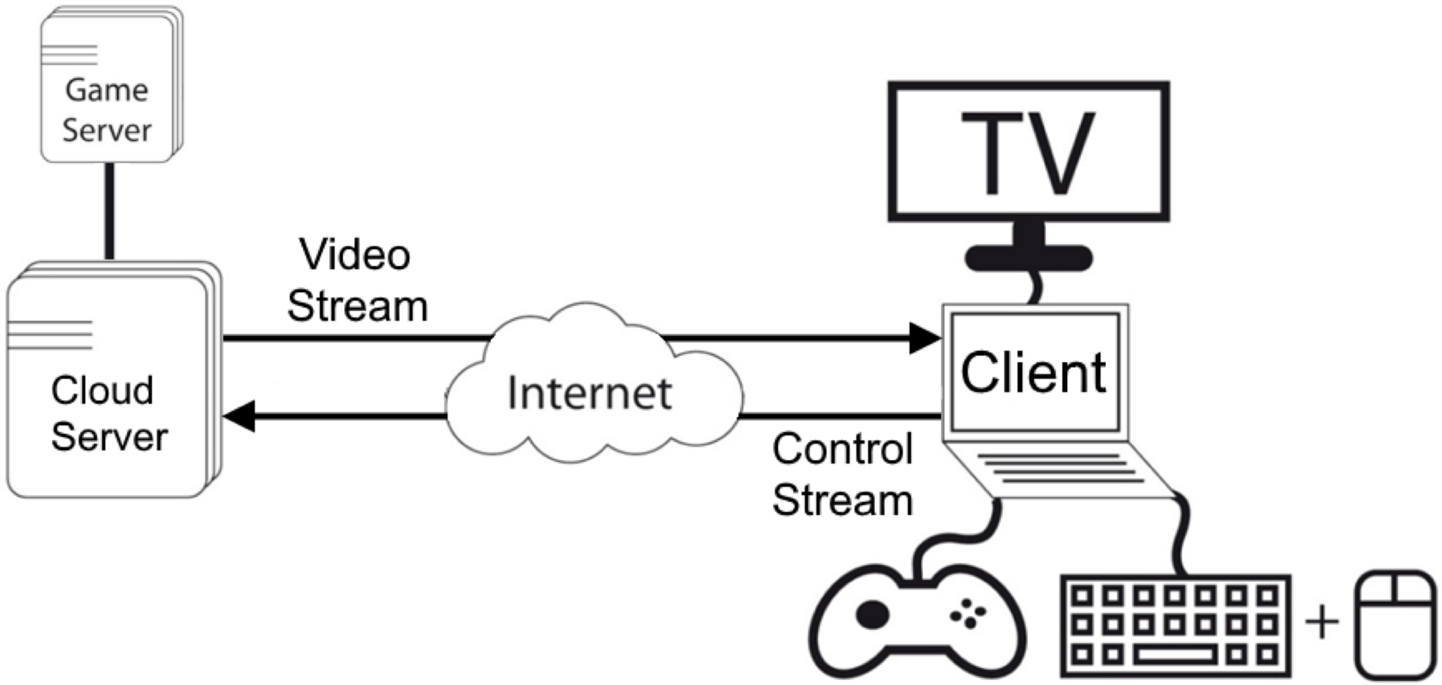
Updates on Standardization Activities with Respect to Gaming Quality of Experience

Steven Schmidt

Quality and Usability Lab – Berlin Institute of Technology

VQEG Meeting - Shenzhen, China, 2019

What is Cloud Gaming?



