

Final VQEG meeting minutes

Ottawa September 10 - 14, 2007

Including Agenda, Participants List, and Minutes from each day's sessions.

Note: the ITU-T JRG-MMQA meeting is held coincident with VQEG during the Multimedia and Hybrid sessions.

Agenda

Video Quality Experts Group

September 10 - 14, 2007- Meeting

Ottawa, Canada

www.vqeg.org

(Note: The ITU-T JRG-MMQA met coincident with the Multimedia and Hybrid Sessions of the VQEG Meeting.)

Final Agenda

Monday, September 10

- | | |
|------------|--|
| 8:00 | Start / Designate Note Taker (Webster/Speranza)
Introductions
Meeting Logistics
Updates (Maximum 15 minutes - each Group)
Independent Lab Group (ILG)(Brunnstrom/Cermak)
RRNR-TV (Bourret/ Lee)
Multimedia (Brunnstrom/Hands)
HDTV (Corriveau/Pinson)
Tools and Subjective Labs Setup Group (Le Callet)
POC for Source and HRC Sequence collection (Lee)
Hybrid – Perceptual/Bitstream (Juric/Lee)
Calibration Verification (Webster/Lee) |
| 10:00 | Begin MM/JRG-MMQA |
| 11:30-1:00 | Lunch |
| 1:00 | MM and JRG-MMQA
1) Status and Exchange of PVSs |

- a) Procedure for exchanging PVSs that don't meet testplan requirements regarding gain, offset, alignment, etc.
- 2) Review of test design
- 3) Common PVS
- 4) Review of changes to testplan (e.g. 50% requirement on own PVSs per proponent test now removed and option D agreed where proponents create and run their own PVSs alongside the comment set)
- 5) Data Analysis
- 6) Screening
- 7) Review the player
- 8) Review the schedule and identify any slippage

5:00 End of Day

Tuesday, September 11

8:00 Start / Designate Note Taker
Review Monday's Decisions (Webster/Speranza)
MM Continues

5:00 End of Day

Wednesday, September 12

8:00 Start / Designate Note Taker
Review Tuesday's Decisions (Webster/ Speranza)
Review MM Decisions

10:30 HDTV

11:30 Lunch

1:00 HDTV

5:00 End of Day

Thursday, September 13

8:00 Start / Designate Note Taker
Review Wednesday's Decisions (Webster/ Speranza)

8:30 HDTV Wrap up

9:30 RRNR-TV

11:30-1:00 Lunch
1:00 – 5:00 Hybrid/Bitstream Project Discussion
5:00 End of Day
7:00 Dinner

Friday September 14

8:00 Start / Designate Note Taker (Christian Schmidmer Opticom)
Review Thursday's Decisions (Webster/ Speranza)
Finish Hybrid (1 hr)
HDTV Monitor (10min)
MM Data Analysis (5 min)

8:30 Review and Documentation of VQEG Decisions
Review action items

10:00 Other Business
Advertisement policy on VQEG reflector
Project for Collaborative model development
Process (Phil/Kjell, 5 minutes)
Replacement for Co-Chair Subj Lab and Tools Group
Next Meetings (Belgium, Japan)
 Between Feb 11- Mar 19 start
 Blackout Feb 18
 June 30 – July14 start
Conference Calls
 Hybrid
 MM 15 Oct
 HDTV

11:30-1:00 Lunch

1:00-3:00 Write Liaison statements to ITU, ATIS-IIF, PRQC, etc.
3:00 Close VQEG meeting

Video Quality Experts Group
September 10– 14, 2007
Ottawa, Canada
Final Participants List

