AVQBits — Adaptive Video Quality Model Based on Bitstream Information for Various Video Applications

Rakesh Rao Ramachandra Rao, Steve Göring, Alexander Raake

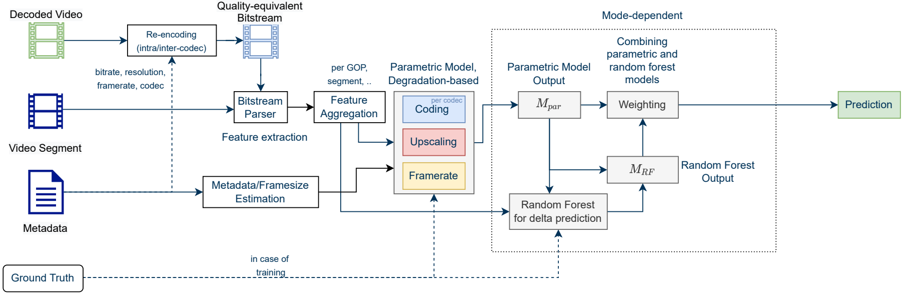
Audiovisual Technology Group, Technische Universität Ilmenau, Germany
VQEG

December 18, 2023

TECHNISCHE UNIVERSITÄT
ILMENAU

AVQBits: Introduction

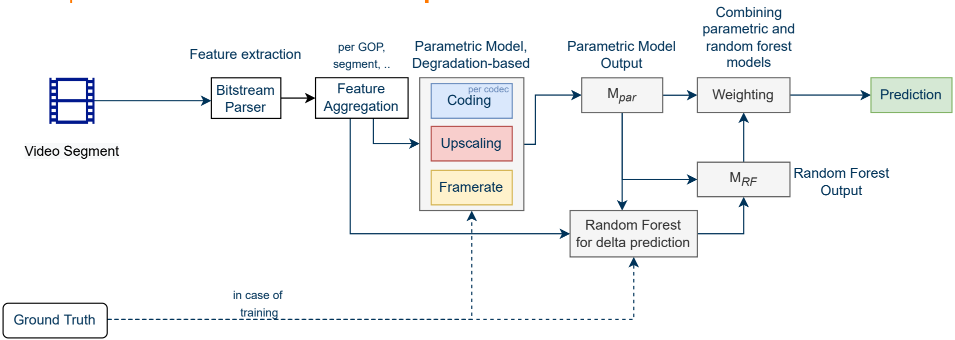
TECHNISCHE UNIVERSITÄT ILMENAU



- 4 Models [1, 2]
  - *AVQBits*|*M*3: Mode 3 (Principal model) → ITU-T Rec. P.1204.3 [3, 4]
  - *AVQBits*|*M*1: Mode 1
  - *AVQBits*|*M*0: Mode 0
  - *AVQBits*|*H*0: Hybrid No-reference Mode 0 (HYN0)
- Publicly available[1,2]

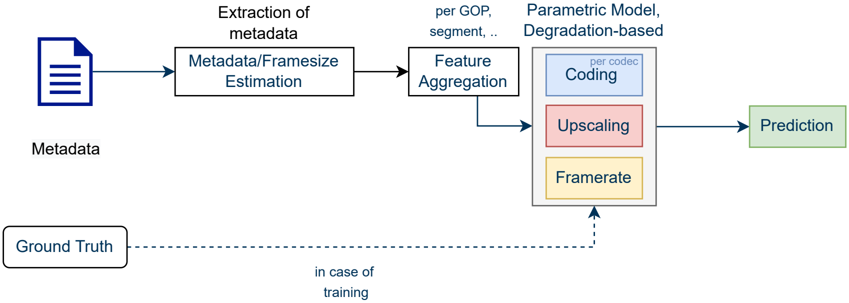
[1] http://git.avt-imt.de/bitstream_mode3_p1204_3
[2] http://git.avt-imt.de/p1204_3_extensions

AVQBits|M3: Model Description



- Input: entire bitstream
- Parametric/Core model: QP, display and coding resolution, display and coding framerate
- Machine-learning-based model: QP, average motion per-frame, horizontal motion, frame size + frame type, bitrate, resolution, framerate, Core model output
- Final prediction: $0.5 \cdot M_{par} + 0.5 \cdot M_{RF}$
- Standardized as ITU-T Rec. P.1204.3 [3]

