VIDEO QUALITY EXPERTS GROUP

Progress report 2015 v.1

For more information contact:

Kjell Brunnström	kjell.brunnstrom@acreo.se	Co-Chair VQEG
Margaret Pinson	margaret@its.bldrdoc.gov	Co-Chair VQEG
Arthur Webster	webster@its.bldrdoc.gov _	Co-Chair VQEG

Introduction

The Video Quality Experts Group was first grounded in basic subjective methodology and objective tool development/verification for video quality assessment such that the industry could be moved forward with standardization and implementation. At the beginning it was focused around measuring the perceived video quality since the distribution path for video and audio were limited and known.

Over the last 20 years from the formation of VQEG the ecosystem has changed dramatically and thus so must the work. Multimedia is now pervasive on all devices and methods of distribution from broadcast to cellular data networks. This shift has the expertise within VQEG to move from the visual (no-audio) quality of video to Quality of Experience (QoE).

The march forward of technologies means that VQEG needs to be react and be on the leading edge of developing, defining and deploying methods and tools that help address these new technologies and move the industry forward. This also means that we need to embrace both qualitative and quantitative ways of defining these new spaces and terms. Taking a holistic approach to QoE will enable VQEG to drive forward and faster with unprecedented collaboration and execution.

Very good progress was achieved during 2015. The 3DTV group completed all goals, leading to a series of three new recommendations within the ITU. To demonstrate the change in VQEG's focus, all three Recommendations focus on subjective test methods for 3DTV.

Two regular meetings were held during 2015. One was hosted by Intel in Santa Clara, California, USA, February 23-27. There were 40 participants, including 12 remote participants. The other meeting was hosted by the University of the West of Scotland in Glasgow, Scotland, September 14-18. A total of 37 people participated in this meeting, either physically or remotely.

In 2015, VQEG work emphasized collaboration to develop new and improved methods for measuring video quality. Meetings emphasized presentations that described research underway or newly finished.

VQEG has 10 active workgroups and 4 support groups. Their activities and progress during 2015 will be described below in this report.

Figure 1: Group photo from the meeting in Glasgow, UK, hosted by University of the West of Scotland.

Active Workgroups

3DTV (3DTV)

Co-chairs: Marcus Barkowsky (IRCCyN), Patrick LeCallet (IRCCyN), Quan Huynh-Thu (CiSRA)

Goals:

The 3DTV Project investigates how to assess 3DTV subjective video quality.

Summary:

This group accomplished all major in 2015 and so was formally closed in early 2016. The remaining 3DTV project was moved to the new immersive media group (IMG) in early 2016.

1. Define suitable methodologies for subjective quality assessment of stereoscopic 3D video: A unique set of video sequences is being produced to conduct the experiments (GroTruQoE dataset). The first step is to conduct a large-scale experiment using the pair-comparison methodology, as we are confident that subjects can provide easily a judgment of overall preference. The second step will use the results of the pair-comparison testing as a groundtruth database to investigate which more time-efficient subjective testing methodology can be used to predict the results of the pair-comparison test.

Achievements 2015

The 3DTV group collaborated to develop subjective test methods, display requirements, and quality of experience guidelines. VQEG worked with ITU-T Study Group 9 to contribute this information to three new ITU Recommendations. These Recs. were approved in early 2016.

ITU-T Rec. P.914, “Display Requirements for 3D Video Quality Assessment”

ITU-T Rec. P.915, “Subjective Assessment Methods for 3D Video Quality”

ITU-T Rec. P.916, “Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video”

Plans for 2016

Future development will take place in Immersive Media Group (IMG)

Audio-Visual HD Quality (AVHD)

Co-chairs: Chris Schmidmer (OPTICOM), Quan Huynh-Thu (CiSRA), Margaret Pinson (NTIA/ITS)

Goals:

The AVHD Quality workgroup has two active projects:

- Objective models for adaptive streaming (AVHD-AS)
- Advanced subjective methods (AVHD-SUB)