	Name	Name	Organization	email	
1	Horita	Yuukou	University of Toyama	horita@eng.u-toyama.ac.jp	x
2	Kawayoke	Yoshikazu	University of Toyama	d0671002@ems.u-toyama.ac.jp	x
3	Gustafsson	Jörgen	Ericsson	jorgen.gustafsson@ericsson.com	x
4	Brunnstrom	Kjell	Acreeo	Kjell.Brunnstrom@acreeo.se	x
5	Ford	Carolyn	NTIA/ITS	cford@its.bldrdoc.gov	x
6	Pinson	Margaret	NTIA/ITS	Margaret@its.bldrdoc.gov	x
7	Webster	Arthur	NTIA/ITS	webster@its.bldrdoc.gov	x
8	Adams	Bruce	Telchemy	bruce.adams@telchemy.com	x
9	Le Callet	Patrick	University of Nantes	patrick.lecallet@univ-nantes.fr	x
10	Cermak	Greg	Verizon	greg.cermak@verizon.com	x
11	Barkowsky	Marcus	Opticom	Marcus.Barkowsky@nt.e-technik.uni-erlangen.de	x
12	Ferguson	Kevin	Tektronix	kevin.m.ferguson@exgate.tek.com	x
13	Schmidmer	Christian	Opticom	cs@opticom.de	x
14	Dhondt	Yves	Ghent University	yves.dhondt@ugent.be	x
15	Corriveau	Phil	Intel	philip.j.corriveau@intel.com	x
16	Speranza	Filippo	CRC	Filippo.speranza@crc.ca	x
17	Teng	Andy	Qualcomm	ateng@qualcomm.com	x
18	Thorpe	Leigh	Nortel	thorpe@nortel.com	x
19	Rahrer	Tim	Nortel	tim.rahrer@nortel.com	x
20	Lee	Chulhee	Yonsei University	chulhee@yonsei.ac.kr	x
21	Watanabe	Keishiro	NTT	watanabe.keishiro@lab.ntt.co.jp	x
22	Okamoto	Jun	NTT	okamoto.jun@lab.ntt.co.jp	x
23	McCarthy	Sean	Motorola	SMcCarthy@modulusvideo.com	x
24	Yamada	Toru	NEC	t-yamada@ap.jp.nec.com	x
25	Sugimoto	Osamu	KDDI R&D	sugimoto@kddilabs.jp	x
26	Coverdale	Paul	ClaraSonic	Coverdale@sympatico.ca	x
27	Berger	Jens	Spirent	Jens.Berger@swissqual.com	x
28	Hands	David	BT	david.2.hands@bt.com	audio
29	Renaud	Ron	CRC	Ron.renaud@crc.ca	x
30	Meier	Rolf	Nortel	rmeier@nortel.com	x
31	Huynh-Thu	Quan	Psytechnics	Quan.huynh-thu@psytechnics.com	x
32	Bourret	Alex	BT	Alex.bourret@bt.com	audio
33	Cotanis	Irina	Ericsson	Irina.cotanis@ericsson.com	audio
34	Desjardins	Eric	CRC	Eric.desjardins@crc.ca	x

VQEG Meeting Minutes

Ottawa, September 10-14, 2007

Monday 10 September 2007

Monday (Thanks to Margaret Pinson of NTIA/ITS for taking notes.)
8:30 am, introductions

Summary of project status

ILG – Invoice matrix for MM fees sent to proponents. Proponents & ILG working on transmitting payment; all sent out and being worked upon, with very few paid. Have not heard whether FUB/Vittorio (NTT, Genista) and FT (Opticom, Swissqual) have sent invoices. CRC, Verizon, Acreeo, IRCCyN have sent invoices and are in the process of finalizing payment.

Progress made on MM common sequences. Completed and ready for distribution.

Request for ILG sent regarding RRNR-TV test, with only a few labs having resources available to support RRNR-TV (Verizon & CRC). We will have to see if sufficient resources exist to proceed.

Designs for ILG & Proponent MM tests should be discussed later.

RRNR-TV – Changes made to RRNR-TV test plan: SSCQE to ACR-HRR, ILG do minimum of work, current deadline February 6, 2008 (depending upon ILG availability), new test plan posted (version 2.0) posted on VQEG web site. Proponents currently: NTIA, NEC, & Yonsei (BT unknown as not present).

MM – There was an audio call between the previous & current meeting. Scene pools were finalized. Also, there was a vote for change to test plan regarding amount of PVS that one single proponent could generate for their own test (changed from 50% to 100%). Lots of activity has occurred since then generating PVSs for ILG and proponent tests. One purpose of this meeting is to redistribute MM PVSs. There has been discussion on reflector concerning data analysis & screening (or not) of subjects. The player for MM test has been updated & put on FTP site for testing & should comply with the test plan.

HDTV – Will discuss proposal of having proponents submit subjectively rated video test along with model, as a way to move forward despite lack of ILG support. Interest has been expressed in moving HDTV forward before MM completes. Potentially interested proponents: NTIA, NEC, Yonsei, NTT, KDDI, Opticom, Nortel.

Tools & Subjective Labs – NTIA has provided tool for checking calibration of PVSs for MM test, which was the fundamental issue in the past. Opticom has proved two tools (a model fitting tool and a temporal delay tool), but both are only for the ILG. One need that we have concerns discarding viewers whose performance differs from other viewers, and we need a tool that would perform such analysis. VQEG web site needs to be improved, so that the tools are identified in one place (e.g., download, web link).

