

VQEG: Shaping Standards on Video Quality

Quan Huynh-Thu¹, Arthur Webster², Kjell Brunnström^{3,4}, Margaret Pinson²

¹Canon Information Systems Research Australia (CiSRA), ²NTIA's Institute for Telecommunication Sciences, ³Acreeo Swedish ICT AB, ⁴Mid Sweden University

The Video Quality Experts Group (VQEG) is an international and independent organisation of technical experts in perceptual video quality assessment from industry, academia, and government organisations. The main goals of VQEG are to advance the field of perceptual video quality assessment, establish best practices for subjective experiments, conduct large-scale subjective experiments, and evaluate new objective video quality assessment models. VQEG pursues several objectives and one of them is to support standardisation work. Based on results produced by VQEG, more than twenty-five international Recommendations on video quality have been approved by the International Telecommunication Union (ITU). This paper provides an overview of VQEG activities and how VQEG plays a major role in research and the development of standards on video quality.

1. Introduction

The Video Quality Experts Group (VQEG) was established in 1997 as a forum of international experts working in the field of perceptual video quality. VQEG is an international and independent organisation that is open to all interested organisations and individuals and does not require any membership or fees. VQEG brings together experts in subjective video quality assessment and objective video quality measurement from industry, government, and academia. The uniqueness of VQEG is that it brings together experts in a variety of areas such as image processing, video processing, psychophysics, cognitive psychology, network communications, displays, mathematics, and statistics. These experts share the goal of understanding perception of video quality and developing useful and validated methods for video quality assessment, both subjective and objective.

Initially, VQEG focussed on a single project that only validated objective models, to support Standards Development Organisations (SDO). Over the years, VQEG grew to include multiple projects that explore different areas, including:

- New or improved subjective testing methods
- Collaboratively developed quality models
- Supplementary metrics beyond Mean Opinion Score
- Improved understanding of novel video technologies
- Objective model validation
- Subjectively rated databases of video content
- Video quality assessment software tools

VQEG meetings emphasise an open exchange of ideas, information, and research through discussions, presentations, and publications. The work and interest of

VQEG covers a broad range of research and development of video quality assessment, where standardisation is one very important aspect. This paper gives an overview of VQEG activities and looks at how VQEG is contributing to the standardisation process.

2. VQEG structure, workgroups and projects

VQEG encompasses people with diverse credentials who may focus on different goals. This encourages communication and collaboration amongst disparate organisations from various countries. The impact of VQEG's work can be seen in academic outputs (e.g. software tools, subjective datasets, research papers) and international standards on video quality assessment.

Standardisation provides a platform of established knowledge for the industry to build upon and a level playing field for competition. It is driven by industrial need, but input from academic research is essential. Academia usually has difficulty contributing to the standardisation process due to the fees and formal structure of traditional SDOs. VQEG can bridge this gap in the area of video quality assessment because it is associated with the ITU and has active members in both academia and industry. This allows scientific knowledge, ideas, and results to be informed by industry needs when it is contributed to the international standards process.

VQEG conducts work via discussions over email reflectors, regularly scheduled conference calls, and two or three face-to-face meetings per year. There are currently more than 500 people registered across the 11 email reflectors, including a main reflector for general announcements relevant to the entire group and different project reflectors dedicated to technical discussions of specific projects.

¹ Main author's email: quan.huynh-thu@cisra.canon.com.au; address: Building A, The Park Estate, 5 Talavera Road, 2113 Macquarie Park, Australia.

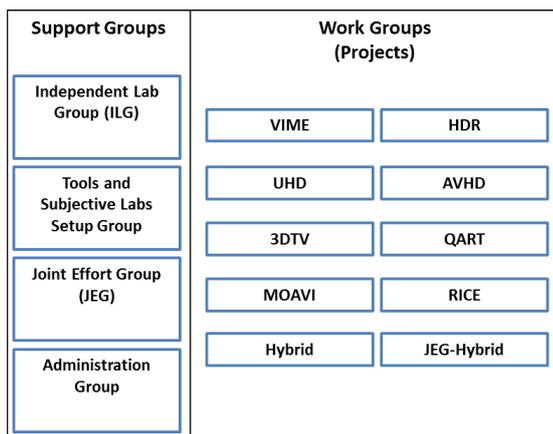


Fig. 1 VQEG structure

VQEG is composed of two types of groups: Support Groups and Work Groups, as shown in Fig. 1.