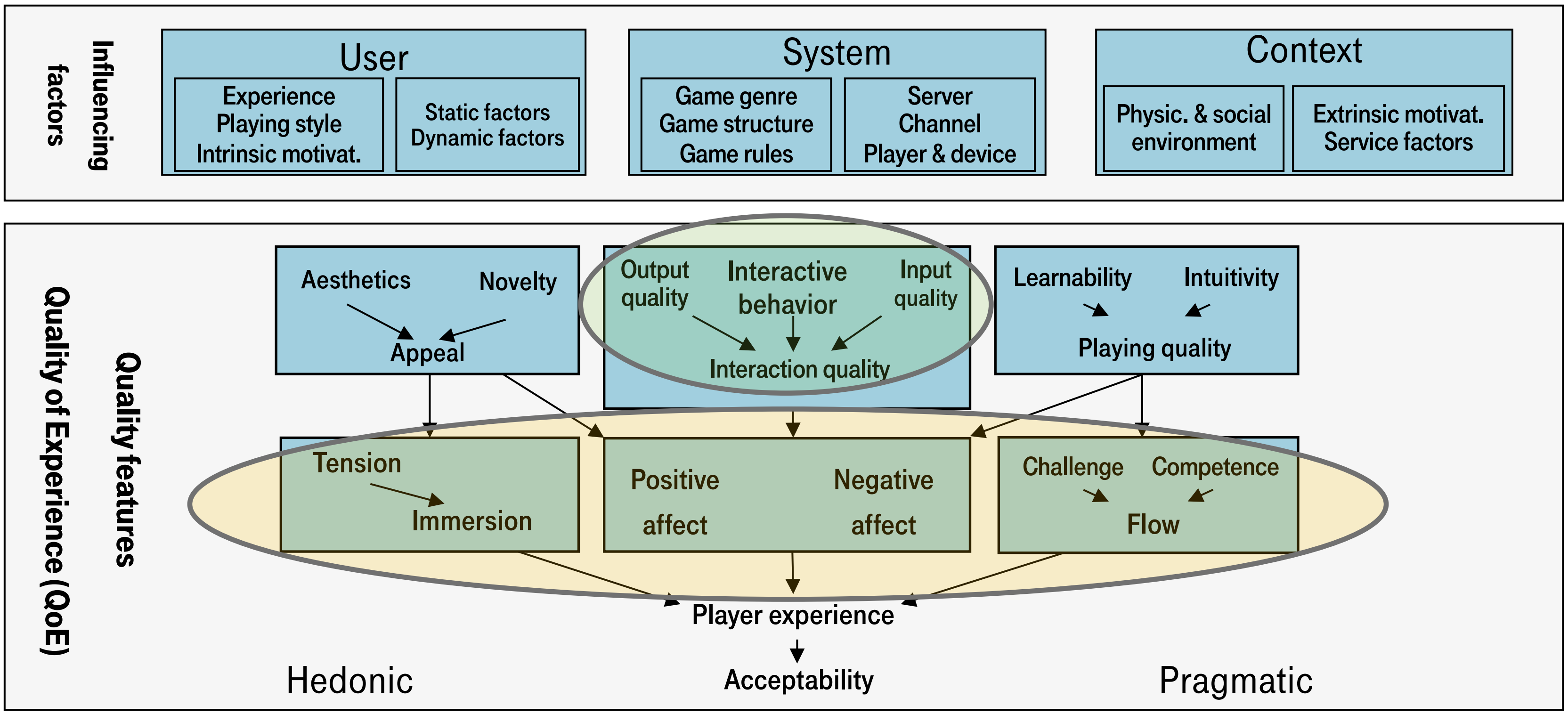
Quality of Experience

Cloud Gaming

- 2.3 billion gamers will spend $137.9 billion on video games in 2018[1]
- Companies in the past could not provide acceptable QoE
- Creation of **3 work items** in ITU-T SG-12
- ITU-T Rec. G.1032 (10/2017) – G.QoE-gaming:
  - **Influence factors** on gaming quality of experience
- ITU-T Rec. P.809 (05/2018) – P.GAME:
  - **Subjective evaluation** methods for gaming quality
- Future ITU-T Rec. G.OMG (studied in Q.13/12):
  - **Opinion model** for gaming applications

Quality & Usability Lab

What is Gaming Quality of Experience?

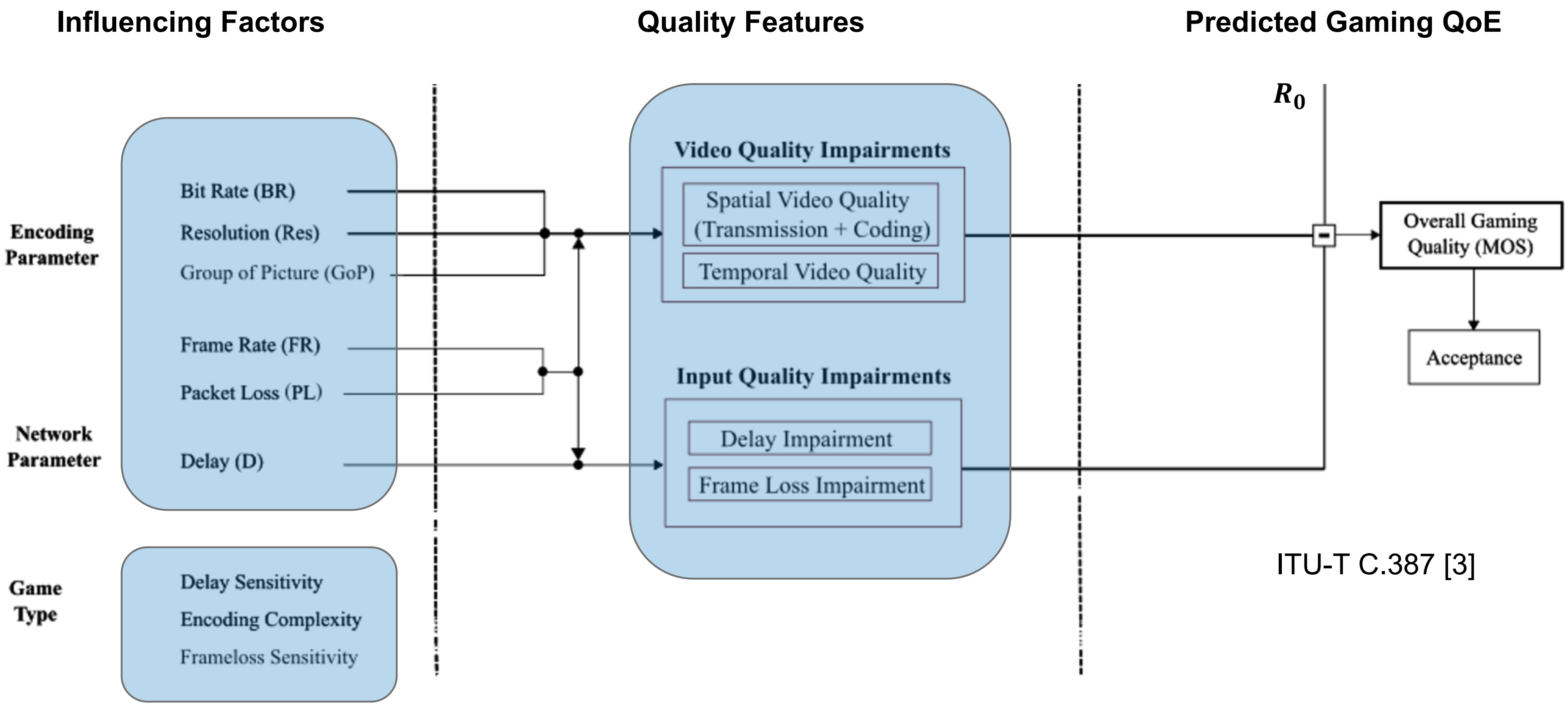


[2]