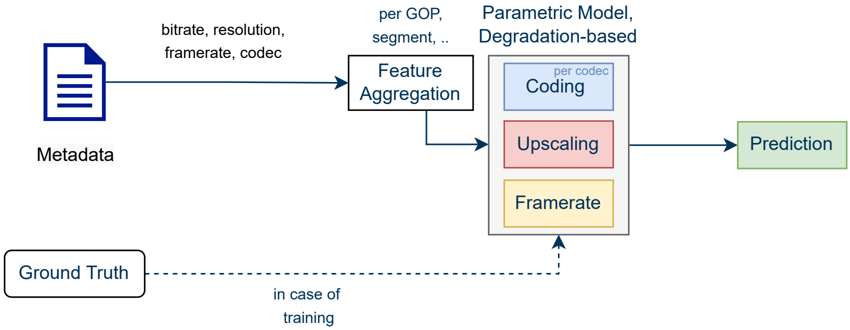
AVQBits|*M*1: Model Description



- Input: metadata (resolution, bitrate, framerate, video codec) + frame-type + frame-sizes
- Uses *AVQBits*|*M*3 Core Model
- Affected component: Quantization/Coding degradation
  - Only mode dependent component
  - QP estimation: metadata + frame-related features

AVQBits|M0: Model Description

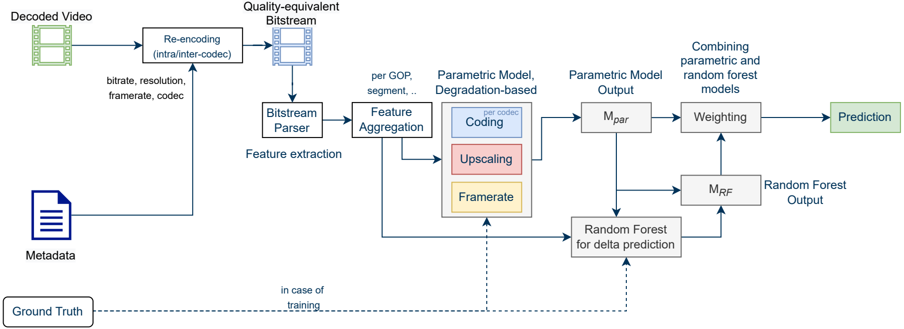
TECHNISCHE UNIVERSITÄT ILMENAU



- Input: metadata (resolution, bitrate, framerate, video codec)
- Uses *AVQBits|M3* Core Model
- QP estimation: metadata

AVQBits|H0: Model Description

TECHNISCHE UNIVERSITÄT ILMENAU



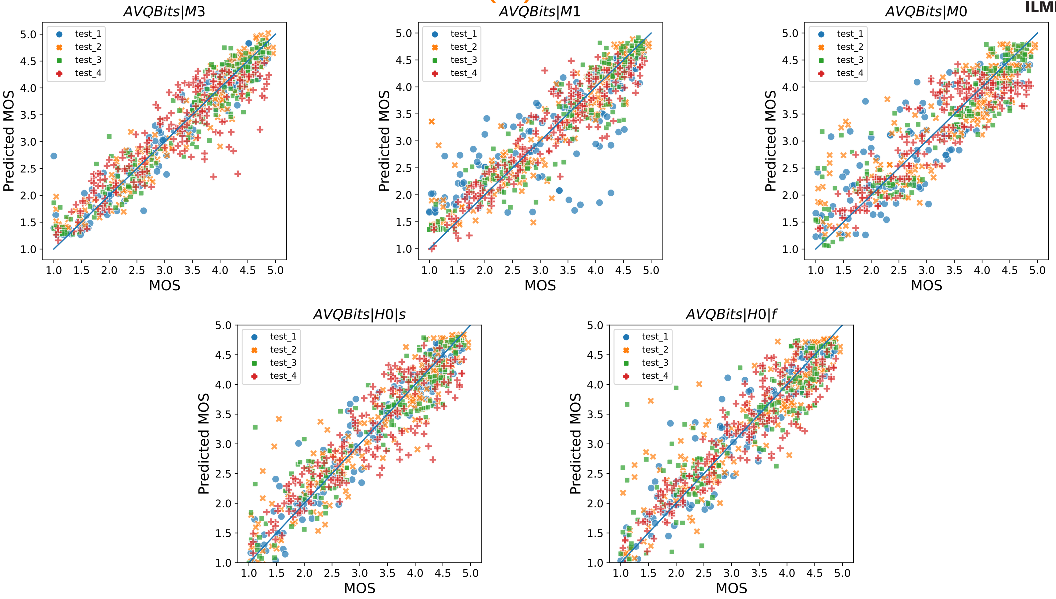
- Input: Distorted video (decoded pixels) + metadata
- Central idea
  - Step 1: "Quality equivalent bitstream" (QEB) creation: decoded pixels + metadata
  - Step 2: Quality prediction: Modified *AVQBits*|*M*3
- Two instances: codec used to create QEB
  - *AVQBits*|*H*0|*s*: same codec as original bitstream
  - *AVQBits*|*H*0|*f*: One pre-defined codec (e.g. H.265)