Achievements 2015

- Objective methods for adaptive streaming (AVHD-AS):
    - The goal of the AVHD-AS project is to validate objective methods for the assessment of adaptive bitrate streaming services from an end users perspective.
    - Targeted scenarios include, but are not limited to streaming of video sequences over fixed and mobile networks, with a focus on consumer entertainment. The objective models are meant to be applied to sequences of up to 5 min duration and yield one quality rating. Full reference, reduced reference, and no reference models will be evaluated. Hybrid models are under consideration.
    - It was decided to modify the duties and responsibilities of proponents and ILG. The goal is to improve the speed of model validation while maintaining a careful and trustworthy validation process.
    - A call for participation was issued.
    - The validation test plan is already very advanced and remaining Details will be established in Q2 2016.
    - Model submission deadline is expected to be in Q3/2016.
- Advanced subjective methods (AVHD-SUB):
    - This goal is to develop improved subjective test methods.
    - In 2015, a collaborative effort began to validate new experiment designs that avoid scene re-use. Such an experiment design is particularly important for adaptive streaming, where video sequences are typically 1 to 5 minutes duration. For such tests, it is not appropriate for the subject to view multiple versions of the same sequence.
    - Subjective test results are expected in 2016.

Joint Effort Group - Hybrid Perceptual Bit-Stream Measurement (JEG-Hybrid)

Co-chairs: Marcus Barkowsky (University of Nantes-IRCCyN), Lucjan Janowski (AGH University), Glenn Van Wallendael (Ghent University-iMinds)

Goals:

The JEG Hybrid Group is an **open collaboration** working together to develop a **robust Hybrid Perceptual/Bit-Stream video quality measurement tool**.

Perceptual video quality measurement has been tackled by many researchers in the past. JEG-Hybrid aims to **unite their independent work into a framework** where individual ideas of modeling the Human Visual System become verifiable and measurable.

Bit-Stream information may be seen as a primary or supplemental information source for video quality assessment. As a primary source, properties such as bit-rate or the quantization parameter may be exploited. As a supplemental source, the video bit-stream delivers information that is otherwise often calculated on the decoded video, notably motion information or frequency analysis. **Rigorous analysis of the value of these information sources** is required. Within JEG-Hybrid bit-stream information is easily exploited beyond its current usage in parametric models and bit-stream models by using a simple interface for interpreting its data in XML files.

Hybrid approaches combine various indicators. These indicators may be quality indicators, degradation indicators, content indicators, and so on. Challenges include the **frame-exact synchronization of bit-stream information with perceptual information** obtained from the decoded video, or innovative approaches on **stable combinations of a large number of indicators** on a comparably small training set in machine learning algorithms.

Robustness of video quality measurement stems from verification and validation on a large dataset. The JEG-Hybrid team maintains and continuously improves a **large-scale database of encoded video sequences** with recent video coding algorithms. Innovative research is performed to **verify prediction performance** of objective measurements with this database because subjective assessment can only be performed on a comparably small and well-chosen subset.

Achievements 2015

The JEG-Hybrid group has advanced on:

- **content characterization**: tools have been produced to calculate Spatial information, Chrominance Information, Contrast information, Spatial perceptual information, Colorfulness,

Gray-Level Co- occurrence Matrix (GLCM), Normalized cross correlation, DCT based features, Laplacian based Features, Temporal information, and MPEG-7 Motion Activity

- **bitstream based indicators:** tools to calculate bitstream features like AvgQP, AvgMVy, and AvgMVx have been produced [7]
- **Objective Measurements**: Several Full-Reference measures have been added on the HEVC database and the lossy HEVC database, P1201.2 scores have been calculated for the AVC database
- **Statistical analysis** of objective Full-Reference measures towards understanding the scope of application of widely used objective measurement algorithms and towards understanding which data needs to be subjectively evaluated [4-6]
- **Reproducible research**: Freely available degradation environment for real-world transmission scenarios, further development of the VirtualBox collaborative environment, inclusion of MOAVI indicators, MySQL database creation for efficient access of the annotations, HRCs, and SRCs.
- **Documentation**: Three scientific publication in 2015, Further details added to the JEG-Wiki pages at http://vqegjeg.intec.ugent.be (collection of SRC locations, Robust decoder description, …)

Plans for 2016

In 2016, progress will be made on enlarging the large-scale database, notably by adding further contents, their characteristics, and their degraded versions, the video quality metric development and the infrastructure for reproducible research. More specifically, the database will get extended and become more specific towards the application of low latency interactive video communication.

