

VQEG Spring 2023 Meeting Minutes

The sessions will be recorded. Let us know if your presentation cannot be recorded, and we will stop the recording.

Email presentations to mpinson@ntia.gov, to be posted to the VQEG website.

Monday, June 26	4
ITU-T Study Group 12 Question 19 Interim meeting	4
Session #2. P.obj-recog	5
ToR and Test Plan for P.obj-recog	5
Session #3. ITU-T Study Group 12 Q19 Interim meeting, P.910 rev	7
Session #4. SAM	8
“(Not so) new findings about Transmission Rating scale and subjective scores” by Pablo Pérez (#119)	8
“ZREC: Mean and Percentile Opinion Scores Recovery” by Jingwen Zhu (#131)	8
“On the Accuracy of Open Video Quality Metrics for Local Decision in AV1 Video Codec” by Andréas Pastor (#135)	8
“Recovering Quality Scores in Noisy Pairwise Subjective Experiments using Negative Log-Likelihood” by Andréas Pastor (#136)	8
“Towards guidelines for subjective Haptic quality assessment: a case study on quality assessment of compressed haptic signals” by Andréas Pastor (#137)	9
Tuesday, June 27	10
Session #1. ITU-T Study Group 12 Question 19 Interim meeting, P.obj-recog	10
Session #2. ITU-T Study Group 12 Question 19 Interim meeting, P.910rev	12
Session #3. ITU-T IRG-AVQA meeting	13
Session #3. ITU-T Study Group 12 Question 19 Interim meeting, P.obj-recog and P.910rev (continued)	14
AVHD	15
“Proposing More Ecologically-Valid Experiment Protocol Using YouTube Platform ” by Lucjan Janowski (#156)	15
Session #4. QAH	16
Wednesday, June 28	17
Session #1: JEG-Hybrid	17
“Updates on JEG-Hybrid Activities and inclusion of IGVQM” by Enrico Masala (#111)	17
"Human-in-the-Loop" Training Procedure of the AIO of a Real Subject", by Lohic Fotio (#109)	17
“Advances on the "template" on how to report DNN-based VQM”, by Lohic Fotio (#110)	17
Multi-Screen Dataset, by Nabajeet Barman (#107)	17
Session #1: AVHD	18
Datasheet for Subjective and Objective Quality Assessment Datasets, by Nabajeet Barman (#117)	18

"Acceptability and Annoyance of Video Quality in Context", by Ali Ak (#142)	18
Session #2: SAM	19
"JPEG AIC-3 Activity on fine-grained assessment of subjective quality of compressed images" by Dietmar Saupe (#121)	19
"Relaxed forced choice improves performance of visual quality assessment methods" by Mohsen Jenadeleh (#122)	19
"Quantization for MOS recovery in ACR experiments" by Mirko Dulfer (#124)	19
"Experiment Precision: Proposing Experiment Precision Measures and Methods for Experiments Comparison" by Lucjan Janowski (#114)	19
Session #3: CGI	20
"Updates on ITU-T P.BBQCG work item, dataset and model development" by David Lindero Mathias Wien (#126)	20
"Study of Subjective and Objective Quality Assessment of Mobile Cloud Gaming Videos " by Avinab Saha (#127)	20
"GAMIVAL: Video Quality Prediction on Mobile Cloud Gaming Content" by Yu-Chih Chen (#118)	20
Session #3: AVHD	21
"Crowdsourcing Pixel Quality Study Using Non-neutral Photos" by Mikołaj Leszczuk (#148)	21
Session #4: AVHD	22
"How more accurate VQA metric would lead to more savings? " by Hojat Yeganeh (#139)	22
"The Role of Theoretical Models in Ecologically Valid Studies: The Example of a Video Quality of Experience Model" by Kamil Koniuch (#113)	22
"Expert viewing and activities in MPEG Visual Quality Assessment" by Mathias Wien (#144)	22
Thursday, June 29	23
Session #1 and #2. NORM (No Reference Metrics)/ ETG (Emerging Technologies Group)	23
#103 Talks on "Deep Learning in Video Coding and Video Quality Domain: Future and Challenges"	23
#103 Panel Session on "Deep Learning in Video Coding and Video Quality Domain: Future and Challenges"	23
Session #3. NORM	24
"Summary of NORM successes and discussion of current efforts for improved complexity metric" by Margaret Pinson and Ioannis (#120)	24
"Blind visual quality assessment for mobile/edge computing" by C.-C. Jay Kuo (#112)	24
"VCAv2.0" by Vignesh V Menon (#134)	24
"Recent Updates on UVQ" by Yilin Wang (#143)	24
"Human Interaction in Industrial Tele-Operated Driving: Laboratory Investigation" by Shirin Rafiei (#141)	24
"Improving Netflix video quality with neural networks" by Christos Bampis (#132)	24
"Evaluate videos with the Advanced Video Quality Tool (AVQT)" by Pranav Sodhani (#154)	24
"Test plan on XR communication systems" by Pablo Pérez and Jesús Gutiérrez (#125)	24
Friday, June 30	26

