A brief introduction to the Video Quality Experts Group

January 2023
VQEG—A Brief History

- Formed in 1997 to advance the field of video quality assessment
- Closely related to ITU-T and ITU-R study groups
  - ITU-T SG9 (Broadband cable & TV)
  - ITU-T SG12 (Performance, QoS and QoE)
  - ITU-R WP6C (Programme production and quality assessment)
- Historically, a primary focus on:
  - Creation of test plans to develop and validate objective quality metrics
  - Particular focus on defining the scope and subjective test methodology
  - Statistical techniques for assessing model performance
  - Recommending approaches/models to be standardized by ITU-R/ITU-T
How is it organized?

- **VQEG board**
  - Kjell Brunnström (RISE Research Institute of Sweden AB)
  - Margaret Pinson (NTIA/ITS, USA)

- **Working groups**
  - Individual co-chairs per group

- **Bi-annual meetings**
  - Historically in-person, worldwide
  - Now mostly online (due to COVID health and traveling restrictions)

- **Next meeting: May/June 2023 (tba)**
# Current Chairs & Co-Chairs

<table>
<thead>
<tr>
<th>Category</th>
<th>Chairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>General chairs</td>
<td>Kjell Brunnström, Margaret Pinson</td>
</tr>
<tr>
<td>5G Key Performance Indicators (5GKPI)</td>
<td>Pablo Perez, Kjell Brunnström</td>
</tr>
<tr>
<td>Audiovisual HD (AVHD)</td>
<td>Shahid Satti, Silvio Borer, Ioannis Katsavounidis</td>
</tr>
<tr>
<td>Computer Graphics Imagery (CGI)</td>
<td>Saman Zadtootaghaj, Nabajeet Barman</td>
</tr>
<tr>
<td>Emerging Technologies Group (ETG)</td>
<td>Nabajeet Barman, Saman Zadtootaghaj</td>
</tr>
<tr>
<td>Human Factors for Visual Experiences (HFVE)</td>
<td>Maria Martini, (vacant)</td>
</tr>
<tr>
<td>Immersive Media Group</td>
<td>Jesus Gutierrez, Zhenzhong Chen, Pablo Perez</td>
</tr>
<tr>
<td>Implementers Guide for Video Quality Metrics (IGVQM)</td>
<td>Ioannis Katsavounidis, (vacant)</td>
</tr>
<tr>
<td>JEG Hybrid</td>
<td>Marcus Barkowsky, Glenn Van Wallendael, Enrico Masala</td>
</tr>
<tr>
<td>No Reference Metrics (NORM)</td>
<td>Ioannis Katsavounidis, Margaret Pinson, Werner Robitza, Cosmin Stejerean</td>
</tr>
<tr>
<td>Quality Assessment for Computer Vision Applications (QACoViA)</td>
<td>Mikolaj Leszczuk, Patrick Le Callet, Lu Zhang</td>
</tr>
<tr>
<td>Quality Assessment for Health applications (QAH)</td>
<td>Lu Zhang, Lucie Lévêque, Meriem Outtas</td>
</tr>
<tr>
<td>Statistical Analysis Methods</td>
<td>Lucjan Janowski, Ioannis Katsavounidis, Zhi Li, Patrick Le Callet</td>
</tr>
<tr>
<td>Tools and Subjective Labs Setup Co-Chair</td>
<td>Glenn Van Wallendael, Werner Robitza</td>
</tr>
<tr>
<td>Video Archives Support</td>
<td>Femi Adeyemi-Ejeye</td>
</tr>
</tbody>
</table>
What’s nice about VQEG?

- Free to join — no membership fees
- No strict or complicated rules
  - Consensus is often reached without lengthy voting procedures
- Simple organization and hierarchy
  - Chairs & co-chairs for different projects
  - Anyone can propose or contribute a new project
- Highly interactive meetings
  - Anyone can present their ideas
  - Focus on discussion time
- Not a commercial venue
  - No sales talks, no commercial advertising
- Mixture between academia and industry
General Topics and Resources
Main Topics

- **General research on video QoE in various fields**
- **Video quality model development**
  - Various types of models (hybrid/bitstream, no-reference, ..)
  - VQEG members may define the scope and test plan
- **Input to standardization forums**
  - ...based on developed and validated models
- **Subjective tests & collection of subjective databases**
  - To develop and validate subjective test methodologies ("ILG" approach)
  - To predict the performance of objective video quality models
- **Joint production of software tools**
  - Helper tools for conducting subjective tests
  - Objective quality analysis
- **Exploration of new application areas**
  - Multimedia, 3DTV, gaming, VR/XR, ..
### Video Quality Model Development — Typical Approach