Point of SRC for Scene & HRC Collection – Must redistribute all MM data as soon as possible at this meeting, preferably on one hard drive & copy. Expect 6 to 7 hours to copy all data from one master hard drive to another drive. Explored commercially available HD content, but this is very expensive (e.g., 30,000 Euros). SMPTE has HDTV material. NTIA can provide some HD material free for research purposes only. The question exists of what footage is appropriate and

how much are we willing to pay. Some of the KDDI and Swissqual material for MM was HDTV. The problem with HD source material is that it much is compressed or upconverted. Send Chulhee Lee of Yonsei information on any HDTV source that you can make available.

Hybrid – At the previous meeting, VQEG agreed on a tool (IPanalyzer) to produce a file that contains the necessary data. From there (Quicktime) we need analysis to produce bit-stream, which task turns out to be very difficult. Maybe we can have our own very simple stream server, but this needs to be discussed at the Hybrid sessions. IP analyzer discussion will be a tricky subject, but from there the test plan should be quick to develop (e.g., starting with another test plan). The goal is to analyze the validity of models that take bit-stream data (e.g., network statistics) and perceptual data (e.g., final image) and produce a quality prediction. G1050 recommendation mentioned during this verbal summary as being a potentially relevant tool – set top boxes appear to have much of the data VQEG wants internally available, and other standards committees are working on making this internal data available externally.

Calibration Verification for Standard Definition – includes spatial shift, temporal delay, gain & offset. Yonsei & NTIA are validating calibration by the end of September so we can get this approved. No subjective testing required. Chulhee has software to create impaired video clips. Thinking of using VQEG FR-TV Phase 1 material, since this was carefully calibrated by hand. MM test data could perhaps also be used. SG-9 meets October 29, 2007 in Colorado, USA.

MM Discussion on PVS Distribution

Acreo – not finished, will finish this week, need to be cut to 8-sec, also have problems with packet loss & calibration that need to be discussed.

CRC/Nortel – 12 Gb, not finished but will be done soon

FT – plan to upload to FTP site, this has not already been done

FUB – no

Genista – have put on FTP site, Marcus may have it available

IRCCyN – yes

KDDI – almost finished, but will finish this week 40 Gb

NTIA yes 51 Gb

NTT almost finished with PVS but some conditions' calibration not good especially conditions with packet loss. Approx 115 Gb.

Opticom yes 19 Gb

Psytechnics yes approx 25 Gb

SwissQual – no status reported

Yonsei yes approx 57 Gb

Only re-distribute 8-second sequences to limit time needed to copy data. If 8-sec sequences are not available, distribute 12-sec sequences and later enough information that every organization can edit identically to 8-sec sequences. Checksum recommended for checking that everyone has the same set of files, particularly when a PVS is replaced due to calibration issues.

Discussion ensued concerning editing (8-sec to 12-sec) either here at Ottawa meeting (discard first 2-sec), or back in labs after meeting (must re-transmit, but can be smarter about editing).

Action Item: All organizations should submit updated test designs (e.g., with final HRCs & updated clip names).

The test plan was modified with the changes from the audio call (e.g., drop of 50% HRC creation in a proponent's test not created by that proponent).

Common Set Sequences were displayed and discussed. Concerns were raised that perhaps a few clips should contain extremely high levels of packet loss. A vote ensued:

- keep Common Set as-is: Yonsei, Intel, KDDI, NTT, Opticom.
- include severe transmission error clip(s) in common set: Ericson, Nortel, Acreo.
- most organizations abstained.

Agreement was reached to leave common set as is (i.e., not add clip with severe transmission errors). This issue may be re-examined if clips can be made available before the meeting ends. (5 to 3 in favor of keeping the set as is).

Leigh Thorpe's presentation on data analysis ensued. See slides (EvaluationTesting_VQEG.ppt) on VQEG meeting ftp site. Discussions ensued on data aggregation.

Tuesday 11 September 2007

Review of Monday's minutes

Thanks for Quan Huynh-Thu of Psytechnics for taking the minutes.

Discussion about data analysis

Start of the discussion based on GC's email addressing the following concerns:

- Amount of data to analyze
- Labs/person(s) that will do the analysis

GC mentioned that due to the amount of data, the analysis cannot be conducted by one lab (like during FRTV2) but will have to be conducted by several labs

=> Need to identify who and which tool(s) will be used for the analysis. In the case that different tools are used, results could be slightly different.