Future ITU-T Rec. G.OMG [8]

- **Predict** overall gaming **QoE based on impairment factors** using encoding and network parameters
- Two modes depending on available information about game content
- **Scope**:
  - Considering relevant factors identified in ITU-T Rec. G.1032
  - Network planning tool (infrastructure and resource allocation)
  - Target services: cloud gaming
  - Target group: non-professional gamer
  - Not: VR gaming, mobile devices, social aspects (but might be applicable)
  - Not: influence of the design of games or the motivation of users to play

Future ITU-T Rec. G.OMG

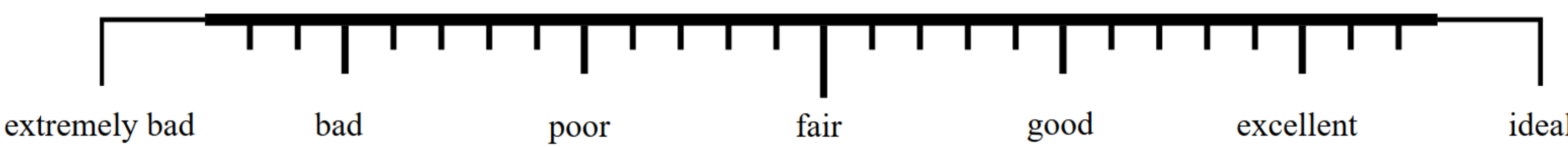


ITU-T C.387 [3]

Assessment of Quality Features

- Guidlines for assessment are given in ITU-T Rec. P.809
- Aggred to use 7-point continous rating scales [ITU-T Rec. P.851]
- Adaptation of Game Experience Questionnaire [4]
- Video Quality dimensions by Schiffner [5]
- Overall Gaming QoE

How do you rate the overall quality of your gaming experience?