References

1. Barkowsky, M., Sedano, I., Brunnström, K., Leszczuk, M., & Staelens, N. (2014). Hybrid video quality prediction: reviewing video quality measurement for widening application scope. Multimedia Tools and Applications, 1–21.
2. Shahid, M., Rossholm, A., Lövström, B., & Zepernick, H. - J. (2014). No-reference image and video quality assessment: a classification and review of recent approaches. EURASIP Journal on Image and Video Processing, 2014(1), 40.
3. Van Wallendael, G., Staelens, N., Masala, E., Janowski, L., Berger, K., & Barkowsky, M. (2014). Dreamed about training, verifying and validating your QoE model on a million videos? VQEG eLetter, 1(2), 19–29.
4. Mikolaj Leszczuk; Lucjan Janowski; Marcus Barkowsky. Freely Available Large-scale Video Quality Assessment Database in Full-HD Resolution with H.264 Coding, Globecom 2013

5. Barkowsky, M., Masala, E., Van Wallendael, G., Brunnstrom, K., Staelens, N., & Le Callet, P. (2015). Objective Video Quality Assessment — Towards Large Scale Video Database Enhanced Model Development. IEICE Transactions on Communications, E-98b(1), 2–11.
6. Van Wallendael, G., Staelens, N., Masala, E., & Barkowsky, M. (2015). FullHD HEVC-Encoded Video Quality Assessment Database. Ninth International Workshop on Video Processing and Quality Metrics (VPQM)
7. Shahid, M., Panasiuk, J., Van Wallendael, G., Barkowsky, M., & Lovstrom, B. (2015). Predicting full-reference video quality measures using HEVC bitstream-based no-reference features. In Quality of Multimedia Experience (QoMEX), 2015 Seventh International Workshop on (pp. 1–2).

HDR (High Dynamic Range)

Patrick LeCallet (IRCCyN), Elaine Jin and Philip Corriveau (Intel Corp)

Goals

Human eyes can perceive approximately 5 orders of luminance magnitude at a time, but current capture and display technologies are limited to at most 3 orders of magnitude.

The aim of HDR imaging: capture and display these 5 orders of magnitude with high fidelity.

The aim of the VQEG HDR project is to develop methods for assessing the quality of HDR video.

HDR technologies aim at filling the gap between capture and display technologies and the abilities of the human visual system. This will provide a more realistic visual experience compared to current Low Dynamic Range (LDR) imaging. All of the current technology on the market is LDR but it is felt that within the next few years HDR will be ready for commercial release and we need to be ahead of it to understand the user benefit and how to measure quality.

Achievements 2014 & 15

There were 4 main accomplishments completed

First was that the group created a database that has enabled the testing for how tone mapping affects human attention and modification of artistic intent. This work enabled a comprehensive evaluation of objective methods and solid statistical analysis was completed. The database is available for other researchers or efforts. It has been shown that TMOs influence the VA for images more than for videos. The relationship of the VA with the resulting quality is not linear. Also, the performance of the three tested automatic saliency estimators has been proven not sufficient to replace human observers in the given application.

Reference

Manish Narwaria and al. IMPACT OF TONE MAPPING IN HIGH DYNAMIC RANGE IMAGE COMPRESSION. VPQM, Jan 2014, Chandler, United States. pp. 1-6, 2014.

Manish Narwaria and al. Tone mapping based HDR compression: Does it affect visual experience?. Signal Processing: Image Communication, Elsevier, 2014, 29 (2), pp.257-273.

Datasets

http://ivc.univ-nantes.fr/test/en/databases/ETHyma/

Secondly another dataset was created for the investigation of tone mapping operators for quality. This yielded two results around the use for investigating quality issues in local and global HDR codec optimization. Also it enabled the creation of a database with corresponding subjective scores. It is still up for consideration on if the database will be able to made available for other researchers or efforts. A study comparing various TMOs with single-exposure photograph in representing the HDR scene has been conducted. Surprisingly, no statistically significant difference has been found, suggesting that the extra details provided by the TMOs do not suffice to better represent the scene if the naturalness is not maintained. We also investigated the methods for content selection for HDR studies and the influence of the experimental design on the final outcome . These studies lead to the preparation of a novel, representative dataset of tone-mapped images.

Reference

Manish Narwaria and al. SINGLE EXPOSURE VS TONE MAPPED HIGH DYNAMIC RANGE IMAGES: A STUDY BASED ON QUALITY OF EXPERIENCE. 22nd European Signal Processing Conference (EUSIPCO), Sep 2014, Libon, Portugal.

Manish Narwaria and al. An objective method for High Dynamic Range source content selection. Sixth International Workshop on Quality of Multimedia Experience (QoMEX), Sep 2014, Singapore, Singapore. 2014.