GC has SAS package and Opticom has made a tool (Optimap) available to the ILG. They have been compared on a common data set and results are very close (correlations within tenth of percent).

MB said that 3 steps are required in the process of data analysis:

1. Gather all the subjective data (subjects ratings, MOS, standard deviation for each PVS) and corresponding output of objective models in one document (e.g. Excel Spreadsheet).
2. Computation of coefficients of the fitting function between subjective and objective scores (fitting per experiment)
3. Computation of the 3 performance metrics (correlation, RMSE and outlier ratio)

MB made the following proposal:

- A common format for step 1
- One common program for step 3 that takes as inputs the fitting coefficients, subjective data and objective data
- Several programs can be used to compute coefficients for step 2

=> Discussion about who will do what.

FS proposed that proponents do all the analysis and that the ILG uses Optimap (or other software) to check a subset of the data => what happens if results are not similar enough?

All the issues will be addressed by a data analysis working group: Opticom, CRC, IrCCyN, Verizon, SwissQual, Yonsei, Intel, Psytechnics (chair: Greg Cermak).

Discussion about post-experiment subject screening

GC proposed that the screening procedure should not be compulsory.

Decision: MM test plan stays as it is. If a lab has excluded and replaced subjects using the screening procedure, the number of excluded subjects must be reported.

Contributions are solicited to revise ITU-T P.910 and/or BT.500 to include the screening procedure and subjective test methodology used in the MM test plan.

Note: a problem has been found in the CommonSet sequences brought by NTIA/ITS. This will be fixed before the end of the meeting.

JB asked a few clarifications about some portions of the text in the MM test plan.

Currently the MM test plan has the following sentence “Coding Schemes that will be used may include, but are not limited to:” in Section 6.3.8. JB mentioned that now that all test designs are known, then this sentence could be replaced by the exact list of codecs included in all tests. MP mentioned that this information (which defines the final scope of the test) is typically included in the final report with the results.

P29: the sentence “These 25% must have a maximum temporal registration error of +3 seconds (added delay)” was modified to “These 25% must have a at most maximum temporal registration error of +3 seconds (added delay).”

Discussion about the ACREO player

KB said he updated the player and sent it to a small group of testers. He received feedback on several issues and bugs. The latest version is beta 4. Feedback was also sent on this version. Some remaining issues are:

1. How to start the experiment (currently the character ‘s’ has to be pressed to start the experiment, which implies the use of a keyboard).
2. Insert of a break during a session
3. Practice trials not included in current version.

Decision: The player will be run twice: once with a setup file to play 6 practice trials (which are not used in data analysis) and once with the setup file to play the actual PVS for the test. The same set of training sequences will be used in all tests.

Discussion on which sequences to use for the training trials: should we produce all new sequences or pick them from existing PVSs that have been created for the different tests. The current Test Plan specifies that the source content for the training trials should be different from the one used in the test.

Decision: 2/3 majority reached to modify the current test plan to allow training sequences to use the same content as the one used in the test

2 methods are proposed to select the training sequences:

Method 1: each of the 6 practice trials is picked up in a different test (that PVS has a content identified to be unique, i.e. the content of that PVS is not used in other experiments of the same resolution).

Method2: the training sequences are picked from the CommonSet

Vote: majority in favor of Method 1

Decision: Section 4.1.7 is modified to 'Practice clips: these test clips allow the viewer to familiarize with the assessment procedure and software. They must represent the range of distortions in the experiment. A number of 6 practice clips will be used. Each of the practice clips will come from a different test. Ratings given to practice clips are not used for data analysis.

QH will pick up the practice trials from the existing sets of PVSs that have been exchanged during the meeting.

Other remaining issues with the software:

1. PixelDepth: currently specified to be 16 but this creates some color artifacts. This parameter will be removed.
 2. NoOfImagesInSequence: currently needs to be specified but cannot be the same number for 25 and 30fps. This parameter will be removed.
 3. Grey background: currently set to 128 but might need to be changed to 108 to comply with ITU recommendation.
 4. MB mentioned he experienced a software crash when trying to run a full test
- ACREO will update the software to take into account all the mentioned issues.

Discussion about swap of PVS

Currently the MM test plan specifies that every proponent test lab runs a subjective experiment using the test material they have produced. Discussion is re-opened to swap test material such that proponent test lab runs a subjective experiment using PVSs prepared by another proponent test lab.