AVQBits: Model Evaluation (1)

- Dataset: AVT-VQDB-UHD-1; Metrics: Pearson correlation, RMSE
- *AVQBits* comparison
  - $AVQBits|M3$: best performing
  - $AVQBits|H0|s$: performs on-par with $AVQBits|M3$
  - $AVQBits|H0|f$: more sophisticated codec mapping → increased prediction accuracy
  - $AVQBits|M1$, $AVQBits|M0$: least well-performing
- Comparison with SoA models
  - VMAF: best performing pixel model
  - $AVQBits|M3$, $AVQBits|H0$: outperform VMAF
  - $AVQBits|M1$, $AVQBits|M0$: outperform best performing NR pixel models (BRISQUE, NIQE)
- P.NATS Phase 2: $AVQBits|M3$ → best performing model

AVQBits: Model Evaluation (2)

TECHNISCHE UNIVERSITÄT ILMENAU



AVQBits: Evaluation of Gaming Video Quality (1)

TECHNISCHE UNIVERSITÄT
ILMENAU

- Motivation
  - Advances in passive and interactive gaming services
  - Increase in popularity of gaming streaming services (e.g: Twitch.tv)
  - Need for video quality models for gaming content (cloud gaming)
- Evaluation using 4 datasets
  - GVS [5], KUGVD [6]: libx264 encoding
  - CGVDS [7]: NVENC (H.264); simulates encoding in cloud gaming services
  - Twitch [8]: proprietary Twitch encoding
  - GVS, KUGVD, CGVDS: publicly available
  - Maximum resolution: 1080p

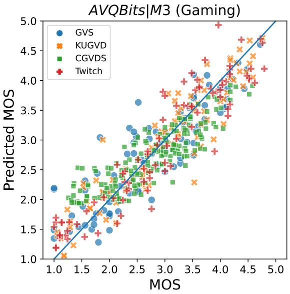
AVQBits: Evaluation of Gaming Video Quality (2)

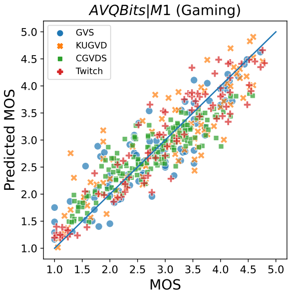
TECHNISCHE UNIVERSITÄT ILMENAU

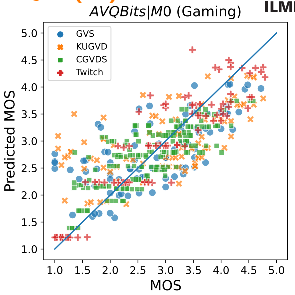
- *AVQBits* models evaluated out-of-the-box; no dedicated retraining
- *AVQBits* comparison
  - *AVQBits*|*M*3, *AVQBits*|*M*1, *AVQBits*|*H*0: perform on-par
  - *AVQBits*|*M*0: dedicated retraining → potential accuracy improvement
- Comparsison with SoA models
  - Best performing pixel models: VMAF (FR), NDNetGaming (NR, CGVDS)
  - *AVQBits*|*M*3, *AVQBits*|*M*1, *AVQBits*|*H*0: perform on-par with VMAF and NDNetGaming
  - *AVQBits*|*M*0: outperforms traditional NR models (BRISQUE, NIQE)

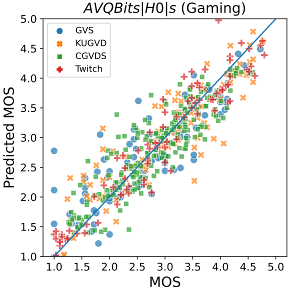
AVQBits: Evaluation of Gaming Video Quality (3)

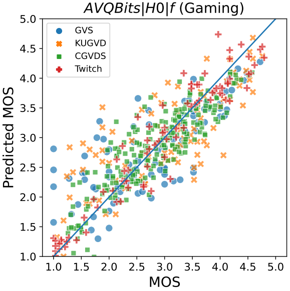
TECHNISCHE UNIVERSITÄT
ILMENAU











AVQBits: Evaluation of 360° Video Quality (1)

TECHNISCHE UNIVERSITÄT
ILMENAU

- Considered dataset: 360 Streaming Video Quality Dataset [3]
- 3 different tests
- Source content duration: 30 s
- Resolution: 1080p - 8K
- 2 different HMDs: HTC Vive (test_1), HTC Vive Pro (test_2 and test_3)