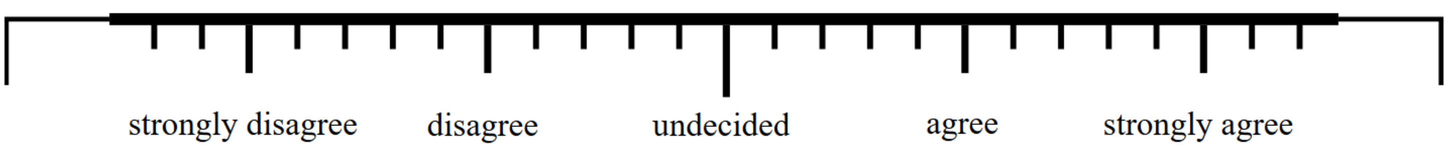


Quality & Usability Lab

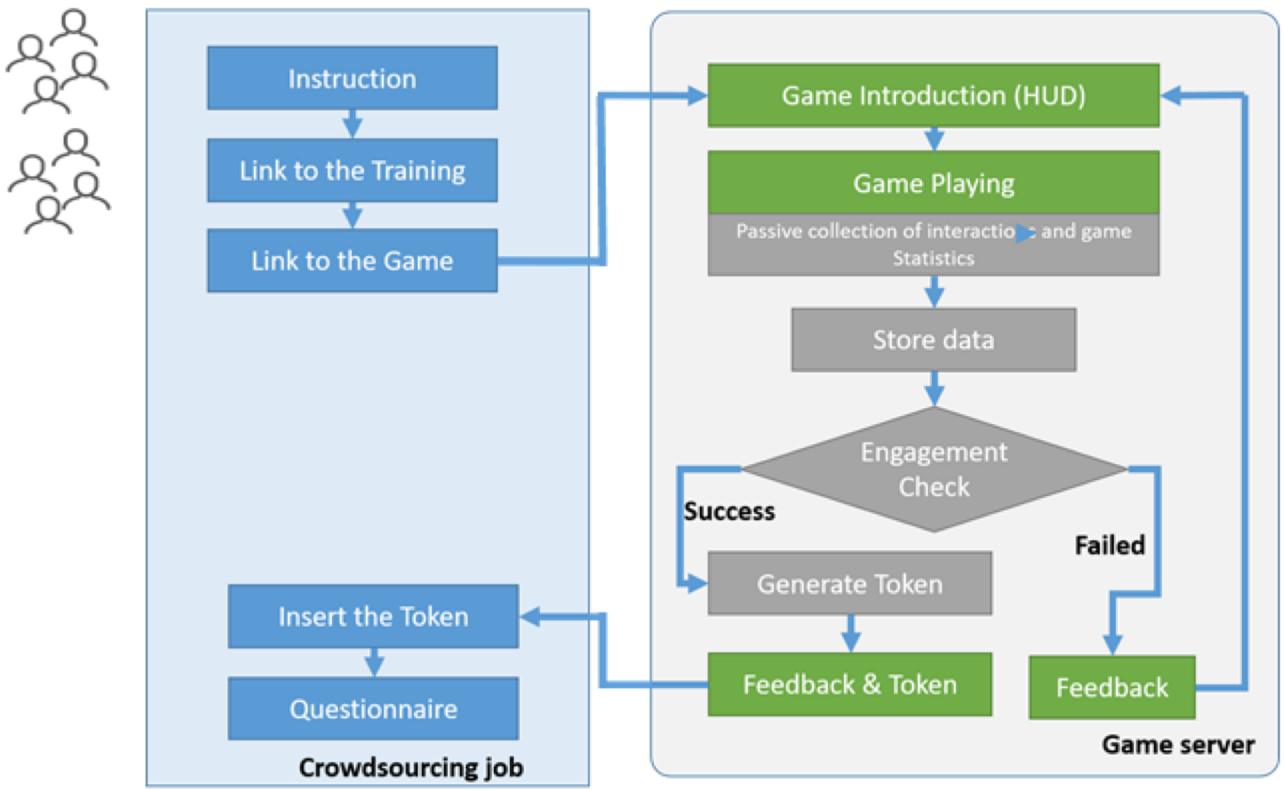
Assessment of Quality Features

- Development of new questionnaire assessing Input Quality (Playability)
  - Pool of items created based on literature and expert interviews
  - Develoment of web-based games (JS) including typical degradations
  - Crowdsourcing approach used to gather ratings from many diverse users
  - 2-level factor analysis (EFA+CFA)
  - 3 sub-dimension: controllability, immediate feedback, responsiveness

I received immediate feedback on my actions.



Assessing gaming QoE using Crowdsourcing



ITU-T C.376 [6]

Quality & Usability Lab

Related documents and dataset for G.OMG

- **Requirement specification** submitted in May 2018 [ITU-T C.200]
- **Data assessment** process will be discussed in Dec 2018 [ITU-T C.293]
- First Draft for an **Opinion Model Predicting Gaming QoE** in May 2019 [ITU-T C.387]

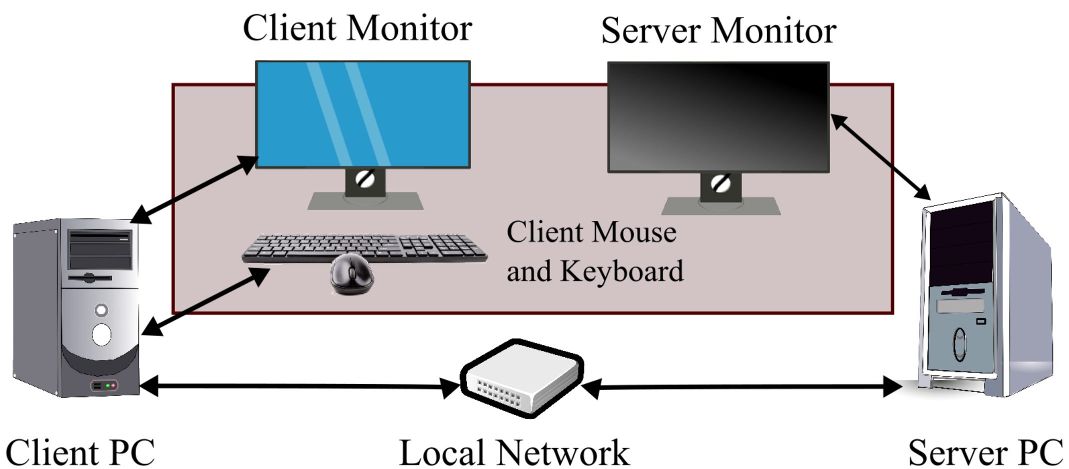
- Interactive dataset:
  - 9 different games
  - 1080p: 2, 4, 6, 50 mbps
  - Frame Rate: 60, 30, 20, 10 fps
  - Frame Loss Rate: 0, 10, 15, 25, 30 %
  - Delay: 0, 25, 50, 100, 200, 400 ms

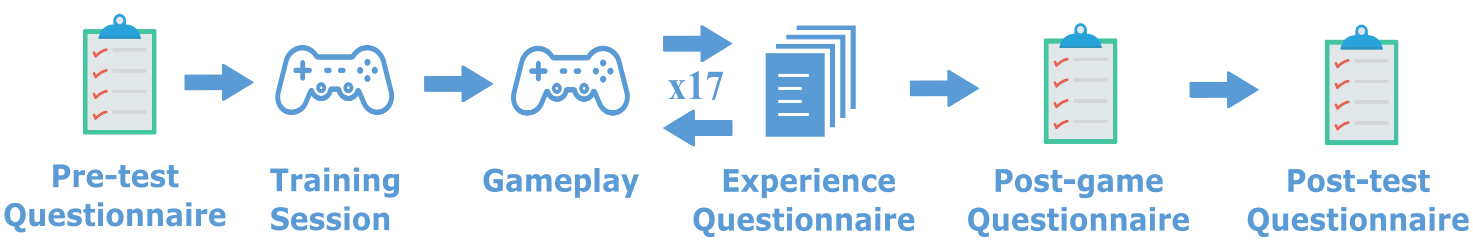
- Passive dataset:
  - 9 (now 18) different games
  - 480p: 0.3, 1, 2, 4 mbps
  - 720p: 1, 2, 4, 50 mbps
  - 1080p: 2, 4, 6, 50 mbps
  - 20, 30, 60 fps
  - (UDP packet loss + GoP)