10 organizations in favor in changing the test plan.

0 organization opposed.

Decision: Each proponent test lab will run their subjective experiment using PVSs prepared by another proponent. Entire experiments are exchanged.

Before the end of the meeting, MP will produce a table with pairs of proponents.

Review of the MM schedule

See MM test plan v1.19

Wednesday 12 September 2007

(Thanks to Greg Cermak of Verizon for taking minutes.)

Tuesday minutes reviewed and approved.

The MM Data Analysis Group proposed that we would use OptiMap (or a MatLab equivalent) to analyze all the test data. Proponents will analyze their own data. If there is a discrepancy between the proponent results and the ILG analysis, then the proponent gives their mapping parameters to the ILG and ILG checks that their parameters do produce the rmse and correlation claimed. As per the testplan, any mapping tool used to produce mapping coefficients will be made available to ILG. A proponent can submit their own mapping coefficients, but must make the software available to the ILG. This proposal was accepted with some discussion about data formats. It was proposed (DH) that we report the raw model data as well as the mapped model data for each model. The raw model data are currently a requested input for OptiMap, and the fitted data are standard output So we automatically will have the data requested by DH. However, the raw data are due 14 Dec. and the fitted data are due 15 Jan., so there will have to be a separate editing step to put in the fitted model data.

Discussion of who will reformat the subjective data from the format of section 4.2.1 into the data format necessary for OptiMap or equivalent program. A group of proponents/ILG (Yonsei, Psytechnics, Opticom, and IRCCyN) are proposing to write software to convert from the section 4.2.1 format to the OptiMap format. ILG (CRC) will run this software and will output data in the format suitable for the OptiMap-type analysis program. These OptiMap-type data will be distributed to all proponents and ILG.

The last few dates in the MM Test Plan schedule were reviewed in the light of these new data-handling procedures.

Status of the PVS files was reviewed again – which files have been received and copied to a central storage site? The updated summary table was uploaded to the ftp site.

Ericsson proposed three PVSs, VGA, CIF, QCIF with low bit rates and high packet loss as possible bottom anchors in the MM Common Set. Proposal is to exchange these for existing Common Set PVSs with smaller transmission errors. Decision: These 3 PVSs are accepted; MP will incorporate them in the Common Set and distribute them to proponents & ILG.

ATIS-IIF sent a liaison regarding IPTV metrics.

HDTV

Co-chair (MP) reports that ILG will not be able to support HDTV currently because of resources going to MM and to RRNR work. MP proposes that HD testing be run in a new way in which the proponents do almost all of the work of assembling PVSs, and that ILG will mainly observe and officiate. (See proposal text – to be inserted here.) Major issues: Availability of source video. Need to keep equipment requirements minimal or generic.

Comments: How do we avoid proponents training their models on their own subject data? How do we avoid a proponent training on everyone else's data? Suggestion: Model should be

submitted before source video (SRC and PVS) is chosen so that proponent cannot influence the subjective data so that other models will fail. (Not clear to me how this would be accomplished.) However, the task of choosing the source material has been very labor-intensive: Who's going to do it? Question: How many proponent-based video source sets are necessary so that no proponent is familiar with too many sources? Suggestion: Move HDTV ahead of RRNR. Issue: Some proponents may not have enough facilities to do a complicated test.

Decision. Informal vote: Option 1. Run some kind of simplified HD Test to be finished in one year: Proponents submit a subjective data set and a model at the same time. Option 2. Approach like MM. Option 3. Pay ILG to run subjective tests. Discussion: If phase 1 only has coding errors, then should VQEG be recommending any quality metrics after Phase 1? Vote for option 1: NTIA, Acreo, VZ, SwissQual, Ericsson, NTT, Nortel, Opticom, Intel, Ghent. Option 2: Yonsei, NEC, KDDI. Option 3: Motorola. Winner is Option 1.

Decision (decision made later). Question of transmission errors in Phase 1 HDTV: Vote for having transmission errors in Phase 1. 7 for. 7 against.

Who can produce transmission errors for HD? Jens of SwissQual; Jun of NTT; Leigh of Nortel; Jorgen of Ericsson; Osamu of KDDI; Marcus of Opticom.

Image resolution and frame rate: How many proponents can handle subjective testing for 720p 59.94 fps? KDDI, Opticom, NTIA

720p50 Opticom, SwissQual

1080i29.97 Yonsei, KDDI, Opticom, NTT, NTIA, NEC

1080i 25 Opticom, SwissQual.