2.1 Support Groups

These groups provide support to any of the on-going projects in the Work Groups. The support groups provide help in many ways such as providing source content, producing test content, designing and conducting subjective experiments, conducting statistical analysis of data, and developing subjective testing software.

Independent Lab Group (ILG) helps make sure that each test to validate a class of assessment methods (e.g. objective video quality models) gives a trustworthy result, is fair to the algorithm developers, and meets industry needs. All validation test plans are designed and executed under the supervision of the ILG. The ILG will typically be involved in the design of the experiments, the development of the validation test plan, and the data analysis. The ILG is composed of organisations that do not propose algorithms, and do not benefit commercially or otherwise from the results of the validation tests.

Tools and Subjective Lab Setup Group develops and collects assets (e.g. software tools) to help meet the needs of a specific VQEG project. These tools are usually subsequently made available to the public for academic research or commercial research.

Joint-Effort Group (JEG) is composed of organisations working collaboratively to develop new video quality assessment algorithms, conduct collaborative studies, and increase expertise.

Administration Group is composed of the VQEG co-chairs. It deals with the general administration of the organisation, maintains the website, and provides external communications.

2.2 Work Groups

VQEG's work is organised in Work Groups (WG). This section provides a bird's-eye view of the WGs currently active (see also VQEG's webpage [1] and latest

status report [2]). Each WG runs one or more projects dedicated to a specific scope or a topic.

Video and Image Models for consumer content Evaluation (VIME) develops no-reference methods to assess the quality of images and videos captured by consumer digital cameras. VIME is the most recently formed VQEG WG and project. It was proposed in July 2014, holds regular conference calls, and is currently defining its first work items and targets.

High Dynamic Range Video (HDR) develops methods to assess the quality of HDR and HDR tone-mapped content. For that purpose, databases of subjectively annotated HDR compressed images [3] and eye-tracking data have been produced to understand the differences between viewing LDR and HDR images. This WG also investigates how to reliably evaluate the subjective quality of experience of HDR and HDR tone-mapped content, including aspects such as naturalness.

Ultra High Definition (UHD) addresses the quality of 4K/8K video. The scope of work is currently under discussion.

Audiovisual High Definition (AVHD) addresses the combined audiovisual quality of video accompanied by sound. This WG is currently defining a test plan to evaluate audio-video quality metrics and investigating improved subjective testing approaches. One recent study assessed the influence of the testing environment [4]. Results from this project have recently produced a standard for subjective testing [5]. The AVHD WG also works on video quality for adaptive video streaming.

3DTV addresses how to assess 3DTV quality of experience. The WG is currently conducting a large-scale study to investigate suitable subjective testing methodologies and the influence of the test lab environment. A very large ground-truth database is being developed and will contain accurate subjective results. When this database is available, new subjective methodologies can be compared to this ground-truth data to assess the merits of such new methodologies. The next target will be the evaluation of objective video quality assessment models for stereo 3DTV. Additionally, this WG is currently collaborating with the European Digital Video Broadcast (DVB) consortium to compare different frame-compatible 3D formats. Results will be used by DVB to define future broadcast 3D formats in Europe.

Quality Assessment for Recognition and Task-based Multimedia Applications (QART) focuses on determining the video quality requirements for human recognition tasks, where object, face, or character recognition is critical. This project produced a new standard method for subjective testing [6]. QART is

currently conducting further studies to improve the standard and develop objective models for predicting video quality in the context of object recognition tasks.

