, kurtosis, IQR, quantiles, last + first value
  - split into $n = 3$ temporal groups: for each group: mean + std

Models for Video Quality Prediction– Speed-up

- full-frame calculation of all features: time consuming
- (A) reduction of frames: breaks movement features
- (B) reduction per frame: focus on 360p center cropped view of videos
  - drastic time reduction, nearly negligible error
  - wider range of features compensate for 360p view

Models for Video Quality Prediction– Evaluation - cc

- 32 runs each 10-fold-cross validation