Therefore all four could be in the testing (and are currently in the Test Plan).

Issues to be considered before deciding on subjective test methodology:

- HRC types
- Transmission errors? Yes/no vote: No=5; Yes=12. Mild vs. severe errors: Opinion: Mild=1; Mild + severe = 10-ish. **Decision:** Transmission errors will be included. Note: Implies striking from Test Plan 1.4 the restrictions on freeze frames and frame skipping.
- Codecs: H.264 (AVC both high and main profile voted in; SVC voted in), MPEG-2 voted in, VC1 voted in, MJPEG-2000 voted out, DivX voted out.
- Explicit post-processing (not part of codec): de-blurring, de-blocking, noise filter. Vote against post-processing = 9; vote for = 1. **Decision:** No post-processing allowed.
- Scaling: PVS must be same scale, resolution, and format as original; HRC can include transformations such as 720p -->NTSC-->720p.
- If a progressive display is used and it needs de-interlacing, then this de-interlacing is performed offline so that the model receives the same signal as the viewer.
- Range of quality (average bit rate): 4 Mbs to 30 Mbs; 1 vote. 1.5 Mbs to 30; 7 votes; Mbs. 0.5 to 30 Mbs; 7 votes. **Decision.** Compromise of 1 Mbs as lower limit: voted in.

- Number of sequences tested per proponents: 160 (probably ACR). 60-70 (probably DSCQS). 30 (method unspecified).
- Length of test: 1.5 hour (with break; total viewing time of 50 minutes) vote=9. 1 hour (with break; total viewing time of 30 minutes) vote= 5. **Decision:** 50 minutes of viewing & voting time; total of 1.5 in-lab time.
- Duration of sequence: 30 sec to 1 min proposed (0 votes). Others propose 10 sec (11 votes). An issue is availability of long HD sequences. Motorola proposes 30 sec (3 votes). 16 sec proposed (1 vote). **Decision:** 10 sec.

Subjective testing methodology:

DSIS (ITU-T BT.500, 5-point impairment scale; approx 111 PVSs) 2 votes; 1 vote
 DSCQS (ITU-T BT.500, double stimulus continuous scale; approx 59 PVS) 3 votes; 2 votes
 SAMVIQ (New method using random access, multiple views, user control, costs under \$5k, and is available from FT; approx 50 PVS) 1 vote; 1 vote
 ACRHRR (ITU-RP.910, single stimulus 5-point scale; approx 200 PVS) 3 votes; 9 votes.

Decision: ACRHRR is clear choice.

ILG labs will be sent a list of the SRCs chosen by each proponent so the ILG can check that multiple proponents do not pick the same SRC. (Sent by email with thumbnail image.)

Models: FR, NR, RR, hybrid models will be accepted. Bandwidths for RR: **56 kbs, 128 kbs, 256 kbs.**

SRC submission date: For donation at next meeting.

SRC potential donors: NTIA, Intel, Nortel, Ghent, Comcast.

Camera specifications: Proposal is that native resolution of camera be 1920 x 1080. Much discussion. Must be careful about naming specific brands and models of cameras. Next meeting the experts will view examples of SRCs from different cameras in order to certify whether cameras are high enough quality. Others say that eyeball is not good enough; only specs should be able to support a camera specification for the Test Plan. Otherwise, no decision about camera specs.

Discussion of buying video clips rather than buying a high-end HD camera: Would proponents be willing to purchase SRCs that other proponents have chosen? Or, is a proponent allowed to submit purchased source that then other proponents must buy (or send an executable of their model to the first proponent to run against the purchased video)?
 Vote: Yes= NTT, Yonsei, NTIA; No=SwissQual; Depends on fee; cannot give away their executable = KDDI, Opticom, NEC.

Note: An ILG lab will need to be available to validate models for proponents who cannot let out their models to other proponents (KDDI, Opticom, NEC).

Tentative model submission date: Sept. 2008.

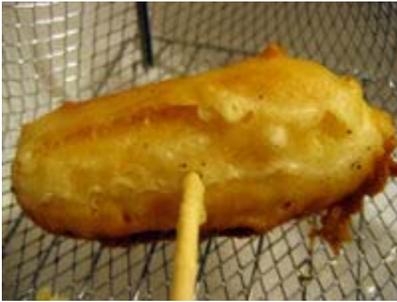
Chair (AW) suggests that we consider applying metrics from J.144R to HD source as an interim action.