Session #1: ETG	26
“Sandwiched Video Compression: Efficiently Extending the Reach of Standard Codecs with Neural Wrappers” by Berivan Isik (#140)	26
“Green Variable framerate encoding for Adaptive Live Streaming” by Vignesh V Menon/ Samira Afzal (#133)	26
“Energy efficient Live Per-title Encoding for Adaptive Streaming” by Prajit T Rajendran/ Vignesh V Menon (#138)	26
Session #1 and #2: CGI, “Special Gaming Session” led by Saman and Nabajeet	27
Session #3. QACOVIA	29
“Subjective Quality Assessment of Video Summarisation Algorithms: A Crowdsourcing Approach” by Avrajyoti Dutta (#147)	29
“Low complexity No-Reference Image Quality Assessment Based on Multiscale Attention Mechanism With Natural Scene Statistics” by Farhad Pakdaman, Li Yu (#145)	29
“Visual Quality Indicators Adapted to Resolution Changes” by Mikołaj Leszczuk (#149)	29
““In the Wild” Video Content as a Special Case of User Generated Content and a System for Its Recognition” by Mikołaj Leszczuk (#149)	29
Session #3: AVHD	30
“Subjective video quality for 4K HDR-WCG content using a browser-based approach for at-home testing” by Lukas Krasula (#151)	30
“Elevating Your Streaming Experience with Just Noticeable Difference (JND)-based Encoding” by Jingwen ZHU (#151)	30

Monday, June 26

Attendees introduced themselves, then VQEG projects were summarized and Introduced.

ITU-T Study Group 12 Question 19 Interim meeting

Begins 10:35 am

See slides in file "Q19 Interim meeting ver1"

Introduction to P.910rev and P.obj-recog

P.910rev = combine and revise ITU-T Rec. P.910, P.911, and P.913

P.obj-recog = parametric model for automated driving

- NTT and Ericsson interested in P.obj-recog. Seeking other interested organizations.

Session #2. P.obj-recog

Contribution from NTT, available on both the ITU website and VQEG meeting files

ToR and Test Plan for P.obj-recog

Masanori Koike and colleagues from NTT

Autonomous driving system with remote monitor, provides

Missing piece: need an object recognition ratio estimation model (estimate recognition ratio).

Depends on the distance needed to stop the car. Long distance = easier to stop but difficult to recognize (small object). Consider safety (e.g., person suddenly appears from behind object, close to the car). If loose portion of video to packet loss, may be difficult or impossible to recognize object.

Must always monitor video, try to recognize objects (if present or not). Conclude need parametric model (bitrate, packet loss, car velocity, target distance) plus a priori information (e.g., object type, weather).

Proposal: design for riskiest case, always need minimum quality, because people are doing the object recognition (i.e., at the monitoring center). Goal is observer has more than one car, but we start with one-to-one mapping (one observer per autonomous card).

Could be public network or a private network.

What happens when there is a critical issue? Hit breaks or stop car.

a priori info may change as drive (e.g., dirt on lens), dynamic instead of static input. Need subjective tests to understand parameter values.

Delay? Could be several seconds, must measure glass to glass: camera, encoder, network, decoder, display

Resolution and **frame rate** are other missing factors. Usually fix at 720p

Another factor: number of pixels on the object, angle of camera to object.

In 2nd paragraph of scope, questions on regulations and laws.