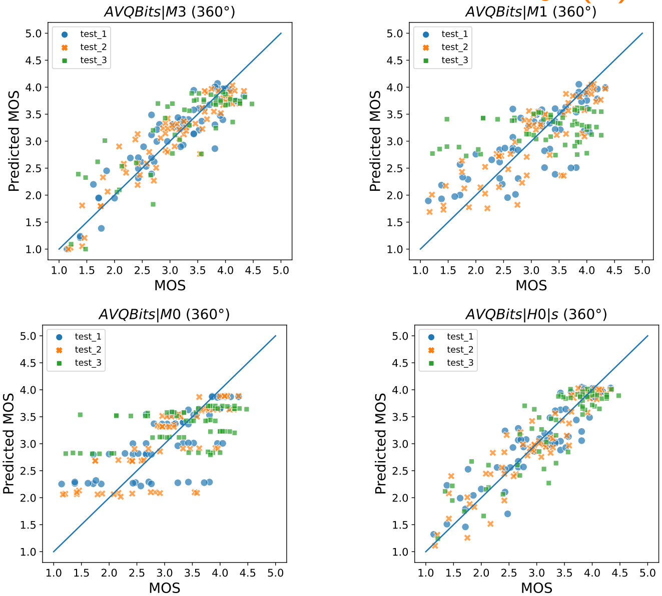
[3] *Fremerey, Göring, Ramachandra Rao, Huang, and Raake* 2020: "Subjective Test Dataset and Meta-data-based Models for 360° Streaming Video Quality"

AVQBits: Evaluation of 360° Video Quality (2)

TECHNISCHE UNIVERSITÄT
ILMENAU

- *AVQBits* models evaluated out-of-the-box; no dedicated retraining
- *AVQBits* comparison
  - *AVQBits*|*M*3, *AVQBits*|*H*0: perform on-par
  - *AVQBits*|*M*1, *AVQBits*|*M*0: lower accuracy; dedicated retraining → potential accuracy improvement
- Comparsison with SoA models
  - Best performing pixel models: VMAF_cc [10], VMAF, WS_SSIM
  - *AVQBits*|*M*3, *AVQBits*|*H*0: outperform VMAF, VMAF_cc and WS_SSIM

AVQBits: Evaluation of 360° Video Quality (3)



AVQBits: Overall Integral Quality Model

TECHNISCHE UNIVERSITÄT ILMENAU

- Based on ITU-T Rec. P.1203.3 [11]
  - Overall audiovisual quality (O.35)
    - Temporal effects related to video-quality fluctuations: per-1-sec audiovisual quality score + 3 additional factors (needed due to limited accuracy of per-1-sec audiovisual scores)
    - Limited accuracy of per-1-sec audiovisual scores: reverse-engineered from overall session quality scores
- New O.35
  - per-1-sec audiovisual quality scores capture temporal effects related to video-quality fluctuations
  - Rationale → per-1-sec and per-segment scores of *AVQBits* more accurate than ITU-T Rec. P.1203.1 [12]
  - Added as appendix to ITU-T Rec. P.1204.3 [3]

AVQBits: Evaluation of Overall Integral Quality (1)

TECHNISCHE UNIVERSITÄT ILMENAU

- PNATS-UHD-1-Long dataset
  - Created as part of P.NATS Phase 2 competition
  - 5 tests conducted by 4 different labs: 1 test conducted by TUIL
  - Source audiovisual content duration: 1-5 min
  - Context: Mobile + TV
  - Distortions: video quality switches, initial loading delay, stalling events

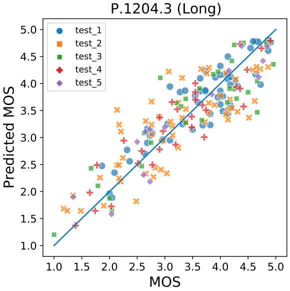
AVQBits: Evaluation of Overall Integral Quality (2)

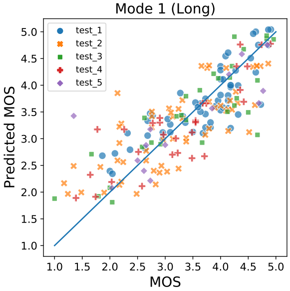
TECHNISCHE UNIVERSITÄT ILMENAU

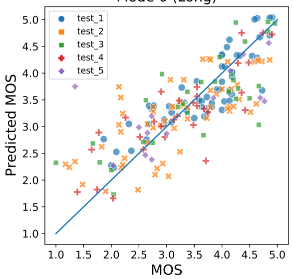
- *AVQBits*|*M*3: best performing model
- *AVQBits*|*H*0
  - Prediction accuracy lower than *AVQBits*|*M*3 unlike short-term video quality
  - Additional mapping for per-1-sec scores may increase prediction accuracy
- *AVQBits*|*M*1, *AVQBits*|*M*0
  - Lower prediction accuracy
  - Possible reason: per-1-sec quality score same as per-segment score $\rightarrow$ no optimal handling of effect of temporal efforts related to video-quality fluctuations