Thursday 13 September 2007

VQEG Meeting, Ottawa, Ontario, Canada
Thursday, September 13, 2007
Thanks to Carolyn Ford of NTIA/ITS for taking notes.

First order of business:

- Carolyn will be taking notes today, on the following conditions:
 - o the membership buys her beers at dinner tonight
 - o She will not take notes at the next meeting, someone else can take a turn



This is a deep-fried twinkie.

The weather is lovely... the first nip of autumn is in the air, and the leaves are just starting to change color [color comment submitted by Yves]

Arthur is grumpy.

HD discussion resumes

Chulhee (Yonsei) proposes a scheme for model sharing to handle the issue of sharing source, it was accepted into the plan. See the ftp site (or the plan).

Carolyn (NTIA) will donate about one day of professional shooting if given a list of potential scenes desired.

Camera/source specs: discuss over the reflector for the next few weeks. Rolf Meier (Nortel), Phil Corriveau (Intel), Margaret Pinson (NTIA), Chulhee Lee (Yonsei), Ron Renaud (CRC), Yves Dhondt (Ghent) will lead the push.

Minimum number of data sets (experiments?) for HD test: three independent sources of the data sets generating at least 4 data sets.

Misc Topic

Arthur presents a request from SG9 to run the models listed in J.144R on clipped HD source and compare to a subjective test of same. This needs to be completed by about April. No one volunteers to do this (IRCCyN, CRC, Verizon will help).

RRNR Discussion

Proponents: NTIA, NEC, Tektronix, Yonsei, BT

Change to test plan: minimum 12 sources, no max, ILG makes the decision.

No more than 75% PVS can be from created by a single proponent.

Oh no, the discussion is veering into statistical analysis. Greg will provide input later.

Adjustments to schedule. See revised plan.

Chulhee proposes a Test Conditions diagram. See ftp site.

Greg returns with decision: straight averaging of MOS from different labs running the same PVSs without normalization of viewers from different labs. See him for explanation.

Lunch

Group Photo

Hybrid Model discussion

Contributions (4) from NTT (see files on ftp server) to use the P.NAMS mode A and B in the hybrid project. No objection to NTT proposals.

Contribution from Telchemy. Propose to include encrypted content. None opposed.

Review of liaison from ATIS regarding a list of parameters to consider to use in QoS measurements. Reply liaison will be written Friday.

Discussion regarding which level of stream will be available for analysis.

Discussion regarding how many decoders will be used for the test (error concealment vs. none).

Proposed possible proponents: Psytechnics, KDDI, BT, NTT, NEC, Yonsei, SQ, Opticom, Telchemy, Tektronix, NTIA, FT, DT, Witbe, QualiDeo, Ghent, Toyama

Interested core group to help Chulhee: Bruce (Telchemy), Carolyn (NTIA), Yves (Ghent)

Discussion about ways to recover if proponent's model can't read the bitstream.

Limit the types of transport streams to 3GPP transport file format and MPEG2 TS.

Formats: Two models: (QVGA and VGA) , (SD and HD)

Codecs: See doc (070913 hybrid models.ppt) on Ottawa meeting ftp site.

End of day

Aaaahhh... can't wait to get my free beers! (She got em).

Friday 14 September 2007

Notetaker: Christian Schmidmer, OPTICOM

Meeting starts at 8:18 (instead of 8:00, which is exceptionally good!)

- Short discussion on email from Stefan Winkler regarding the length of his sequences. It was decided that Quan from Psytechnics shall reply that all the sequences must be fixed since some models might otherwise be confused.

Review of yesterday's minutes

- Short discussion on Carolyn's comments. Disagreement on her comment re. Arthur's mood
- Meetings accepted after several changes

Update of status of PVS creation

- ➔ See PVS_Status.xls
- ➔ Quan reminds proponents to update their test design description with information on which filename correlates with which HRC

Discussion on Hybrid Models (8:51, continued)

- ➔ See also Chulhee's powerpoint slides!

- Short summary by Chulhee of yesterdays' decisions.
- Discussion on the scope:
 - Codecs: H.264, MPEG2, MPEG4-Part2
 - Transport: H.264-RTP, H264-MPEG2TS, MPEG2-TS..... (something is missing here, see Chulhee's notes)
- Chulhee presents proposal from IPTV regarding monitoring points. Agreement in the group that our focus is on measurements at the end of the transmission chain.
- Reference codecs to ensure compliance with the standard for error free bitstreams are:
 - H.264: Public JM model. We must fix a version number to be used at a later stage since the model is often updated.
 - MPEG2: ffmpeg
 - MPEG4: ??????
- Discussion on whether we shall limit us to using only one decoder for each codec type or multiple codecs. If multiple decoders should be used which produce sequences of very different quality, the bistream models have no chance to predict both correctly without knowing which decoder was used. The drawback is that we can not prove that the models would work with other decoders....