A prior information may be very difficult to quantify (e.g., object color, object size, object background).

P.obj-recog was presented and discussed. Recommended changes are noted in the updated document.

- **Glass to Glass Latency.** Add upper limit and note relationship between latency and distance driven (consumed by latency lag). H.265 could be 5 to 6 sec delay, so 100 meters at 40 kph. Include all steps (camera, encoder, network, decoder, and display)
- **Pixels on Object.** Add as another parameter, influenced by the number of pixels, field of view, and distance.
- **Johnson criteria.** Starting point to understand classification type. Note object exists vs recognize object.
- **Scope,** add limitation: camera problems omitted (e.g., dirt on lens, lens flare)
- **Add parametric model factor:** delay from camera auto white balance
- **Table.** Add lower limit on video resolution and frame rate
- **Table.** expansion for future: other codes (H.265 has longer delay but used by current market)
- **Table.** H.265 has long delays, note need for minimum encoding and decoding delays. Add which GOP structures are supported. Short GOP structure will help.
- **Table.** 10% packet loss seems much too high. Maybe reduce to 1%. Specify that this is at the raw level, and not the final packet loss level. Maybe more realistic to have 0.5% packet loss in the decoder. Note or mandate packet loss strategies. Transport architecture may have its own parameters that must be tuned.
- **A priori information.** Add information from camera, such as delay for auto white balance.
- **Figure. Add perception factors.** Viewing angle, impact of high driving speeds on the monitor, contrast sensitivity, etc.
- **Scope.** Add video quality for remote driving and teleoperation.

Session #3. ITU-T Study Group 12 Q19 Interim meeting, P.910 rev

Edits in revised document on VQEG meeting site. This document may need to be removed.

Session #4. SAM

“(Not so) new findings about Transmission Rating scale and subjective scores” by Pablo Pérez (#119)

e-model, designed for narrowband voice. extend to video, as a means to map multiple experiments to a single scale. Underlying structure seems to be similar for all experiments (e.g., speech and video).

“ZREC: Mean and Percentile Opinion Scores Recovery” by Jingwen Zhu (#131)

proposed new subject screening methods. Similar goals to MLE (content ambiguity, subject bias, subject inconsistency all considered).

paper link: <https://hal.science/hal-04017583/document>

github link: <https://github.com/kyillene/ZREC>

“On the Accuracy of Open Video Quality Metrics for Local Decision in AV1 Video Codec” by Andréas Pastor (#135)

Divide videos into small blocks (tubes of 64x64) to analyze visibility of impairments on a small scale, as used by codecs. Experimented with various subjective methods. Analyzed FR metrics, in comparison.

paper link: <https://ieeexplore.ieee.org/document/9897469>

abstract: VMAF is a popular objective quality metric used for video quality evaluation. The power of VMAF has been demonstrated for a wide variety of video scales and encoding processes. However, its ability to evaluate the quality of small video patches has not yet been tested, despite its importance for encoding algorithms. We applied Maximum Likelihood Difference Scaling (MLDS) methodology to estimate supra-threshold perceptual differences in localized sections in videos, also known as tubes, encoded using AV1. We further used the results to assess the performance of VMAF in this scenario and proposed a recalibration of the algorithm to improve its agreement with the subjective data.

“Recovering Quality Scores in Noisy Pairwise Subjective Experiments using Negative Log-Likelihood” by Andréas Pastor (#136)

Large scale dataset, crowdsourcing, reliability of subjects is a major concern. Spammer detection for indirect annotation?

paper link: <https://hal.science/hal-04132007/>

abstract: To gather larger datasets to train data-angry deep learning quality assessment models, crowdsourcing has become essential to recruit participants. These participants are asked their opinion by directly rating stimuli, e.g., using single or double stimulus methodologies, or indirectly by ranking stimuli or comparing distances as in the Maximum Likelihood Difference Scaling method. In crowdsourcing, participants' behaviors and environmental distractions are not controlled. So, the researcher must pay attention to the answers' reliability. Cleaning methods exist for direct annotation

subjective methodologies. However, solutions for indirect annotation methods are limited. In this work, we propose a method based on the negative log-likelihood to detect spammers among participants from their answers. To demonstrate its use, we applied it in a quadruplet preference-based scenario. The proposed method requires low computation and can be integrated into active-sampling strategies, where annotations available per comparison are small. We demonstrate that our method is robust to various spammer behaviors and accurate by removing only spammers. It helps reduce the gap between data collected in in-lab conditions (i.e., no spammer) and through crowdsourcing: our method reduces estimated uncertainties around data-points by 50%, and RMSE between estimations from an in-lab experiment and the same experiment performed in crowdsourcing by 1.8.