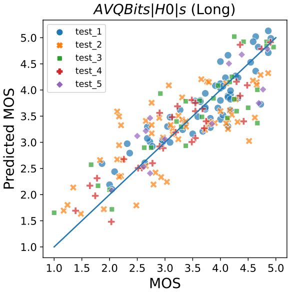
AVQBits: Evaluation of Overall Integral Quality (3)

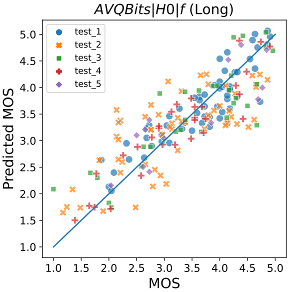
TECHNISCHE UNIVERSITÄT ILMENAU











Thank you for your attention

TECHNISCHE UNIVERSITÄT
ILMENAU

References I

TECHNISCHE UNIVERSITÄT ILMENAU

[1] Rakesh Rao Ramachandra Rao, Steve Göring, and Alexander Raake. "AVQBits - Adaptive Video Quality Model Based on Bitstream Information for Various Video Applications". In: *IEEE Access* 10 (2022).

[2] 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 IEEE International Conference on Quality of Multimedia Experience (QoMEX)*. Athlone, Ireland, May 2020.

[3] 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.

References II

TECHNISCHE UNIVERSITÄT ILMENAU

[4] 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).

[5] 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.

[6] Nabajeet Barman et al. "No-reference video quality estimation based on machine learning for passive gaming video streaming applications". In: *IEEE Access* 7 (2019), pp. 74511–74527.

References III

TECHNISCHE UNIVERSITÄT ILMENAU

[7] Saman Zadtootaghaj et al. "Quality Estimation Models for Gaming Video Streaming Services Using Perceptual Video Quality Dimensions". In: *Proceedings of the 11th International Conference on Multimedia Systems*. ACM. 2020.

[8] Rakesh Rao Ramachandra Rao et al. "A Large-scale Evaluation of the bitstream-based video-quality model ITU-T P.1204.3 on Gaming Content". In: *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. 2020.

[9] Stephan Fremerey et al. "Subjective Test Dataset and Meta-data-based Models for 360° Streaming Video Quality". In: *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE. 2020.

References IV

TECHNISCHE UNIVERSITÄT
ILMENAU

[10] Steve Göring, Christopher Krämmer, and Alexander Raake. "cencro – Speedup of Video Quality Calculation using Center Cropping". In: *2019 IEEE ISM*. Dec. 2019, pp. 1–8.

[11] ITU-T Rec. P.1203.3. *Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - Quality integration module*. Geneva, Switzerland: International Telecommunication Union, 2020.

[12] ITU-T Rec. P.1203.1. *Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport – Video quality estimation module*. Geneva, Switzerland: International Telecommunication Union, 2019.

References V

TECHNISCHE UNIVERSITÄT ILMENAU

[13] John Allnatt. "Subjective rating and apparent magnitude". In: *International Journal of Man-Machine Studies* 7.6 (1975), pp. 801–816.

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

[15] Nils Olof Johannesson. "The ETSI Computation Model: A Tool for Transmission Planning of Telephone Networks". In: *IEEE Communications Magazine* Jan. (1997), pp. 70–79.

[16] Sebastian Möller. *Assessment and Prediction of Speech Quality in Telecommunications*. Springer Science & Business Media, 2000.

References VI

TECHNISCHE UNIVERSITÄT ILMENAU

[17] ITU–T Rec. G.107. *The E–Model, a Computational Model for Use in Transmission Planning*. International Telecommunication Union. CH–Geneva, 2009.

[18] ITU-T. *P.1401 : Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*. Tech. rep. Int. Telecommunication Union, 2014.

[19] ITU-T. *RECOMMENDATION ITU-R BT.500-13 – Methodology for the subjective assessment of the quality of television pictures*. Tech. rep. Int. Telecommunication Union, 2014.