Monitoring of Audio Visual Quality by Key Indicators (MOAVI) collaboratively develops no-reference key performance indicators (KPI) of different types of video artefacts for monitoring quality of audio-visual network services. Several KPIs have been developed [7] and experiments are being conducted to evaluate their accuracy.

Real-Time Interactive Communications Evaluation (RICE) develops new methods for subjective assessment and objective measurement of interactive communications services. In this effort, novel investigation methods based on psycho-physiological measurements are considered in addition to the more established subjective testing methodologies for interactive communication services.

Hybrid Perceptual/Bitstream (Hybrid) addresses video quality assessment models using both the encoded bitstream and the decoded video information as inputs. This WG executed a test plan that evaluated a number of proposed models. Evaluation results were finalised in July 2014 and have recently produced a new set of ITU recommendations [8].

JEG-Hybrid is the flagship project of the JEG group and is an open collaboration to develop a hybrid (perceptual/bitstream) video quality assessment model for AVC and HEVC encoded videos. The goal is to produce one objective quality estimation model that combines metrics developed separately by different contributing organisations. This project has created a set of 10,000 encoded videos to train objective metrics for video quality assessment.

3. Objective video quality validation testing

VQEG has expended tremendous effort into assessing the performance of candidate algorithms for video quality assessment for the purpose of standardisation. VQEG has conducted more than 60 subjective experiments in test labs located around the world. These efforts have both advanced best practices for subjective testing and provided data for the validation of objective video quality assessment algorithms in the context of standards development.

VQEG validates objective video quality models to directly support the ITU and industry needs. The goal is to assess each model's ability to predict video quality as rated by a panel of viewers. The VQEG validation process emphasises integrity and reliability, which may impact speed of project progress.

VQEG intentionally makes no decision as to which model should be standardised. That is the responsibility of the SDO. VQEG only conducts validation tests, reports the results to SDOs, and makes the reports publicly available on the VQEG website: www.vqeg.org.

A key ingredient in validation is to ensure that the test content is unknown to the models under evaluation. It is important not to reuse training data for the model validation. Novel data is needed to characterise how the model performs in a general scenario. In general, the purpose of an objective video quality model is to predict subjective results. These models are typically based on tuneable perceptual parameters that are optimised or trained using subjective data. VQEG creates several new subjective databases for use during the validation process.

VQEG test plans define exact procedures for performing model validation. These test plans describe the scope of the validation project, characteristics of source content, the scope and nature of video quality degradations, the subjective rating method, the subjective test environment, restrictions on the objective quality models, submission procedures, and model performance evaluation metrics; they also specify which test laboratories will conduct the test. Importantly, the test plans are worked out and approved by consensus in advance amongst all VQEG participants, which include model proponents, subjective test laboratories, industry representatives, academia, and representatives of several SDOs. Proponents are organisations that propose new objective models to be evaluated.

Roles and responsibilities are clearly delineated between proponents and the ILG. The ILG creates new and unknown content to evaluate the model and provides test oversight to ensure a fair and unbiased evaluation. Subjective test laboratories in the ILG conduct the experiments that produce the subjective data used to evaluate the proposed technologies. Subjective test laboratories are usually part of the ILG but proponents also conduct subjective tests in some projects. The ILG provides oversight to all subjective tests and is the main safeguard of the independence of the evaluations conducted by VQEG. An organisation can be a technology proponent in one project but at the same time act as an independent laboratory in another project.

A list of VQEG validation tests, final reports, and resulting ITU Recommendations can be found on the VQEG website by selecting the link "Reports". These validation tests are typically competitive tests in which algorithms proposed by different proponents are evaluated on the same subjective databases for direct comparison. Some recent projects use a collaborative

approach, where organisations work together in the development of new tools and algorithms.

4. Role of VQEG in standards development

VQEG does not develop or publish standards. VQEG conducts carefully designed and independent technical evaluations, and reports results to SDOs. Based on these results, SDOs develop international Recommendations and standards. Since 2000, the International Telecommunication Union (ITU) has approved more than 25 Recommendations based on technical results produced by VQEG (see also [9]). Table 1 shows the different VQEG projects that have produced international Recommendations on subjective or objective video quality assessment methodologies or algorithms (dates indicate the latest version of each international Recommendation).