“Towards guidelines for subjective Haptic quality assessment: a case study on quality assessment of compressed haptic signals” by Andréas Pastor (#137)

Explored three types of haptic feedback. Analyzed rating task difficulty by comparing with other modalities datasets (e.g. audio, video, 360 video, spatial audio and audiovisual QoE tasks). Rating task does not seem to be more difficult than, for example, audiovisual rating task.

Then, outliers strategies developed for audio and video QoE experiments are tested, and conclusion are that basic methods proposed in standard as BT.500 and P.913 are not suited for outlier removal in haptic QoE.

In the full paper, more experiments are conducted to recommend on the number of participants to include in the subjective study. paper link: <https://hal.science/hal-04132004>

Abstract:

Modern systems are multimodal (e.g., video, audio, smell), and haptic feedback provides the user with additional entertainment and sensory immersion. Standard recommendation groups extensively studied and focused on video and audio subjective quality assessment, especially in signal transmission. In that context, subjective quality assessment and Quality of Experience (QoE) of Haptic signals is at its infant age. We propose further analyzing the collected data from a recent subjective quality assessment campaign as part of the MPEG haptic standardization group. In particular, we are addressing the following questions: 1) How the emerging field of haptic signal QoE can benefit from existing efforts of video and audio quality assessment standards? 2) How to detect possible outliers or characterize the rater's reliability? 3) How does the discriminability of haptic tests increases with the number of raters? Towards this goal, we question if traditional analysis as proposed for audio or video signal are suitable, as well as other state-of-the-art techniques. We also compare the discriminability of the haptics quality assessment tests with other modalities such as audio, video, and immersive content (360° contents). We propose recommendations on the number of raters required to meet the usual discriminability obtained for other perceptual modalities and how to process ratings to remove possible noise and biases. These results could feed future recommendations in standards such as BT500-14 or P.913 but for haptic signals.

Tuesday, June 27

Session #1. ITU-T Study Group 12 Question 19 Interim meeting, P.obj-recog

Discussion on test method. Question is “can you recognize objects?” Recommend change, ask what object the person recognizes. Otherwise might have something else in mind (other than what experimenter had in mind).

Change to “what object do you recognize? <list 5 objects>”. Unless this change is made, you cannot be sure that “the subject recognizes the object correctly” as stated in section 3.2. This list of potential objects could include “mistake / I did not intend to press stop”.

Alternatively, if you want a yes/no answer, could change the question to “Do you see or recognize a dangerous object in the direction of the car? yes/no”. But yes/no might get annoying for the subject. maybe give 5 sec to press escape if wrong?

Need clear definition of an “object” due to conflicting definitions in different fields of research.

Make sure text says that the video disappears when the subject presses “stop”.

Also, ideally, videos should have the same object on different roads, and under different conditions (e.g., shadow, night, day, rain). This will help make the test more robust.

Only time of video is recorded currently, not object recognized.

Change: Do not let people rewind and try again. Some subjects will try to watch again and again until they choose the ideal time to press breaks.

Change: Must decide exactly what experiment is being run. Text is ambiguous.

Option #1. Ability of the driver to react. Test is “when will the remote operator hit the breaks”. Can use video content to decide if this decision was right or wrong. The current yes/no question is appropriate. Delete from text any mention of correctly recognizing the object. Matches real situation. Remote drivers will not have a 2nd chance. This option was preferred by attendees, because the experiment is simple.

Option #2. Ability of the driver to identify objects. Want to know if recognize the object correctly (Must change the question (as above) to list possible objects). Must record more than just the timing of when pressed stop; must also record the object recognized and test the correctness of this answer. This may not be as valuable; prior research may already have this information. List of items should include option like “I recognize a possible danger, I am not sure what it is, but I want to slow down”. This is a much more complex experiment that will require careful design.