<table>
<thead>
<tr>
<th>Test Plan</th>
<th>Training Databases</th>
<th>Model Training</th>
<th>Validation Databases</th>
<th>Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creation of scope and terms of reference</td>
<td>Joint development of subjective test databases</td>
<td>Development of the models by proponents based on training data</td>
<td>Development of new databases with previously unknown content</td>
<td>Validation of model performance according to predetermined statistical criteria</td>
</tr>
<tr>
<td>Definition of inputs/outputs</td>
<td>Conduction of tests by proponents or independent labs (“Independent Laboratory Group”, ILG)</td>
<td>Collaborative or in form of a competition — various advantages of either method</td>
<td>Conduction of tests by different labs</td>
<td>Suggestion of which models may be standardized — ITU-T contributions</td>
</tr>
<tr>
<td>Rules for model submission and validation</td>
<td>Sharing of training data</td>
<td>Submission of model candidates</td>
<td>Collection of validation data; merging with training data to form complete dataset</td>
<td></td>
</tr>
<tr>
<td>Open call for participation to industry/academia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

→ historical approach — model development may now be iterative/collaborative or done within ITU-T itself
Model Development — Previous Projects

- **FRTV Phase 1:**

- **Multimedia Phase 1:**

- **HDTV Phase 1:**
  - Five video datasets available

- **AVHD-AS:**
  - Joint project with ITU-T Study Group 12 P.NATS Phase 2
  - UHD/4K, 60p, H.264, H.265, VP9
  - IEEE Access Paper summarizing the effort
Input to Standardization Forums

- Historically, VQEG has recommended model algorithms to become standardized
- Based on performance against subjective data
- Sometimes no models could be standardized due to low performance or unreliability (e.g., mostly the case with no-reference pixel-based models)
- Newer VQEG projects have a more collaborative and iterative approach to developing algorithms
- Contributions for subjective evaluation techniques
Video Datasets & CDVL

- Consumer Digital Video Library
- High-quality, royalty-free test material, mostly from previous VQEG projects
- Another list of datasets on VQEG website
Software Tools

- Public **software tools repository**
- Various software packages developed over the years
- Grouped by topic and searchable
  - Quality Analysis
  - Encoding
  - Streaming
  - Subjective Test Software
  - Helper Tools

Please submit your tools! →
https://github.com/VQEG/software-tools/
Dissemination

- Extensive reports of previous subjective test and model development activities
  - Can be found on VQEG website
  - ...even for historical activities
- VQEG has contributed columns to **SIGMM Records** (ACM SIG Multimedia's quarterly newsletter) with recent updates
  - Latest issue:  
  - Detailed Summary of the December meeting
Current Activities

Details:

https://vqeg.org/projects-home/
https://vqeg.org/meetings-home/
Statistical Analysis Methods (SAM)

- **Overview:** Statistical Analysis Methods
- **Minutes:** VQEG SAM Monthly
- **Developed a new ratings recovery method based on SUREAL**
  - How can you better recover “real” ratings from noisy subjective data?
- **Further additions to ITU-R BT.500-14 and ITU-T Rec. P.910/913**
- **Revisions and merging of ITU-T Rec. P.910 and P.913**

Li et al., 2020
Statistical Analysis Methods (SAM) ctd.

- Various historical activities related to designing and evaluating test methods
- How to obtain the most valid and reliable ratings?
- Examples:
  - Comparison of different rating scales (2011 paper)
  - Impact of the test environment on MOS (2012 paper)
  - Experiments with unrepeated scenes (2019 paper)

[Diagram showing 5-pt discrete vs. continuous scale]

Janowski et al. 2019

Huynh-Thu et al., 2011
No-Reference Metrics (NORM)

- Design of a no-reference pixel-based metric
- **Various resources** for NR metrics
- Collection of video datasets with new scope (e.g. security applications, user-generated content)
- **Open framework** for collaborative development of no-reference quality indicators
- **Journal paper** (2022) showing why NR metrics lack accuracy and reproducibility

Various NR metrics vs. dataset, Pinson, 2022
No-Reference Metrics (NORM) ctd.