Table 1. ITU Recommendations based on VQEG's projects.

VQEG project	ITU-T Rec.	ITU-R Rec.
FRTV-I	J.143 (2000) J.144 (2001) J.149 (2004)	
FRTV-II	J.144 (2004)	BT.1683 (2004)
Multimedia I	J.148 (2003) P.910 (2008) J.246 (2008) J.247 (2008) J.340 (2010)	BT.1866 (2010) BT.1867 (2010)
RRNR-TV	J.244 (2008) J.249 (2010)	BT.1885 (2011)
HDTV-I	J.341 (2011) J.342 (2011)	
QART	P.912 (2008)	
Hybrid/Bitstream	J.343 (2014) J.343.1-6 (2014)	
3DTV	J.3D-disp-req J.3D-fatigue P.3D-sam (expected 2015)	
AVHD	P.913 (2014)	

Historically, VQEG has had a very close relationship with ITU-T Study Group 9 (SG9). As a result, most Recommendations so far produced from VQEG results have been developed by ITU-T SG9. However, VQEG has also established liaisons with other SDOs and consortiums such as ITU-T SG12 and SG16, MPEG, ITU-R SG6, ATIS, IEEE P.3333 and P.1858, DVB, and ETSI.

Each VQEG meeting hosts an Intersector Rapporteur Group on Audiovisual Quality Assessment (IRG-AVQA) meeting. The IRG-AVQA provides VQEG an official conduit to the ITU. This official relationship helps VQEG contribute to the ITU standardisation process.

5. VQEG eLetter

Currently, there is no journal dedicated specifically to the field of perceptual video quality. For this reason, in 2014 VQEG started to publish an eLetter [10] to encourage timely publication of the latest developments, trends, challenges, and technical advances in research related to video quality. The eLetter is published once or twice a year. Each issue includes papers contributed by invited authors or authors responding to a call-for-papers.

6. Conclusion

VQEG provides a unique forum for exchanges between technical experts from different areas of the industry and academia, all working towards the same goal of advancing the field of perceptual visual quality assessment. With its current and future projects, VQEG will continue to play a fundamental role in the development of standards on subjective and objective video quality assessment by providing large-scale independent evaluations of new video quality assessment technologies.

References

- [1] Video Quality Experts Group: Projects, accessible at <http://www.its.bldrdoc.gov/vqeg/projects-home.aspx>
- [2] Video Quality Experts Group: Progress Report, accessible at: <http://www.vqeg.org>
- [3] M. Narwaria M. Perreira Da Silva, P. Le Callet, R. P epion: Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality, *Optical Engineering*, vol. 52, no. 10, pp. 1-15 (2013).
- [4] M. Pinson, L. Janowski, R. P epion, Q. Huynh-Thu, C. Schmidmer, P. Corriveau, A. Younkin, P. Le Callet, M. Barkowsky, W. Ingram: The influence of subjects and environment on audiovisual subjective tests: an international study, *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 6, pp. 640-651 (2012).
- [5] International Telecommunication Union: ITU-T Rec. P.913 Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment (2014).
- [6] International Telecommunication Union: ITU-T Rec. P.912 Subjective video quality assessment methods for recognition tasks (2008).
- [7] MOAVI metrics accessible at: <http://vq.kt.agh.edu.pl>
- [8] International Telecommunication Union: ITU-T Rec. J.343 Hybrid perceptual bitstream video quality assessment (2014).
- [9] K. Brunnstr om, D. Hands, F. Speranza, and A. Webster: VQEG Validation and ITU Standardisation of Objective Perceptual Video Quality Metrics, *IEEE Signal Processing Magazine*, vol. 26, pp. 96-101 (2009).
- [10] Video Quality Experts Group: eLetter, accessible at: <http://www.its.bldrdoc.gov/vqeg/eletter.aspx>