L. Krasula and al. “Influence of HDR reference on observers preference in tone-mapped images evaluation,” 7th International Workshop on Quality of Multimedia Experience (QoMEX), 2015.

Dataset

http://ivc.univ-nantes.fr/test/en/databases/PairCompTMO/

http://ivc.univ-nantes.fr/test/en/databases/JPEG_HDR_Images/

Third there was a drive in the core metrics development for HDR. Tools were created for objective HDR assessment validation of a new version of the tool HDR-VDP-2.2 which has been released for use in the industry.

Reference

Manish Narwaria and al.. On Improving the Pooling in HDR-VDP-2 towards Better HDR Perceptual Quality Assessment. Human Vision and Electronic Imaging 2014, Feb 2014, San Francisco, United States.

Manish Narwaria and al.. HDR-VDP-2.2: A calibrated method for objective quality prediction of high-dynamic range and standard images: Journal of Electronic Imaging, Society of Photo-optical Instrumentation Engineers, 2014, 24 (1), pp.010501. <http://dx.doi.org/10.1117/1.JEI.24.1.010501>

Lastly the group completed testing in support of validation of an HDR metric extended to Video. All previous tools and metrics have been for still images. HDR-VQM has been extended to enable the algorithms scale to motion video.

Reference

Manish Narwaria et al., “HDR-VQM: An Objective Quality Measure for High Dynamic Range Video.”, Signal Processing: Image Communication, Elsevier, 2015, 35, pp.46-60. <10.1016/j.image.2015.04.009>

Manish Narwaria et al., “Study of high dynamic range video quality assessment.”, Proc. SPIE 9599, Applications of Digital Image Processing XXXVIII, 95990V (September 22, 2015); doi:10.1117/12.2189178

Manish Narwaria et al., “High Dynamic Range Visual Quality of Experience Measurement: Challenges and Perspectives.”, Visual Signal Quality Assessment - Quality of Experience (QoE), Springer International Publishing, pp.129-155, 2015, 978-3-319-10368-6.

Manish Narwaria et al.,”Quality of Experience and HDR: Concepts and How to Measure it.”, High Dynamic Range Video. Elsevier Ltd. 2016.

Rafal Mantiukand et al., “HDR Image and Video Quality Prediction.”, High Dynamic Range Video. Elsevier Ltd. 2016.

Plans for 2016

The group has been renamed HDR-WCG to cover Wide Color gamut application scenario

Current Objectives are

- Extend HDR-VQM profile to support: WCG conditions, Ultra High definition
- Evaluate robustness of HDR-VQM on new HDR video compression scheme
- Perform subjective experiments related to WCG scenarios, especially color volume processing
- Identify a way to build out a research data base from which the community can benefit from
- Drive the consideration of HDR-WCG into the Vime group for consideration in the inclusion of the no-reference image tool released in 2015
- Define the classification of usages for this category of images and videos so that guidance can be provided to the industry.

Monitoring of Audio Visual Quality by Key Indicators (MOAVI)

Co-chairs: Silvio Borer (SwissQual) and Mikołaj Leszczuk (AGH)

Goals:

The MOAVI group is an open collaborative for developing No-Reference models for monitoring audio-visual service quality. The goal is to develop a set of key indicators (e.g. blocking effects, blurring effects, freeze/jerkiness effects, ghosting effects, slice videostripe errors, aspect ratio problems, field order problems or photosensitive epilepsy flashing effects, silence, clipping) describing service quality in general and to select subsets for each potential application. Therefore, the MOAVI project concentrates on models based on key indicators contrary to models predicting overall quality.

Achievements 2015

As a result of the MOAVI project metrics and tools were integrated for the assessment of individual distortion of the (moving) image and sound like an blockiness artifact or blur in a system measuring the quality of video sequences, allowing validation of the achieved results. Due to the characteristics of the design, the No-Reference (NR) metric function models quality based on pixels and audio samples, rather than based on packets.

Analysis of the distortions comes with the implementation of a number of different distortion metrics, such as blockiness artifact, blur, exposure, interlacing, noise, frame flickering, brightness, pillarboxing, letterboxing, contrast, block loss, mute and clipping. It also incorporated in two metrics for dynamics: temporal activity (the amount of motion) and spatial activing (level of detail) of the video.

The calculation of all these metrics have been optimized to the sequence playback real-time.