Change: Don't use the same video two or more times.

Consider break latency: Similar experiment, remote driving a toy truck. latency was very important. In addition to the glass to glass (camera to monitor) latency, the another latency must be added: from the reaction between operator and to the reaction of the car's break.

Consider in test structure that actual operators will work long hours. Might be 8 hour work days with breaks every 2 hour, for example. Should the test structure mimic this? Opposing argument is that workers will not have such a high rate of incidences, which will cause fatigue. Maybe just explain your thinking and why the structure was chosen.

Section 4.1, two other conditions to consider: an object on the road (e.g., heap of rocks) or a car (e.g., coming the wrong direction, car stopped in the road).

Cost. Need special videos that have metadata: distance from camera to object. This could be expensive.

Question: how much will it cost to create the test material? Participate as a subjective test lab? Could we use deep fake, to artificially create objects on the road? Could a drone be used, with forward facing camera, to record moving quickly toward a person or object without endangering the person or driver? Maybe an auto manufacturer in collaboration, should have an auto simulator for artificial test track. Use totally animated footage would be worst (not realistic response to compression) but easy.

Recording videos, consider privacy laws.

Concerns raised about 10% packet loss rate. See Monday minutes. If possible, replace this with the actual packet loss rate (e.g., rural area, urban area). More realistic test design.

Session #2. ITU-T Study Group 12 Question 19 Interim meeting, P.910rev

All updates was captured in the P.910rev document which available on the VQEG website

Session #3. ITU-T IRG-AVQA meeting

Report on ITU-R and ITU-T activities. Discussion on how to progress J.no-ref, a work item that has remained in progress but incomplete for 20 years.

Session #3. ITU-T Study Group 12 Question 19 Interim meeting, P.obj-recog and P.910rev (continued)

Article “Teleoperation steps in when autonomous vehicles does not know what to do” presented as an example of why P.obj-recog is needed. Further edits to P.910rev discussed and made.

Two SG12 meetings will be held in the 2021-2024 study period.

September ITU-T SG12 meeting :

- 19-28 September, 2023 in Mexico city
- Contribution deadline: September 6th

April ITU-T SG12 meeting:

- 15-25 April 2024, location is not decided

AVHD

“Proposing More Ecologically-Valid Experiment Protocol Using YouTube Platform ” by Lucjan Janowski (#156)

Session #4. QAH

[Draft_Recom_June2023.docx - Google Docs](#)

Wednesday, June 28

Session #1: JEG-Hybrid

“Updates on JEG-Hybrid Activities and inclusion of IGVQM” by Enrico Masala (#111)

Maragaret Pinson asks about necessary adaptations of descriptions on the VQEG website to reflect the merge of activities. The JEG-Hybrid chairs will make a proposal for changes to the website. Additionally, a mail should be sent to the main VQEG mailing list to inform the people of the IGVQM merge and plans that aren't here today.

Biweekly conference call (next: Jul 10, 2023, 4pm CEST) open for anyone to join.
JEG-Hybrid website <https://vqegjeg.github.io/jeg-hybrid/>
Subscribe to the VQEG-JEG mailing list: jeg@vqeg.org

"Human-in-the-Loop" Training Procedure of the AIO of a Real Subject”, by Lohic Fotio (#109)

Questions are asked about the accuracy reported in the work. It is clarified that the accuracy measure was a simplified representation of reality by looking for the discrete output.

“Advances on the "template" on how to report DNN-based VQM”, by Lohic Fotio (#110)

A comment is made to ask the people that made the datasets to include the information themselves.

Questions are asked about possible metrics for benchmarks. A discussion on this topic takes place.

Questions are asked whether concepts from explainable AI could be incorporated.

It is commented that there is value in the idea of a guideline.

A question is asked whether the link can be made with P.1401 Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models.

Multi-Screen Dataset, by Nabajeet Barman (#107)

Appreciation about the work is expressed. Clarification is asked about which P1204 model was used. It is clarified that the right model is used for each setting. Clarification is asked about the order in which the devices is used. Half of participants watched the TV, then tablet and then phone for each sequence; for the other half started from phone, then tablet, then TV.