Back-up

AVQBits|M3: Parametric Part – Core Model

TECHNISCHE UNIVERSITÄT ILMENAU

- Degradation-based modeling approach [13, 14, 15, 16, 17]
- 3 types of degradation
  - Quantization/Coding degradation (D_q): $f(QP)$
  - Upscaling degradation (D_u): $f(coding\_res, display\_res)$
  - Temporal degradation (D_t): $f(coding\_framerate, display\_framerate)$
- Degradation values expressed on a 0 to 100 scale
  - Compensates for the compression of the 5-point ACR scale at the scale ends

AVQBits|*M3*: Parametric Part – Prediction

TECHNISCHE UNIVERSITÄT
ILMENAU

$$M_{p_{[0,100]}} = 100 - (D_q + D_u + D_t) \quad (1)$$

$$M_{p_{[1,4.5]}} = MOSfromR(M_{p_{[0,100]}}) \quad (2)$$

$$M_{par} = scaleto5(M_{p_{[1,4.5]}}) \quad (3)$$

AVQBits|*M*3: Machine Learning Part

TECHNISCHE UNIVERSITÄT ILMENAU

- Estimates the "residual"

$$target\_residual = MOS - M_{par} \quad (4)$$

- Based on Random Forests (RF)
  - Hyper-parameters: $trees = 20$ and $depth = 8$
  - One RF for PC/TV and Mobile/Tablet
- Meta-features
  - QP
  - Average motion per-frame
  - Motion in the x-direction (horizontal motion)
  - Frame sizes + frame type
  - Codec, bitrate, resolution, framerate, M_{par}

AVQBits|*M*3: Final Prediction

TECHNISCHE UNIVERSITÄT ILMENAU

- Prediction of the ML-part

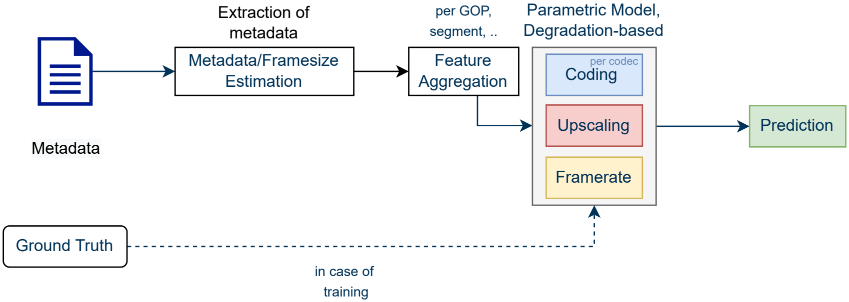
$$M_{RF} = M_{par} + predicted\_residual \tag{5}$$

- Final prediction

$$Prediction = w \cdot M_{par} + (1 - w) \cdot M_{RF} \tag{6}$$

where, $w = 0.5$

AVQBits|M1: Model Description



- Available input information
  - Metadata (bitrate, resolution, framerate and codec) + frame size and frame type information
- Uses *AVQBits|M3* Core Model
- Affected component: Quantization/Coding degradation
  - Only mode dependent component
  - QP estimation: metadata (resolution, bitrate, framerate, video codec) + frame-related features

AVQBits|*M*1: QP Prediction

TECHNISCHE UNIVERSITÄT ILMENAU

- Adaptation of the Quantization/Coding degradation
  - True QP unavailable → estimation of QP using metadata and frame-related features
- QP Prediction

$$QP_{pred} = a_{qp\_m1} + b_{qp\_m1} \cdot log(ms\_nl) + c_{qp\_m1} \cdot log(resolution) + d_{qp\_m1} \cdot log(framerate) + e_{qp\_m1} \cdot log(fsratio) \quad (7)$$

AVQBits: Model Evaluation

TECHNISCHE UNIVERSITÄT ILMENAU

Table: Comparison between *AVQBits* and SoA models.

Model	RMSE	PCC	SROCC	Kendall	R^2 Score
VMAF	0.531	0.880	0.889	0.721	0.774
Brisque	0.653	0.815	0.838	0.653	0.660
NIQE	1.009	0.432	0.445	0.301	0.187
PSNR	1.109	0.131	0.682	0.531	0.017
SSIM	0.956	0.520	0.761	0.569	0.270
MS-SSIM	0.896	0.599	0.752	0.563	0.358
ADM2	0.580	0.855	0.874	0.698	0.731
VIFP	0.757	0.736	0.756	0.562	0.542
$AVQBits\|M3$	0.306	0.962	0.948	0.804	0.925
$AVQBits\|M1$	0.486	0.901	0.904	0.738	0.812
$AVQBits\|M0$	0.503	0.894	0.891	0.701	0.799
$AVQBits\|H0\|s$	0.373	0.943	0.935	0.778	0.889
$AVQBits\|H0\|f$	0.439	0.920	0.914	0.749	0.846

NOTE: RMSE calculated on a 5-point ACR scale after linear mapping between model output and subjective scores as described in ITU-T P.1401 [18]

AVQBits: Evaluation of Gaming Video Quality (1)

TECHNISCHE UNIVERSITÄT
ILMENAU

Table: Overview of the used gaming datasets.