The created system was made available free of charge (for non-commercial purposes) in the form of software downloadable from the website: [VQ]. The software was prepared in versions for Linux, Mac OS and Windows (32-bit and 64-bit). The software can use multiple processor cores, which accelerates the execution.

The developed system allows one to calculate a number of metrics for different video sequences played on different devices and allows one to create a database for estimation metrics used in the JEG-Hybrid project, which will be one of the final results.

Plans for 2016

- H.264:
    - More precise evaluation of exposure detection

    - Development of frame drop detection functionality
    - Constant development of overlaying results onto video
- H.265:
    - Further experimenting with new codec
- Funding only until end of September 2016
- New idea: standardising quality indicators (spanning beyond funding period)
- Discussed commercial exploitation by Net Research

References

[Video_Quality_website] VQEG JEG (2014), *Video Quality website*, VQEG JEG Wiki.
•http://vqegjeg.intec.ugent.be/wiki/index.php/Video_Quality_website•

[VQ] (2016), *Video Quality*.
•http://vq.kt.agh.edu.pl•

Quality Assessment for Recognition and Task-based multimedia applications (QART)

Co-chairs: Mikołaj Leszczuk (AGH)

Goals

Users of video to perform tasks require sufficient video quality to recognize the information needed for their application. Therefore, the fundamental measure of video quality in these applications is the success rate of these tasks (such as recognition), which is referred to as visual intelligibility or acuity. One of the major causes of reduction of visual intelligibility is loss of data, through various forms of compression. Additionally, the characteristics of the scene being captured have a direct effect on visual intelligibility and on the performance of a compression operation-specifically, the size of the target of interest, the lighting conditions, and the temporal complexity of the scene. The QART project is performing a series of tests to study the effects and interactions of compression and scene characteristics. An additional goal is to test existing or develop new objective measurements that will predict the results of the subjective tests of visual intelligibility.

Achievements 2015

Problems of quality evaluation procedures and measurements for Target Recognition Videos (TRV) have been partially standardized in ITU-T Recommendation P.912, entitled: *"Subjective Video Quality Assessment Methods for Recognition Tasks"* (published in 2008). P.912 introduces basic definitions, methods of testing and psycho-physical experiments.

Based on research and observations with VQEG, since 2014, QART has introduced revisions to P.912. Last year, QART continued to participate in ITU-T Study Group 9 (SG9) P.912 standardization activities. QART took part in the meeting of the SG9 in Beijing, China (June 2015) and Geneva, Switzerland (January 2016).

During the ITU-T SG9 meeting in Beijing, QART revised:

- Whole Recommendation (conditions for testing)
- Clause 6 ("Test Methods")
- Clause 6.1 ("Multiple Choice Method")
- Clause 6.2 ("Single Answer Method")
- Clause 7.4 ("Instructions to Subjects")
- Clause 8 ("Statistical Analysis & Reporting")

During the ITU-T SG9 meeting in Geneva, QART revised:

- Clause 7.5 (“Crowdsourcing Environment”)
- Clause 8.1 (“Data Analysis”)

QART was also seeking there for final consent (approving, closing).

Plans for 2016

Revised Recommendation P.912 (03/16) has been already approved in 2016-03 and it is in force now. In 2016, P.912 is going to be published.

References

[P.912] *P.912 : Subjective video quality assessment methods for recognition tasks*
•https://www.itu.int/rec/T-REC-P.912•

PsyPhyQA former RICE (Real-Time Interactive Communications Evaluation)

Co-chairs: Sebastian.Arndt (Technische Universität Berlin), Ulrich Engelke (CSIRO), Naeem Ramzan (University of the West of Scotland) and Kjell Brunnström (Acreo Swedish ICT AB)

Goals

The PsyPhyQA the former Real-Time Interactive Communications Evaluation (RICE) Project is directed towards the development of new methodologies for subjective assessment and objective measurement of interactive communications services. In this effort novel investigation methods based on psycho-physiological measurements are considered..

Psychophysiological measurements

Video quality assessment is typically performed using questionnaires, either open-ended or based on psychometric scales, such as n-point Likert scales. As valuable as these studies are, they are based on conscious responses by the participants and often do not provide sufficiently deep insight into underlying perceptual and cognitive processes. In order to gain a deeper understanding about the perceptual and cognitive processes underlying video quality perception, psychophysiological measurements can be performed. For instance,

- Eye gaze tracking provides valuable information about overt visual attention in visual space,
- Electroencephalography (EEG) measurements inform about cognitive activity, such as cognitive load, situational awareness, emotional responses, and covert attention, and
- Galvanic skin response (GSR) provides insight into arousal and hence emotional states.