<https://github.com/NabajeetBarman/Multiscreen-Dataset>

Session #1: AVHD

Datasheet for Subjective and Objective Quality Assessment Datasets, by Nabajeet Barman (#117)

<https://github.com/NabajeetBarman/datasheet-for-qoe-datasets>

It is commented that Qualinet tried a similar approach. It would be good to contact the involved people for best practices.

It is mentioned that there is a metadata format for such datasets and it is questioned whether this could be combined. It is also suggested that maybe it would be useful to make it machine readable. It is suggested to add the github link to the tools website.

“Acceptability and Annoyance of Video Quality in Context”, by Ali Ak (#142)

Question is asked about what the effect could be from performing the test in a country where people are confronted with restricted data plans in their daily life.

A comment is made that the research works towards Big Noticeable Differences instead of Just noticeable Differences. ACR does not immediately translate to satisfaction of users. Different contexts such as country, data plan matter and it is questioned what other aspects have an influence. Also, free plan as compared to a paid plan can be considered as a possible context.

It is commented that such results indicate that results from tests depend on the experiment setup.

It is questioned what the consistency was among different participants. Sufficiency can be linked to the proposed work. Content and personal differences can also be linked to annoyance. Could user preference be included as an additional item in the rating scale (annoying, annoying but acceptable, not acceptable)?

Session #2: SAM

“JPEG AIC-3 Activity on fine-grained assessment of subjective quality of compressed images” by Dietmar Saupe (#121)

“Relaxed forced choice improves performance of visual quality assessment methods” by Mohsen Jenadeleh (#122)

“Quantization for MOS recovery in ACR experiments” by Mirko Dulfer (#124)

“Experiment Precision: Proposing Experiment Precision Measures and Methods for Experiments Comparison” by Lucjan Janowski (#114)

code: <https://github.com/Qub3k/qoe-experiment-precision>

Session #3: CGI

“Updates on ITU-T P.BBQCG work item, dataset and model development” by David Lindero Mathias Wien (#126)

Updates on the ITU-T SG12 P.BBQCG work item.

“Study of Subjective and Objective Quality Assessment of Mobile Cloud Gaming Videos ” by Avinab Saha (#127)

Presenting the gaming video quality dataset created for mobile cloud gaming using Meta cloud gaming service.

600 videos are used from 30 source sequences.

Source sequences are displayed by streaming at 50 mbps.

Max resolution was 720p.

The dataset will be available soon at

<https://live.ece.utexas.edu/research/LIVE-Meta-Mobile-Cloud-Gaming/index.html>

“GAMIVAL: Video Quality Prediction on Mobile Cloud Gaming Content” by Yu-Chih Chen (#118)

Presented the development of the gaming video quality model, based on fusing different features from different models.

The model performs better on the gaming dataset, and also faster than other gaming-specific models compared.

Access to the model: <https://github.com/utlive/GAMIVAL>

Two presenters did not show up.

Session #3: AVHD

“Crowdsourcing Pixel Quality Study Using Non-neutral Photos” by
Mikołaj Leszczuk (#148)

Session #4: AVHD

Two liaison statements were discussed. See meeting files for details.

“How more accurate VQA metric would lead to more savings? ” by Hojat Yeganeh (#139)

Method to compare metrics to JND, to make decisions of reducing bitrate without impacting quality. Attendees raised concerns raised with analysis technique, which are based on minimum function (too unreliable; should use mean, median, or 95th percentile instead) and assume high precision of the metrics (VMAF, SSIMPLUS).

“The Role of Theoretical Models in Ecologically Valid Studies: The Example of a Video Quality of Experience Model” by Kamil Koniuch (#113)

“Expert viewing and activities in MPEG Visual Quality Assessment” by Mathias Wien (#144)

Thursday, June 29

Session #1 and #2. NORM (No Reference Metrics)/ ETG (Emerging Technologies Group)

#103 Talks on “Deep Learning in Video Coding and Video Quality Domain: Future and Challenges”

- Towards Real-Time 4K Image Super-Resolution (Marcos V. Conde, University of Würzburg, Sony Interactive Entertainment)
- Generative Compression & Challenges for Quality Assessment (David Minnen, Google)