- How can we determine the complexity of a video before encoding it?
  - Clarifying the use of Spatial and Temporal Information (SI, TI)
    - ITU-T P.910 was updated
  - Siti-tools code released as open-source
  - Taking into account motion using a motion search framework
  - Integrating approaches like Video Complexity Analyzer

- Video quality metadata standard
  - How to embed metadata on source/encoded video quality directly in containers or bitstreams
  - Payload definition and liaison with MPEG and AOM (Alliance for Open Media)
Computer Generated Imagery (CGI)

- Analyzing and evaluating of computer-generated content
  - Creation of open source datasets
  - Machine learning/deep learning based quality enhancement of gaming content
  - Development of models and metrics for assessing gaming QoE
- Aligned with ITU-T Study Group 12 work item “P.BBQCG”
  - Developing a gaming QoE model that uses the bitstream metadata for video quality
  - Interactive and passive subjective tests planned
- Previous activities:
  - Creation of various ITU-T recommendations
  - ITU-T Rec. G.1032: Influence factors on gaming QoE
  - ITU-T Rec. P.809: Subjective evaluation methods for gaming quality
  - ITU-T Rec. G.1072: Opinion model for gaming applications
Joint Effort Group – Hybrid (JEG-Hybrid)

- Project homepage
- Original idea to develop a hybrid quality model, but activities evolved to become more diverse, more research-oriented
- Research questions:
  - Modeling single observer behavior in subjective experiments with neural networks → predicting individual quality ratings
  - Modeling disagreement in video quality metrics
  - Templates for publishing results in image/video quality assessment → reproducible research

Predicting individual quality ratings, Fotio T. et al., 2023
Implementer's Guide for Video Quality Metrics (IGVQM)

- Create a guide on how to properly use video quality metrics
- **Working document** available
  - Currently moving part of the activities into JEG-Hybrid
- **Scope**:
  - Collect full-reference metrics and open-source solutions
  - Highlight differences between metrics
  - Determine temporal aggregation methods for image-based metrics (PSNR, SSIM)
  - Mappings between nonlinear objective metrics and linear scales (e.g. 0-100)
Quality Assessment for Computer Vision Applications (QACoViA)

- Methods for determining precision of computer vision approaches
- Identifying the limits of computer vision methods with respect to the visual quality of the ingest
- Recent highlights:
  - Method for Assessing Objective Video Quality for Automatic License Plate Recognition Tasks
  - Assessing Rail 8KUHD CCTV Forward Facing Video
  - Comparing the Robustness of Humans and Deep Neural Networks on Facial Expression Recognition
  - Video Coding for Machines: Large-Scale Evaluation of Deep Neural Networks Robustness to Compression Artifacts for Semantic Segmentation
5G Key Performance Indicators (5GKPI)

- **Goals:**
  - Defining use cases for video in 5G
  - Studying QoE aspects for video in mobility and industrial scenarios
  - Identifying the relevant network KPIs and application-level video KPIs (e.g. picture quality, A/V sync, ..)
  - Building open datasets for algorithm testing and training

- **Recent highlights:**
  - ITU-T Technical Report GSTR.5GQoE (2022): Specific QoE requirements and required performance and features from the network
**Immersive Media Group (IMG)**

- **Quality assessment of immersive media**
  - 360-degree content, virtual/augmented/extended reality, light field/plenoptic content, 3D content (stereo, multiview, FVV, etc.).

- **Goals: Baseline quality assessment of today’s systems**
  - Datasets of immersive media content
  - QoE guidelines, subjective test methods, objective metrics, etc.

- **Activities:**
  - Finalized test plan on quality assessment of 360-degree videos → ITU-T P.919
  - Ongoing test plan: Evaluation of immersive/interactive communication systems.
  - Light field quality assessment
Quality Assessment for Health Applications (QAH)

- Assemble databases for medical image and video quality, eye-tracking
- Define subjective experiment methodologies for diagnostic or surgery tasks, eye tracking, …
- Evaluate and develop quality metrics for medical imaging/video, visual attention prediction models, …
- Study quality requirements and QoE for telemedicine
Emerging Technologies Group (ETG)

- Newly formed group
- AI-based technologies
  - Super Resolution
  - Learning-based video compression
  - Enhancement, Denoising and other pre- and post-filter techniques
- Greening of streaming
  - Saving energy and its impact on visual quality
- Blockchain in Media and Entertainment
- Liaison with other standards activities
  - 3GPP, SVTA, CTA WAVE, UHDF, etc.
  - MPEG/JVET
Summary
Summary

- VQEG is a free-to-attend forum for video quality experts
- Both academia and industry is invited to contribute
- 25 years of history with great achievements:
  - Standardization of video quality models
  - Development and application of subjective test procedures
  - Collection of resources for video quality research (databases, software)
  - Discussion of new and emerging technology
Getting Involved

- Have a look at the projects page
- Subscribe to the mailing lists
- Join our regular working group online meetings
  - Any contribution is welcome
  - Bring your questions!
- Take part in the bi-annual VQEG general meetings