Dataset	GVS [5]	KUGVD [6]	CGVDS [7]	Twitch [8]
# Sources	6	6	15	36
Resolution	480p, 720p, 1080p	480p, 720p, 1080p	480p, 720p, 1080p	160p, 360p, 480p, 720p, 900p, 1080p
Framerate	30 fps	30 fps	20, 30, 60 fps	30, 60 fps
# PVS	90	90	72 * 5	90
Duration	30 s	30 s	30 s	30 s
# Participants	25	17	>100 (5 tests)	29
Encoder	ffmpeg x264	ffmpeg x264	NVENC (H.264)	H.264
Encoding mode	CBR	CBR	CBR	Twitch default
Preset	veryfast	veryfast	llhq	Twitch default

AVQBits: Evaluation of Gaming Video Quality (2)

TECHNISCHE UNIVERSITÄT ILMENAU

Table: Comparison between *AVQBits* and SoA models.

Dataset	Model	RMSE	PCC	SROCC	Kendall	R^2 Score
GVS	PSNR	0.63	0.74	0.74	0.57	0.55
GVS	SSIM	0.57	0.80	0.80	0.61	0.62
GVS	VMAF	0.47	0.87	0.86	0.69	0.75
GVS	NIQE	0.64	0.77	0.71	0.53	0.52
GVS	*AVQBits*|*M*3	0.45	0.88	0.87	0.69	0.77
GVS	*AVQBits*|*M*1	0.42	0.89	0.87	0.71	0.79
GVS	*AVQBits*|*M*0	0.69	0.67	0.65	0.49	0.45
GVS	*AVQBits*|*H*0|*s*	0.48	0.86	0.86	0.69	0.74
GVS	*AVQBits*|*H*0|*f*	0.62	0.75	0.73	0.56	0.56
KUGVD	PSNR	0.62	0.80	0.84	0.67	0.64
KUGVD	SSIM	0.48	0.89	0.91	0.74	0.79
KUGVD	VMAF	0.41	0.92	0.92	0.77	0.85
KUGVD	NIQE	0.55	0.85	0.84	0.66	0.72
KUGVD	*AVQBits*|*M*3	0.39	0.93	0.92	0.77	0.86
KUGVD	*AVQBits*|*M*1	0.50	0.87	0.86	0.69	0.76
KUGVD	*AVQBits*|*M*0	0.84	0.59	0.57	0.41	0.35
KUGVD	*AVQBits*|*H*0|*s*	0.46	0.90	0.89	0.72	0.80
KUGVD	*AVQBits*|*H*0|*f*	0.65	0.78	0.76	0.58	0.61

AVQBits: Evaluation of Gaming Video Quality (3)

TECHNISCHE UNIVERSITÄT ILMENAU

Table: Comparison between *AVQBits* and SoA models.

Dataset	Model	RMSE	PCC	SROCC	Kendall	R^2 Score
CGVDS	PSNR	0.60	0.64	0.65	0.47	0.41
CGVDS	SSIM	0.59	0.67	0.78	0.60	0.45
CGVDS	VMAF	0.38	0.88	0.87	0.69	0.77
CGVDS	NIQE	0.66	0.54	0.56	0.41	0.29
CGVDS	$AVQBits\|M3$	0.38	0.85	0.84	0.65	0.72
CGVDS	$AVQBits\|M1$	0.36	0.90	0.88	0.70	0.78
CGVDS	$AVQBits\|M0$	0.47	0.78	0.75	0.56	0.60
CGVDS	$AVQBits\|H0\|s$	0.36	0.89	0.88	0.70	0.79
CGVDS	$AVQBits\|H0\|f$	0.38	0.87	0.87	0.68	0.76
Twitch	NIQE	0.96	0.24	0.11	0.17	0.04
Twitch	$AVQBits\|M3$	0.40	0.93	0.93	0.77	0.87
Twitch	$AVQBits\|M1$	0.37	0.94	0.93	0.77	0.89
Twitch	$AVQBits\|M0$	0.43	0.92	0.89	0.71	0.85
Twitch	$AVQBits\|H0\|s$	0.31	0.96	0.95	0.82	0.92
Twitch	$AVQBits\|H0\|f$	0.30	0.96	0.95	0.81	0.92

NOTE: RMSE calculated on a 5-point ACR scale after linear mapping between model output and subjective scores as described in ITU-T P.1401 [18]

AVQBits: Evaluation of 360° Video Quality (1)

TECHNISCHE UNIVERSITÄT ILMENAU

Table: Details of the 360 Streaming Video Quality Dataset [9].