Decision: *Only one decoder per standard shall be used to generate the PVS for the bitstream only models.*

- Do we need / have a reference bitstream server? VLC (open source, www.videolan.org) was proposed by several attendees. Helix (Real Networks, limited version is available as open source) and Darwin (Apple, open source) servers were proposed too. Both could also generate 3GPP streams. Live555 (open source) was also proposed for H.264 over RTP and MPEG2TS.

Decision: *If possible, only one reference streaming server shall be used for all experiments. This server shall be selected before Dec. 31st, 2007.*

- Capture format for IP traces:

Decision: *PCAP as generated by e.g. Wireshark or tcpdump.*

→ Discussion on schedule:

- IP trace analyzer to generate text files for hybrid models:
Decision: *Proposals required until next Meeting (est. Feb. 2008).*
- Finalization of IP Analyzer:
Target: *finished 3 months after next Meeting.*
- Finalisation of the testplan:
Target: *Should be almost finished after the next meeting.*
- Model submission:
Target: *6 months after model submission.*
- Submission of parsing programs:
Target: *3 months after model submission.*

→ OPTICOM volunteers to edit the schedule in MS Project under the condition that the MS Project version shall be the reference for future changes.

→ Reminder: hybrid@its.blrdoc.gov is the hybrid reflector

Other Business

Unsolicited Advertising over VQEG Reflectors

11:11 Coffee Break

11:40 Meeting Resumed

Discussion on unsolicited advertising continues.

A text was agreed and will be posted on the VQEG web site.

This the agreed upon text:

1. Although most parties participating in VQEG have commercial interests, VQEG shall remain a commerce free space.
2. Companies and individuals which abuse any of the VQEG mailing lists for unsolicited advertising will be deleted from all lists.
3. Unsolicited advertising, dispatching of brochures, public product presentations etc. are not allowed at VQEG meetings. Excepted from this are of course private discussions between participants, excursions to/through the meeting hosts facilities and invited presentations.
4. Questions to the reflector regarding products are okay. The responses should not be sent to the whole reflector but to the requesting party only.
5. Information regarding conferences/journals related to VQEG work are okay to post to the reflector.

Project for collaborative model development

BT and IRCCyN are interested in a collaborative approach to develop modules that measure different aspects of quality. These modules could then be combined to form a kind of “super model”, which is potentially the most accurate model since it contains the combined knowledge of all experts. The two proposing parties would like to start such an initiative between now and the next meeting. ->See also the email from the proposing parties.

Comments:

ITS is only able to participate if the group is completely open and under the umbrella of an international organisation like eg. the ITU JRGMMQA. If VQEG itself fulfils this criterion, is uncertain.

It was mentioned that issues regarding the IPRs may exist and legal issues should be cleared before the development can really start.

A short poll showed that most attendees are interested in such a work.

Decision: *A new group or forum shall be founded to follow that collaborative development approach. proposals on how the legal issues can be overcome are requested for the next meeting.*

Meeting Procedure

Decision: *Meetings shall start at sharp 8:30 in the future.*

New Co-Chair for Subjective Lab and Tools Group

Decision: *Yves Dhont will be a cochair instead of David Bottoms.*

Meeting Schedule

NTT and University of Gent can certainly host the spring meeting only. Later dates are uncertain. It was mentioned that due to the workload three meetings might be required in 2008. The MM statistics shall be discussed in February 2008 according to the testplan. This is proposed as a meeting date.

Proposals for the next two meetings:

1. Between Feb. 11 and March 19, Blackout Feb. 18
2. June 30 -July 14

Most likely the spring meeting will be in Kyoto and the summer meeting will be in Belgium.

Conference Calls: proposals shall be made within a month.

Discussion on Monitor for HD

Problems with consumer LCDs were reported, that could potentially affect the picture quality, especially when they are driven by a PC. This is also valid for digital inputs.

Tentative decision: *Professional CRT and LCD monitors or high quality consumer TVs displaying the full native resolution can be used for the test.*

12:58 Meeting adjourned, everybody is invited to join Arthur for writing Liason statements after 1:00pm