VQEG MEETING SEPT. 2008

Monday

Opening reports from groups. I gave a brief report of ILG activities. Chulhee reports that RRNR is making good progress – may have data by Oct. and a possible final report by Jan. Arthur is thinking about maybe having a meeting in Jan. The MM Final Report has been submitted to SG9 of ITU. I talked about the HD Test Plan and about the statistical work I did on correlation vs RMSE for the MM project. Riccardo (of Orange) again offered the use of SAMVIQ. [Note: This is a very well-attended meeting, mainly by the Europeans.] KWILL has a couple of representatives. Also, supposedly the tool that Alan Clark of Telchemy offered us may be put up for everyone on the VQEG website. Some discussion of the problem of getting open-source video test sequences. Hybrid Project is waiting on getting an “IP Analyzer.” Patrick LeCallet mentioned a prospective project to put together a collaborative best-practices joint model for video (as had been done for voice, leading to PESQ).

Liaison from ITU-R WP6C: There has historically been trouble with WP6. 6C has been re-constituted to have a Rapporteur with VQEG. They will submit the MM report to ITU-R through the new rapporteur, Alina. Vittorio will help liase with VQEG. ITU-T and ITU-R do not play well together. WP6C is calling for proposals for objective measurement of surround sound (5.1 audio).

Liaison from ITU-T SG12: Working on an MM standard P.NAMS, possibly in collaboration with VQEG. It may be closely related to the Hybrid project. They seem to be interested in developing subjective test plans – which seems to be not such a big deal to me. Maybe it’s the management of multiple tests and secrecy that’s difficult.

Liaison from SG9: MM reports have been approved by ITU-T, i.e., the 2 recommendations (J.246 & J.247) coming from VQEG on MM have been approved (Aug. 12). The Final Report was submitted to SG9 (last week), and is scheduled to be submitted to WP6. Meanwhile, ITU-T is talking about restructuring itself at a meeting in Oct/Nov.

Liaison from ATIS IIF: Proposing the fast-track model validation/certification process I’ve already heard about. The big question is how an institution can be constituted to do the certification. Another question is how to keep out too many models.

MONDAY AFTER LUNCH—lunch cost 10 Euro

At lunch there was news that Carolyn Ford is recovering from a stroke.

More discussion about the ATIS IIF proposal. Vittorio is shooting new video that might be used for potential IIF work. Vittorio raises the issue of cost of operating an IIF program. Opticom raises the issue of how often a standard or certification can be updated and still remain useful as a standard. It’s an issue of barriers to entry of new algorithms. The idea is that a proponent would submit an executable model and would never see the video that’s being used for testing. Vittorio reports that FUB was offered a fee to sell the VQEG MM MOS data. Someone suggests that a dual-key security system be set up for VQEG data.

Topic now is supplementary analyses of the MM data and the more general question about ownership, usage, and copyright of VQEG data. Vittorio objects to providing individual raw data to the outside world, as opposed to MOS per PVS. Right now, VQEG is offering to have MOS be available in Sept. 2009. Decision: MOS and DMOS will become available on the VQEG website, but individual raw data will be available only through the individual labs.

Regarding the Supplementary Analysis and the NTIA Special Analysis: Should there be a VQEG Supplementary Report. Question: Do proponents want to release their objective data from MM? Do proponents want to be anonymous in any subsequent reports & data analyses? Much discussion. Discussion will continue tomorrow when the last of the proponents (SwissQual) arrives at the meeting.

Discussion of MM Phase II: Should there be audio in a Phase II? Vittorio suggests using task performance as responses. Arthur raises possibility of using audio in the Hybrid project. !9 people indicated that they want to participate in a Phase II. Volunteers for co-chairs of MM2? Kjell did not refuse to volunteer. Opticom wants to make a couple of small changes to their comments section in the MM Final Report.

TUESDAY MORNING MEETING

K-Will is a proponent in RRNR, along with NTIA, NEC, and someone else.

Vittorio volunteered to be the second co-chair of MM Phase II.

Discussion of whether Opticom can revise their section of the MM Report – Proponents’ Comments. Decision postponed until the afternoon.

SG12, Q14, liaison from Vincent Barriac (FT) regarding the relevance of the MM work to SG12 and P.NAMS. I’m having a hard time understanding what P.NAMS is and how it relates to Hybrid, but the idea is that they are logically related. SG12 is requesting collaboration with VQEG. The domain of P.NAMS is an IP stream; the output would be some sort of quality metric?? They are looking for proponents in a collaborative rather than competitive project. The issue, as with Hybrid, is the need for tools for analyzing IP streams that come from different codecs; also for better packet-loss simulators?? P.NAMS is not yet clear about its scope. P.NAMS sounds more theoretical than actual. A key name is Bruce Adams. P.NAMS is aimed only at packet headers, not at payloads. Interesting point: Proponents would like to know something about the SRC, such as complexity and motion, if they do not see the payload. Bottom line: Subjective test data from Hybrid could be used in the validation of P.NAMS.