Factors and application ranges

Application information	Value range, unit
Sequence duration (secs)	30 (passive), 90 (interactive)
Screen size	24″
Input devices	Mouse and keyboard
Packetization	RTSP (over RTP/UDP/IP)
Video codec	NVENC
Resolution	480p, 720p, 1080p
Coded video bitrate (kbps)	300-50000
Frame rate (fps)	10, 20, 30, 60
Pre-set	llhq (low latency, high quality)
Encoding Mode	CBR
Video Compression	Standard H.264, Main 4.0
Audio codec	AC3
Coded audio bitrate (kbps)	192 (stereo)
Audio sample rate (Hz)	48,000
Packet loss degradation	uniform loss (0-1 %)

Setup and Test Structure





Data Transformation

Continuous 7-points ratings has to been transformed to the R-scale

1. Transformation of extended continuous 7-point ratings (EC) to 5-point ACR ratings using the transformation presented in [7]:

$$\widehat{MOS}_{ACR} = f(\widehat{MOS}_{EC}) = -0.0262 \cdot \widehat{MOS}_{EC}^{3} + 0.2368 \cdot \widehat{MOS}_{EC}^{2} + 0.1907 \cdot \widehat{MOS}_{EC} + 1$$

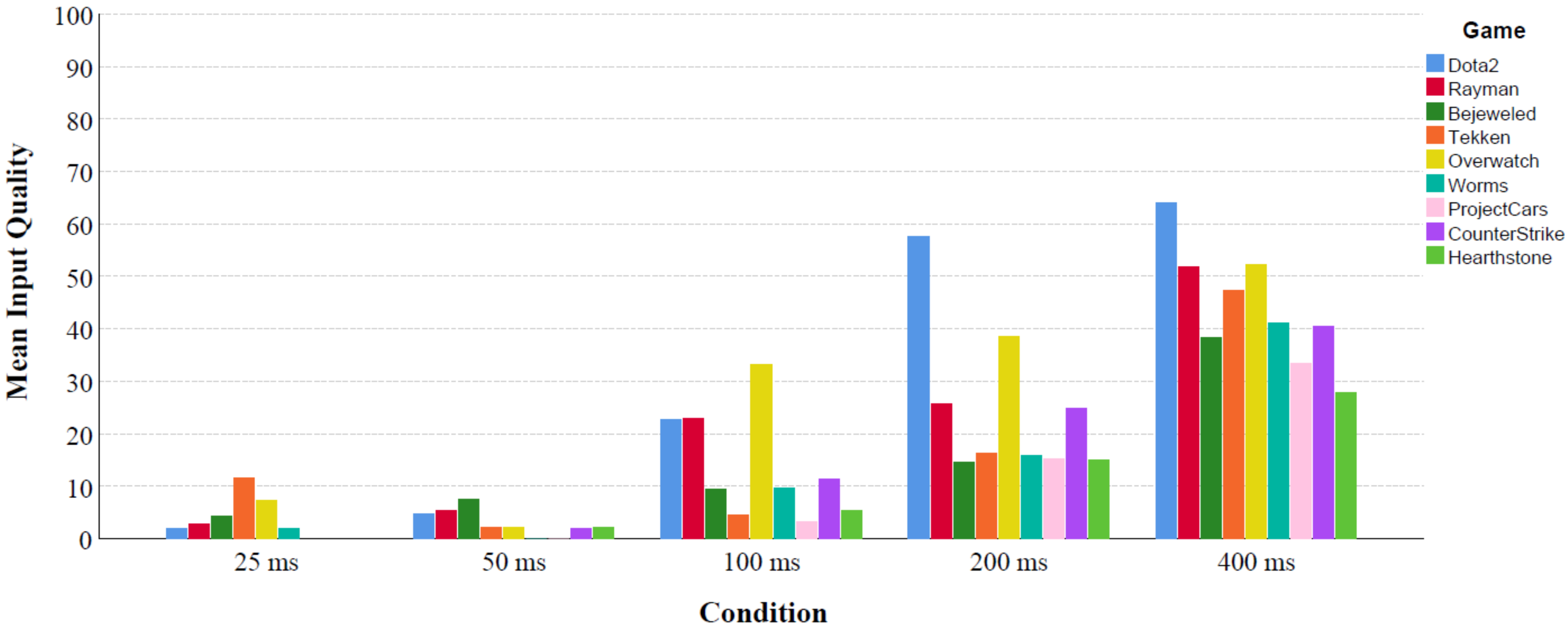
2. Normalize MOS values as follows, as some condition means were larger than 4.5:

$$MOS_{norm,i} = \frac{MOS_i - 1}{MOS_{max} - 1} \cdot 3.5 + 1$$

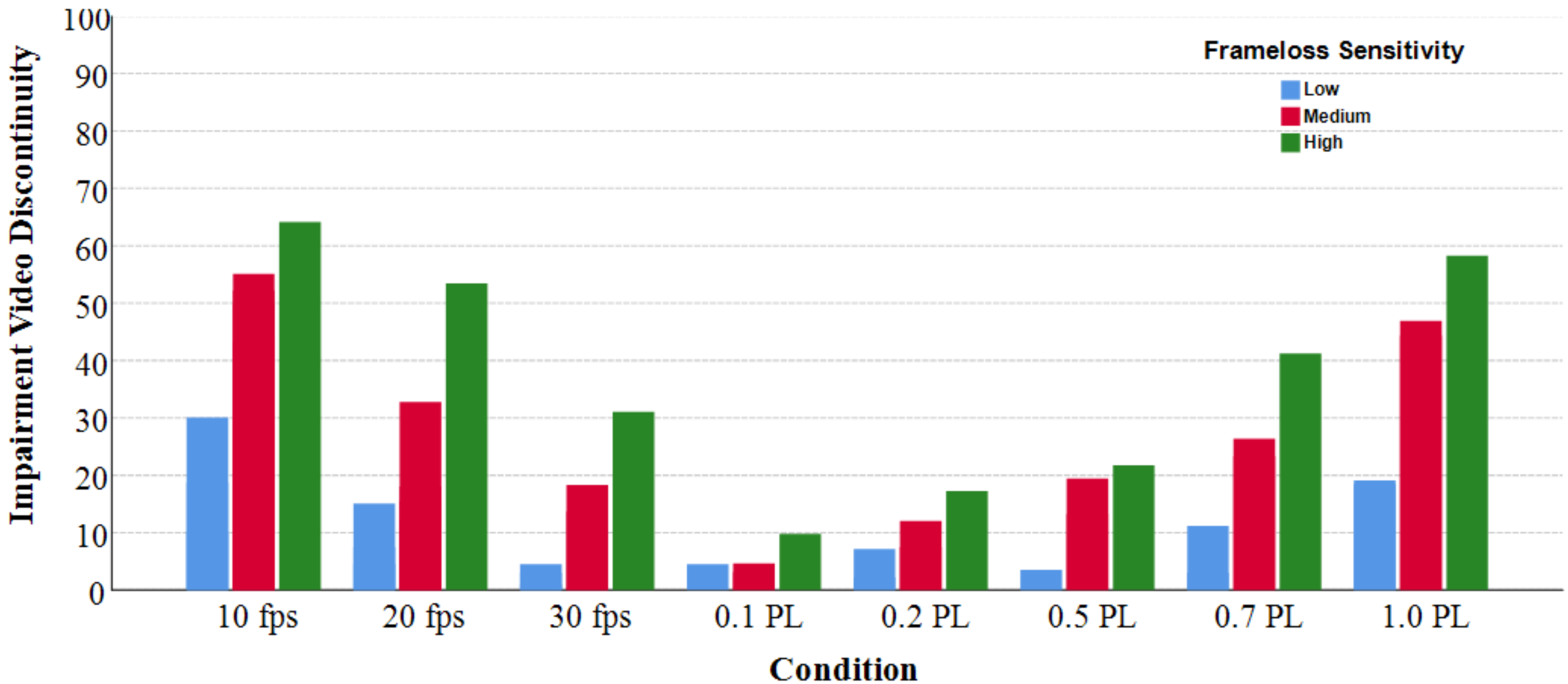
3. Calculation of R-value according to ITU-T Rec. 107
4. Calculation of Impairment (differential R-values from reference):