We used our open-sourced crowdsourcing implementation of P.910 ([microsoft/P.910: Crowdsourced implementation of P.910 \(github.com\)](https://github.com/microsoft/P.910)) for CLIC2022 challenge (we did both ACR and DCR test). See here for details: [Shared Copy of CLIC 2022 - CVPR Presentation \(storage.googleapis.com\)](https://storage.googleapis.com/CLIC2022/Presentation/Shared%20Copy%20of%20CLIC%202022%20-%20CVPR%20Presentation.pdf)

Crowdsourcing platform available at: <http://github.com/microsoft/P.910>

Paper: [A crowdsourced implementation of ITU-T P. 910](#)

#103 Panel Session on “Deep Learning in Video Coding and Video Quality Domain: Future and Challenges”

- Ioannis Katsavounidis (Meta)
- Zhi Li (Netflix)
- Richard Zhang (Adobe)
- Mathias Wien (RWTH Aachen)

Session #3. NORM

“Summary of NORM successes and discussion of current efforts for improved complexity metric” by Margaret Pinson and Ioannis (#120)

“Blind visual quality assessment for mobile/edge computing” by C.-C. Jay Kuo (#112)

“VCAv2.0” by Vignesh V Menon (#134)

“Recent Updates on UVQ” by Yilin Wang (#143)

Open source metric software: <https://github.com/google/uvq>

“Human Interaction in Industrial Tele-Operated Driving: Laboratory Investigation” by Shirin Rafiei (#141)

Questions asked about the possible impact of increasing the driver’s field of view. Drivers want to have side views and other views strategically arranged around them (U-shape bank of 3x3 monitors, including front view, side view, back view, etc.).

“Improving Netflix video quality with neural networks” by Christos Bampis (#132)

“Evaluate videos with the Advanced Video Quality Tool (AVQT)” by Pranav Sodhani (#154)

“Test plan on XR communication systems” by Pablo Pérez and Jesús Gutiérrez (#125)

Exploring four interactive task: (1) survival task, compromise verbal discussion, (2) charades or physiotherapy, visual-only or visual dominant communication, (3) block building, as per lego

construction with learner and instructor, (4) treasure hunt, wide space, clues hidden in real or virtual world, remote support may have map or more clues.

12 proposals for experiments, to try these proposed techniques for interactive subjective tests.

Will feed into draft ITU Recommendation. Expect results at the next VQEG meeting.

Friday, June 30

Session #1: ETG

“Sandwiched Video Compression: Efficiently Extending the Reach of Standard Codecs with Neural Wrappers” by Berivan Isik (#140)

Neural image compression: flexible and performance improvement, but slow, high memory, slow training. Alternative: standard encoder with neural pre- and post-processors. .

“Green Variable framerate encoding for Adaptive Live Streaming” by Vignesh V Menon/ Samira Afzal (#133)

Samira started the presentation and continued by Vignesh.

“Energy efficient Live Per-title Encoding for Adaptive Streaming” by Prajit T Rajendran/ Vignesh V Menon (#138)

Session #1 and #2: CGI, “Special Gaming Session” led by Saman and Nabajeet

Luleå University of Technology, PlayStation, Kingston University, TU Berlin, UT Austin, Aalto University, Simon Fraser University, Sydney University, WPI, University of Alberta, InterDigital, Ericsson, RISE, University of Alberta, University of Oslo (UiO), Simula Metropolitan Center for Digital Engineering (SimulaMet)

Short presentation by Saman Zadtootaghaj about the purpose of Special Gaming Session.

Intro presentation by Maria Martini (Kingston University):

- Subjective quality datasets for gaming content
- Objective quality assessment models for gaming streaming services
- Content Characterization and Classification
- Light field video QA and streaming
- Involvement at standardization in ITU-T, MPEG-DASH, IEEE.
- Center of augment and virtual reality + TV studios
- QoMEX, ICASSP, ICME, ACM MMSys, VQEG.
- VR gaming
- Dataset be released by next year on VR QoE
- Involved in ITU-T SG12, IEEE Standards Association, ETSI (mobile communication), VQEG.