These psychophysiological responses are not intended to replace well established psychophysical assessment techniques, but to augment them and provide additional sub-conscious information.

- The aim of the RICE project is to establish novel psychophysiology-based techniques and methodologies for video quality assessment and real-time interaction of humans with advanced video communication environments. Specifically, some of the aspects that the project is looking at include video quality assessment based on human psychophysiology, including, eye gaze, EEG, EKG, EMG, GSR, etc,
- computational video quality models based on psychophysiological measurements,
- signal processing and machine learning techniques for psychophysiology-based video quality assessment,
- experimental design and methodologies for psychophysiological assessment, and
- correlates of psychophysics and psychophysiology.

Achievements 2015

The PsyPhyQA working group has met regularly online during 2015. A plan for a cross lab video quality experiment with EEG was developed during the online meetings and presented at the Santa Clara meeting in February 2015. Tentative interest from the organisations (Netflix) NTIA/ITS and University of West Scotland were expressed after the presentation.

The planned study aims on the one hand, at comparing results between labs, and on the other hand, at comparing obtained results between different classes of devices (i.e. low-cost consumer grade EEG products vs high-end clinical grade products). Therefore, typical video sequences are used which have a length of several minutes; while subjects are watching these an EEG will be recorded. In a post processing step the recorded data will be analyzed concerning their power spectral energy using this, assumptions about the cognitive state can be drawn. This inter-lab study aims at validating the methodology of EEG in the context of QoE, which may lead in the future of obtaining non-intrusive correlates of quality.

The PsyPhyQA working group has also organized a special session on Psychophysiological Measures for Visual Quality at Human Vision Electronic Imaging, San Francisco, 17 Feb 2016 and submitted an overview journal article to a scientific peer reviewed journal.

Plans for 2016

- Sept 2016: having test plan ready
- Nov 2016: starting subjective tests

Ultra HD

Co-chairs: Vittorio Baroncini (Fondazione Ugo Bordoni), Naeem Ramzan (University of West Scotland), and Chulhee Lee (Yonsei University)

Goals

Three activities are defined within the scope of Ultra HD project: (1) Creation of Ultra HD database, (2) Defining subjective quality testing methodologies for Ultra HD, (3) Objective video quality metrics for Ultra HD.

1. Creation of Ultra HD database:
   10 4K video contents are available on request. It will be shared through external HDD however the source of the contents is http://medialab.sjtu.edu.cn/web4k/index.html and is free to download.
   Initial testing and subjective evaluation is performed for quality evaluation. H.264/MPEG-4, H.265/HEVC and VP9 codecs are used for quality evaluation.
2. Defining subjective quality testing methodologies for Ultra HD
   Work will be carried out in 2016
3. Objective video quality metrics for Ultra HD
   Work will be carried out in 2016

Achievements 2015

In JEG group, there is more than 40 4K UltaHD sequences is available to download

Comparison of AVC, HEVC is performed for different resolutions including 4K and results are presented in the paper given below.

Thiow Keng Tan, Rahitha Weerakkody, Marta Mrak, **Naeem Ramzan**, **Vittorio Baroncini**, Jens-Rainer Ohm, and Gary J. Sullivan, "Video Quality Evaluation Methodology and Verification Testing of HEVC Compression Performance HEVC Subjective Video Quality Test Results", in IEEE Trans. of Circuit and System for Video Technology, Vol. 26, No. 1, Jan, 2016.

Plans for 2016

The work will continue on item 2 and 3 in 2016.

The study between UWS and AGH will be conducted to explore the correlation of subjective data with the full reference objective metrics

VIME (Video and Image Models for consumer content Evaluation)

Co-chairs: Michele Saad (Intel Corp.), Quan Huynh-Thu (CiSRA), James Goel (Qualcomm)

Goals:

No-reference [or blind] image and video quality assessment is still not accurate and robust to be usable in the industry. Current approaches of no-reference image or video quality assessment target the detection and measurement of specific types of degradations. Furthermore, current approaches have focused on the compression and delivery of the content, and not on the capture and rendering of the content. Current approaches also do not consider that image processing can enhance the content, making the quality of the modified image better than the quality of the original image. Increasingly, digital content (images and videos) is being captured by digital cameras and mobile devices, and viewed on a variety of displays and viewing conditions. For these reasons, full-reference approaches are not suited and the requirements for no-reference approaches to measure the quality of images, including capture and rendering aspects, are increasingly important for the industry.