Test	test_1	test_2	test_3
# Sources	8	8	7
Codecs	H.265	H.265	H.265
Resolution	1080p, 4K	1080p, 4K	4K, 6K, 8K
Framerate	30 fps	30 fps	30 fps
# PVS	64	64	63
Duration	30 s	30 s	30 s
# Participants	27	27	27
Display	HTC Vive	HTC Vive Pro	HTC Vive Pro
Test Method	5-point ACR [19]	5-point ACR [19]	5-point ACR [19]

AVQBits: Evaluation of 360° Video Quality (2)

TECHNISCHE UNIVERSITÄT
ILMENAU

Table: Comparison between *AVQBits* and SoA models.

Model	**RMSE**	**PCC**	**SROCC**	**Kendall**
VMAF_cc	0.384	0.898	0.872	0.700
VMAF	0.431	0.870	0.834	0.664
ADM2	0.494	0.825	0.819	0.640
WS_SSIM	0.500	0.820	0.864	0.671
VIFP	0.554	0.773	0.656	0.502
WS_PSNR	0.598	0.729	0.767	0.582
SSIM	0.622	0.702	0.730	0.563
PSNR	0.762	0.489	0.627	0.469
$AVQBits\|M3$	0.377	0.894	0.870	0.679
$AVQBits\|M1$	0.581	0.709	0.677	0.497
$AVQBits\|M0$	0.627	0.658	0.686	0.401
$AVQBits\|H0\|s$	0.356	0.906	0.886	0.695

NOTE: RMSE calculated on a 5-point ACR scale after linear mapping between model output and subjective scores as described in ITU-T P.1401 [18]

AVQBits: Overall Integral Quality Model (1)

TECHNISCHE UNIVERSITÄT ILMENAU

- Based on ITU-T Rec. P.1203.3 [11]
- Overall audiovisual quality (O.35) according to ITU-T Rec. P.1203.3
  - $f(per - 1 - sec\ quality, negBias, oscComp, adaptComp)$
  - $negBias, oscComp, adaptComp$ $\rightarrow$ take into account certain temporal effects related to video-quality fluctuations
- New O.35 $\rightarrow f(per - 1 - sec\ quality)$
  - Rationale $\rightarrow$ per-1-sec and per-segment scores of *AVQBits* more accurate than ITU-T Rec. P.1203.1 [12]

AVQBits: Overall Integral Quality (2)

TECHNISCHE UNIVERSITÄT ILMENAU

Table: Details of the P.NATS-UHD-1-Long dataset [1].

Test	test_1	test_2	test_3	test_4	test_5
# Sources	60	60	30	30	14
Codecs	H.264, H.265, VP9	H.264, H.265, VP9	H.264, H.265, VP9	H.264, H.265, VP9	H.264, H.265, VP9
Resolution	240p - 1440p	360p - 2160p	240p - 1440p	360p - 2160p	240p - 1440p
Framerate	15 - 60 fps	24 - 60 fps	15 - 60 fps	24 - 60 fps	15 - 60 fps
# PVS	60	60	30	30	14
Duration	1 min	1 min	2 min	2 min	5 min
Display device	Mobile	TV	Mobile	TV	Mobile
# Participants	24	31	24	31	36
Test Method	5-point ACR [19]	5-point ACR [19]	5-point ACR [19]	5-point ACR [19]	5-point ACR [19]

AVQBits: Evaluation of Overall Integral Quality

TECHNISCHE UNIVERSITÄT
ILMENAU

Table: Performance of *AVQBits* on the P.NATS-UHD-1-Long dataset [1].

Test	Model	RMSE	PCC	SROCC	Kendall	R^2 Score
All	$AVQBits|M3$ / P.1204.3	0.479	0.864	0.844	0.660	0.747
All	$AVQBits|M1$	0.596	0.780	0.787	0.602	0.608
All	$AVQBits|M0$	0.694	0.686	0.683	0.500	0.471
All	$AVQBits|H0|s$	0.570	0.797	0.768	0.584	0.635
All	$AVQBits|H0|f$	0.582	0.787	0.756	0.572	0.619

NOTE: RMSE calculated on a 5-point ACR scale after linear mapping between model output and subjective scores as described in ITU-T P.1401 [18]