Intro presentation by David Lindero (Ericsson):

- VR, compression and movement amount, delay and viewport and controller
- Cloud gaming - P.BBQCG
- Reproducibility tool, P.NATS processing chain

Intro presentation by Karan Mitra, Irina Cotanis (Luleå University of Technology/Infovista):

- Mobile cloud gaming
-
- XR telecommunication on extended - P.1320, P.TeleMeTax on taxonomy of telemeetings
- Involved in ITU-T SG12, 3GPP, VQEG.

Intro presentation by Tanja Kojic (TU Berlin)

- User experience and eXtended reality, VR and AR
- How do human users experience virtual environments produced by new media technologies?
- Human-machine interaction is the main focus
- Exergaming and multiplayer, visual biofeedback
- Active in ITU-T SG12

Intro presentation by Cise Midoglu, Carsten Griwodz (University of Oslo (UiO), Simula Metropolitan Center for Digital Engineering (SimulaMet)

- QoE, VR, Gaming
- Carsten focuses on haptic and VR.
- XR, communication and VR
- SimulaMet involves in Gaming QoE on ITU-T P.BBQCG, in both passive and interactive.
- VR and Gaming in social context, like protection of minors
- Crowdsourced user studies

Intro presentation by Edouard François (InterDigital):

- Gaming media compression
- 2D video coding, implicit visual representation for 2D and 3D video data

- Point cloud and dynamic mesh
- XR gaming, in defining characteristics of 5G, QoE framework, rendering media service enabler
- Active in MPEG, JVET and 3GPP SA4

Intro presentation by Matthew Guzdial (University of Alberta):

- Learning Game Mechanics from video/stream
- The lab focus on procedural content generation + core machine learning
- Reinforcement learning for game development and adapting content to users
- Smart adaptation of AI partners to users

Intro presentation by Mohamed Hefeeda (Simon Fraser University):

- Project on cloud gaming
- Cloud gaming challenges:
 - latency, modularity and heterogeneity + sensitivities
- AMD collaboration, reduce the bandwidth for cloud gaming
- RoI encoding: DeepGame (36% bandwidth reduction)
 - spatio-temporal content gaming session
- Up sampling, super resolution: GameSR → 30 ms
- RL controller

Intro presentation by Mark Claypool (Worcester Polytechnic Institute):

- Research on online and cloud games
- Cloud-based game streaming - Game bottleneck sharing compared Luna, GeForce, Stadia
- Gaming QoE and effect of latency
- Latency compensation techniques

Intro presentation by Avinab Saha (UT Austin):

- Two Gaming VQA dataset (UGC gaming VQA and PGC Gaming VQA) - each 600 videos
 - LIVE-Youtube gaming video dataset- WACV 2022
 - Live-Meta mobile cloud gaming - IEEE TIP 2023
- Gaming video quality assessment models:
 - Game-VQP - IEEE SPL 2023
 - GAMIVAL - IEEE TIP (under review)
- Work on data efficient label free- pre-training algorithms
- Efficient deep architecture design,
 - Video transformers and ConvNext

Intro presentation by Kjell Brunnström (RISE):

- RISE has a special mission to support small and medium sized enterprises.
- QoE for remote control
- Measurements of perceived and sound and eye movement
- Gaming connection: nintendo kart live
- Human Interaction in industrial tele-operated driving
- Perceptual aspect of real-time video streaming using drones
- Active at ITU-T SG12 through VQEG.

Session #3. QACOVIA

“Subjective Quality Assessment of Video Summarisation Algorithms: A Crowdsourcing Approach” by Avrajyoti Dutta (#147)

“Low complexity No-Reference Image Quality Assessment Based on Multiscale Attention Mechanism With Natural Scene Statistics” by Farhad Pakdaman, Li Yu (#145)

“Visual Quality Indicators Adapted to Resolution Changes” by Mikołaj Leszczuk (#149)

““In the Wild” Video Content as a Special Case of User Generated Content and a System for Its Recognition” by Mikołaj Leszczuk (#149)

Session #3: AVHD

“Subjective video quality for 4K HDR-WCG content using a browser-based approach for at-home testing” by Lukas Krasula (#151)

“Elevating Your Streaming Experience with Just Noticeable Difference (JND)-based Encoding” by Jingwen ZHU (#151)