Current databases of images used in research do not contain distortions arising in the capture of images by (compact and high-end) digital cameras and mobile devices (smartphones, tablets). Images and videos captured by these devices contain distortions that are very complex in nature. These distortions may result from the sequence of steps during the capture, from optical system aberrations to the post processing embedded in the devices after the physical light is captured.

The current goals of the VIME project are to investigate and develop:

1) New approaches to subjective study design for the purpose of addressing emerging quality assessment needs (as market and consumer demands evolve).

2) No-reference models for image quality assessment of consumer content with real distortions (as opposed to simulated artificial distortions), including capture and rendering aspects.

3) No-reference models for video quality assessment of consumer content with real distortions, including capture and rendering aspects.

Achievements 2015

The workgroup has had regular conference calls. Meeting minutes are accessible on the VQEG ftp server: ftp://vqeg.its.bldrdoc.gov/Documents/Projects/vime/

Key accomplishments in 2015 include:

- Released the open source no-reference image quality evaluation tool VIQET (VQEG Image Quality Evaluation Tool). VIQET is an objective, no-reference photo quality evaluation tool. VIQET is a free open source tool designed to evaluate quality of consumer photos. In order to perform photo quality evaluation, VIQET requires a set of photos from the test device. It estimates an overall Mean Opinion Score (MOS) for a device based on the individual image MOS scores in the set. The estimated MOS by VIQET falls in a range of 1 to 5, where 1 corresponds to a low quality rating and 5 corresponds to excellent quality.

    - VIQET is an open source project that is available at www.GitHub.com/VIQET.
    - The desktop tool installer can be downloaded at: https://github.com/VIQET/VIQET-Desktop/releases
    - The source code can be found at: https://github.com/VIQET/VIQET-Desktop

- Started the development of the VIME Image database that is freely accessible to the public via Flickr.
    - The current dataset contains more than 1100 images at the beginning of 2016
    - Instructions on how to contribute image to the database can be found in VQEG eLetter (vol.2, issue 1): ftp://vqeg.its.bldrdoc.gov/eLetter/Issues/VQEG_eLetter_vol02_issue1.pdf
    - VIME Flickr site: https://www.flickr.com/groups/vime/

Plans for 2016

- Continue the extension of the VIME Image database
- Investigate more systematic methods to curate and annotate (e.g. image labels) the database content
- Extend the development of VIQET to broader image categories (potentially no-reference HDR image quality evaluation).

Support Groups

Tools and Subjective Labs Setup (STL)

Co-chairs: Glenn Van Wallendael (Ghent University - iMinds), Bert Vankeirsbilck (Ghent University - iMinds)

Goals

The VQEG Tools and Subjective Labs Setup group tries to bundle tools which can aid in the development of subjective quality metrics. These tools can be found on the following website:
http://vqegstl.ugent.be/

Achievements 2015

For 2015, we made available the **ITU-T P.1201.2 Audiovisual Quality Estimation Tool** implemented in Python. It estimates the audiovisual, video, audio coding quality for IP-based video streaming applications. It works on audiovisual bitstream (PCAP) in combination with the modified JM H.264/AVC codec. It also works with encrypted data, but the I/P/B frame sizes have in that case to be estimated, and the performance of the model depends on the accurate estimation of the I/P/B frame sizes.
http://vqegstl.ugent.be/?q=P.1201.2

Additionally, a set of **No Reference metrics** to evaluate the level of distortion (like blocking and blurring) present in a video file is offered by Department of Telecommunications AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków. These metrics includeBlockiness, Blur, Exposure time distortion, Interlace, Noise, Framing, Spatial activity, Temporal activity, Flickering, Blackout, Pillarboxing, Letterboxing, Brightness, Contrast, Slicing, Block loss**,** Mute, and Clipping.
http://vqegstl.ugent.be/?q=Video%20quality%20indicators%20-%20AGH
or http://vq.kt.agh.edu.pl/

For completeness, the following tools which got contributed previous years are still provided on the website:

- VQEGPlayer: a software for performing subjective video quality experiments for Windows 7 in 64bit.