$$I_{factor}(condition) = R_{max,factor} - R_{factor}(condition)$$

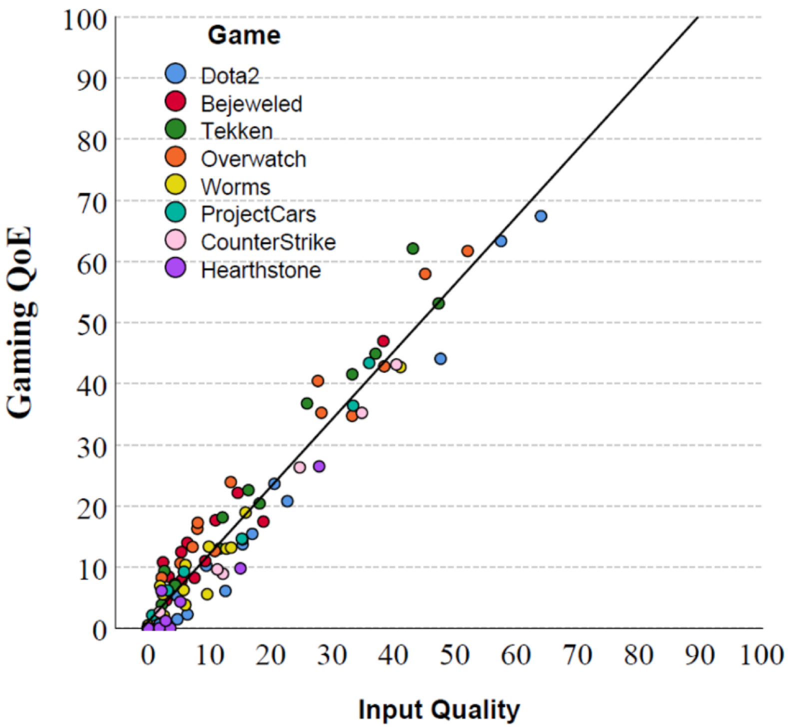
Impact of network delay

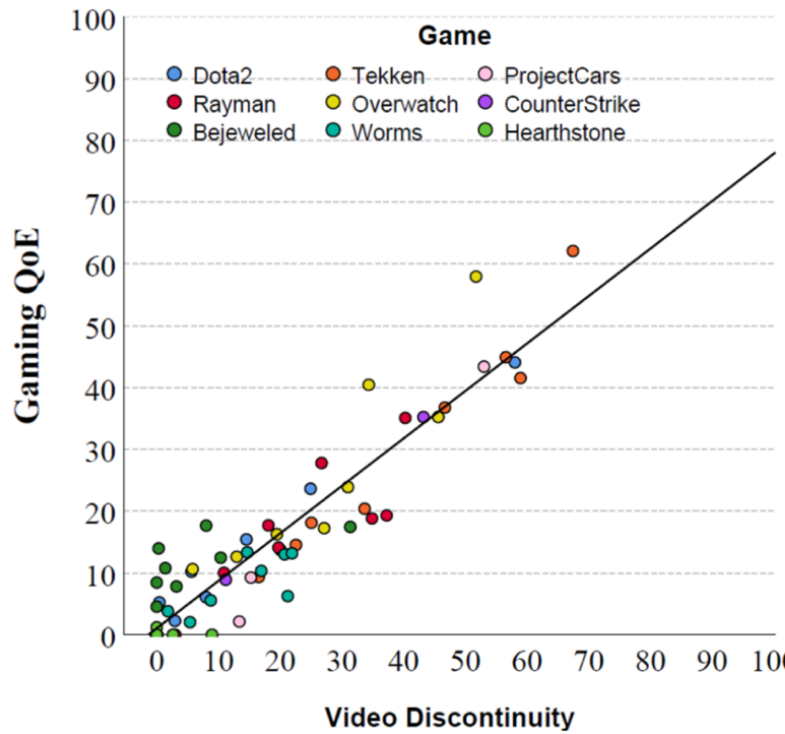


Impact of network delay



Suitability of quality aspects for prediction





Impairment Factors

$$R_{QoE} = R_{max,QoE} - a \cdot I_{VQ_{trans}} - b \cdot I_{VQ_{cod}} - c \cdot I_{TVQ} - d \cdot I_{IPQ_{delay}} - e \cdot I_{IPQ_{frames}}$$

- Rmax, QoE = 100
- I_VQ_{trans} : video quality impairment factor for video transmission errors
- I_VQ_{cod} : video quality impairment factor for video compression artefacts
- I_{TVQ} ITVQ : Temporal Video Quality (discontinuity) caused by low encoding framerates (FR_enc) or frame losses due to packet loss
- I_IPQ_{delay} : Input Quality caused by network delay, and
- I_IPQ_{frames} : Input Quality caused by low encoding framerates (FR_enc) or frame losses due to packet loss.

Modeling of Impairment Factors

Has to be updated as also bitrate and delay are important

Average displayed frame per second

$$Avg_{FPS} = FR_{enc} \cdot exp\left((f1 + \frac{f2}{FR_{enc}}) \cdot PL)\right)$$

Avg_FPS values for different encoding framerates and packet loss rates

FR_enc [fps]	**60**	**30**	**20**
PL [%]	Average Displayed Frames per Second [fps]		
0	60	30	20
0.1	55	28	19
0.2	50	26	18
0.4	42	23	16
0.5	39	21	15
0.7	32	18	14
1	25	15	12
2	10	8	7

Video Model

- The video quality is estimated based on G.1071 prediction for Transmission Impairment and Coding Impairment

I_VQ_Cod

$$I_{VQ_{cod}} = a_{1V} \cdot \exp(a_{2V} \cdot BitPerPixel) + a_{3V} \cdot ContentComplexity + a_{4V}$$

$$ContentComplexity = a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33}$$

$$BitPerPixel = \frac{Bitrate \cdot 10^6}{NumPixelPerFrame \cdot Framerate}$$

I_VQ_Trans

$$I_{VQ_{trans}} = c_{1V} \cdot \log(c_{2V} \cdot LossMagnitudeE + 1)$$

$$LossMagnitudeE = q_1 \cdot \exp(q_2 \cdot LossMagnitudeNP) - q_1$$

Impairment on Temporal Video Quality

- For the impairment factor I_TVQ the parameters Average displayed frame per second (Avg_FPS) and encoding framerate (FR_enc) were used:

$$I_{TVQ} = d_1 + d_2 \cdot exp\left(\frac{1}{Avg_{FPS}}\right) + d_3 . exp\left(\frac{1}{FR_{enc}}\right) + d_4 . Avg_{FPS} + d_5 . FR_{enc}$$

- The performance of the impairment factor I_TVQ on the training dataset

	Low complex class	Medium complex class	High complex class
RMSE	4.00	8.084	3.528
PLCC	0.887	0.889	0.986

Impairment on Input Quality

- Delay Impairment on Input Quality:

$$I_{IPQ_{Delay}} = \frac{d_1}{1 + \exp(d_2 - d_3 \cdot Delay)} - d_4$$

- Frame Loss Impairment on Input Quality:

$$I_{IPQ_{Frames}} = \frac{e_1}{\exp(e_2 \cdot Avg_{FPS})} + e_3 \cdot FrameLossRate$$

$$FrameLossRate = \frac{FR_{enc} - Avg_{FPS}}{FR_{enc}} \cdot 100$$

G.OMG Core Model

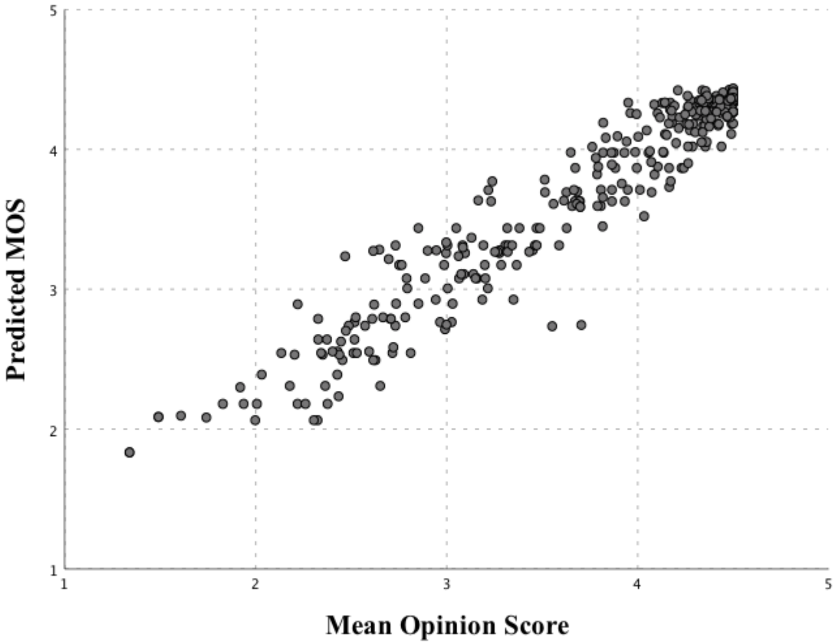
- It turned out that the temporal video quality impairment factor I_TVQ does not contribute to the model. Thereby, the following model equation remained:

$$R_{QoE} = R_{max,QoE} - a \cdot I_{VQ_{trans}} - b \cdot I_{VQ_{cod}} - c \cdot I_{IPQ_{delay}} - d \cdot I_{IPQ_{frames}}$$

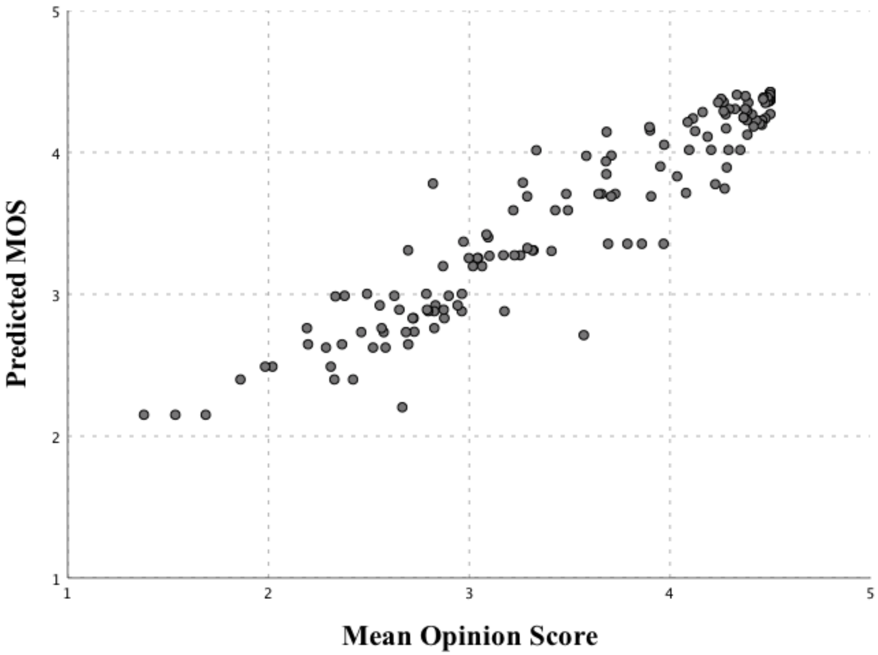
Coefficient	a	b	c	d
Value	0.923	0.922	1.036	0.966

Scatter Plots

Very good performance IF the content classifications are applied:
RMSE of 0.292 and PLCC of 0.94



Scatter plot of MOS ratings of training dataset and predicted gaming QoE ratings on MOS scale



Scatter plot of MOS ratings of test dataset and predicted gaming QoE ratings on MOS scale

Remaining challenges

- Input are system influence factors, potentially augmented by human and context influence factors
- Game characteristics largely determine the impact of system influence factors
- Parametric description of game characteristics missing
- Combination with Online Gaming missing
- Fast development of cloud gaming systems
  - Individual network protocols (packet loss influence)
  - Different compression methods (e.g. AV1)

Thank you for your Attention!

Updates on the ITU-T Activities with
Respect to Gaming Quality Assessment

Steven Schmidt

We are always searching for collaborations ☺

Visit
www.qu.tu-berlin.de
for more information.

References

[1] H.B. Duran, Newzoo quarterly update of its Global Games Market Report, 2018, URL: http://www.alistdaily.com/entertainment/mobile-games-market-newzoo-april-2018/ (last accessed: 15.10.2019)

[2] S. Möller, S. Schmidt, and J. Beyer, “Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming qoe,” in *Quality of Multimedia Experience (QoMEX), 2013 Fifth International Workshop on*, 2013.

[3] International Telecommunication Union, Study Group 12 (Source: TU Berlin), “ITU-T Contribution SG12-C387: First Draft for an Opinion Model Predicting Gaming QoE (G.OMG)“, Geneva, 2019.

[4] K. Poels, Y. A. W. de Kort, and W. A. Ijsselsteijn, “FUGA-The fun of gaming: Measuring the human experience of media enjoyment. Deliverable D3. 3: Game Experience Questionnaire,” FUGA project, 2008.

[5] F. Schiffner and S. Möller, “Direct Scaling & Quality Prediction for perceptual Video Quality Dimensions,” in 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), 2018.

[6] International Telecommunication Union, Study Group 12 (Source: TU Berlin), “ITU-T Contribution SG12-C376: Guidelines for the Assessment of Gaming QoE Using Crowdsourcing”, Geneva, 2019.

[7] Köster, Friedemann, et al. "Comparison between the discrete ACR scale and an extended continuous scale for the quality assessment of transmitted speech." *Fortschritte der Akustik-DAGA,* 2015.