- MATLAB Code for Popular Subject Screening Algorithms including ITU-R Rec. BT.500 Annex 2 Clause 2.3.1, ITU-R BT.1788, VQEG HDTV Phase I Test Plan, and VQEG Multimedia Phase I Test Plan.

- Web-Enabled Subjective Test (WEST) software package offering a solution to the problem of gathering subjective testing data from multiple locations and multiple portable or computing devices.

- Definitely Lossless: a solution to reliably transmit video data from one location to another

- H264AnnexBExtractor: a C++ application extracting the H.264 byte stream from network traces (.pcap files) containing H.264 streaming.

- PcapLossGenerator

- Sirannon: , formerly known as xStreamer, aims at being a modular multimedia streamer and receiver. The modularity is inspired by the Click Modular Router project an Direct Show filters.

- Telchemy PCAP/MPEG2-TS Scrambling Application: This application scrambles and unscrambles live streams and offline PCAP files.

- Telchemy PCAP Loss Insertion Tool: This software introduces losses to a pcap capture file using a 2-state or 4-state Markov model.

- IPTV-interface: Open source interface software for video quality monitoring. The monitoring of the service quality is crucial to the successful operation of commercial internet services to end customers.

- 3D Video Player: a full-featured 3D movie player.

- Modified JM H.264/AVC codec: Version 16.1 of the JM Reference software has been adjusted in order to enable the generation of an XML-based trace file.

Plans for 2016

There are currently no new tools in the pipeline for 2016, but we would like to encourage anyone who wants to make available tools that can help the video quality research community to contact us.

Independent Lab Group (ILG)

Co-chairs: Phil Corriveau (INTEL), Margaret Pinson (NTIA/ITS)

Goals

The goal is to ensure that all VQEG validation testing is unbiased and done to high quality standards. And provide guidance on the structure and direction of subjective research within VQEG.

Achievements 2015

The ILG made a proposal to make the validation process go more quickly. Velocity has always been an opportunity area for VQEG in how we execute our work. One area that the ILG wishes to accelerate is the VQEG objective algorithm validation process. The ILG proposes that the AVHD project try a different division of labor, to try to foster speed and efficiency. This proposal was accepted by the AVHD group. For a full description of the modified ILG and proponent roles, see VQEG_AVHD_2015_111_ILG Proposal to AVHD.docx in the Glasgow 2015 meeting files.

Direction in 2016

In 2016 the role of the ILG continues to evolve to find ways to accelerate the testing process. With all of the collaborative efforts that have been initiated the requirement for test processes to be closely monitored and governed has reduced. We will continue to operate to provide an unbiased and independent direction and monitoring when and where required.

Joint Effort Group (JEG)

Co-chairs: Kjell Brunnström (Acreo Swedish ICT AB), Patrick Le Callet (IRCCyN)

Goals

Promotes the idea of joint collaboration within VQEG.

Achievements 2015

The number of JEG-related projects has increased and now almost all the projects are currently JEG.

Plans for 2016

Proposal is to change VQEG group names to reflect whether or not the effort is currently collaborative, through a “JEG-” prefix.

VQEG Administration and Web Support

Co-chairs: Kjell Brunnström (Acreo Swedish ICT AB), Margaret Pinson (NTIA/ITS), Arthur Webster (NTIA/ITS)

Goals

To take care of issues that do not fall within a VQEG workgroups and answering general questions about VQEG.

Achievements 2015

Planning and organizing two face-to-face meetings: one in Santa Clara, USA, Feb. 23-27, 2015, hosted by Intel Corp. and one in Glasgow, Sept 14-18, 2015, hosted by University of the West of Scotland.

The home page, group pages, and mailing lists have been updated and maintained.

Plans for 2016

Two face-to-face meeting are planned, one in San Diego, USA, Feb. 29- March 4, 2016, hosted by Qualcomm and one in London, United Kingdom hosted by Sky, Oct 24-28.

eLetter

Co-chairs: Margaret Pinson (NTIA/ITS), Naeem Ramzan (University of West of Scotland)

Goals

The VQEG eLetter provides timely updates on recent developments, hot research topics, and society news in the area of video quality.

Achievements 2015

The VQEG eLetter was established in 2014. One issue was organized in 2015, and published in early 2016. This issue focuses on issues related to the VIME group.

Plans for 2016

It was agreed at the VQEG meeting in San Diego, Feb 16, 2016 that for the next issue of the eLetter should cover:

- Immersive Media
- Visually lossless compression