There’s a side conversation running along, which is about subjective testing methods. The current revised P.910 admits ACRHR. It would be nice to be on top of the Perception & Psychophysics literature regarding experimental methods for audio + video.

AFTER LUNCH TUESDAY

Had a tour of IBBT demos. They develop apps and some hardware for multimedia use in many areas. They are funded by the government of Flanders, and are associated with the University of Gent. The lab has over 100 people, and is much more impressive than anything in VZ N&T.

HYBRID PROJECT. Co-chairs Chulhee Lee and Jens Berger (SwissQual). Starting with requirements for the PVSs and SRCs in terms of length, alignment, other technical issues. Then arguments about subjective judgment task and scale. Voted for an 11-point ACR HRR scale.

Preferred input type: Options are text vs. bit-stream. [Maybe send my MM analysis to the guy from Lancaster and the guy from AGH.

Method for capturing bit-stream data: Several kinds of apps available.

Server for streaming video: Several possible choices.

IP analyzer converter & decoder. Acreo has a tool. Telchemy may have something. “TCP Replay” is a packet-stream player. No real resolution to this issue.

Question: Model applied to packet loss vs. bit error? Ericsson suggests PLR 0-15% for QCIF/VGA in a mobile environment. PLR for SD/HD said to be available from DSL Forum (now Broadband Forum).

Schedule: Hope for a proposal for a working system by Jan. 2009; Test Plan finalized July 2009. Model submission 6 months after Test Plan finalized.

[Returning to the MM Final Report: Proposal for all proponents to revise their comments section to delete reference to competitors’ models. Proponent opinion is split. Resolution: Arthur is not going to change the Final Report.]

WEDNESDAY MORNING

Notes from yesterday that I missed: 1. Marie-Neige Garcia will chair a group on subjective scales. [Maybe I should try to become involved.] 2. Acreo will donate open-source code for handling IP video streams.

RR/NR

Proponenets = NEC, NTIA, Yonsei, K-WILL. ILG = CRC.

Ran into problems in creating PVSs.

Using only 16 subjects per test.

Data collection and analysis done by end of Oct.

Final Report by end of November.

K-WILL is having trouble doing their subjective test, hiring NHK for more money than they want to spend. Maybe FUB will run the test. Maybe FT will run the test.

HD TEST PLAN

[Most of the notes will be in the HD Test Plan itself.]

Discussion of 720p vs. 1080i: Both kinds of displays are in the market, but the idea is that anything broadcast at 720p would be rescaled by home screens to 1080i. Should source be 1080p or 720p? Supposedly in Kyoto it was decided that source should be 1080p. Referring to the Kyoto meeting notes. Also, much discussion. 1080p that had been upscaled from 720p can be taken as source? Answer: Yes. Then, is codec applied to 720p or to 1080p? Answer: to 1080p.

WEDNESDAY AFTERNOON

[Note from lunch: Mu Mu of Lancaster U. should be introduced to Amy Reibman of ATT.]

Proponents: NEC, SwissQual, Psytehnics, NTT, Opticom, Yonsei, NTT, K-WILL.

ILG: Verizon, FUB, Orange, IRCCyN, IBBT, Acreo, Intel, DT

Large discussion about what analysis can be done on the test data outside of the Test Plan. Universities want access to all data for publication. Potential user organizations such as FT/Orange also want access to data for extra analyses. The proponent companies want to restrict all objective (model) data. Impasse? No agreement. Requires ILG-proponent negotiation offline.

Regarding objective models: What counts as a different model? Are 1080i and 1080p separate models? Should each model be required to handle all of the combinations of 1080i and p at 30/25 and 60/50 fps. Question is turned into what the labs are capable of testing. (I’m not quite sure what the AJA/Xena does.)

A related issue is how much spatial and temporal calibration must be done by the model, and how much needs to be done manually in preparation of test videos? If the limits are loose, then the models may have to spend 2 hours per video clip just doing the registration calibration (says Chulhee). The proponents vote for removing the manual inputs. People raise strange PVS scenarios in which, e.g., delay varies dynamically within a video clip. Need to go to section 8 – calibration limits – and agree there before returning to section 3.2 about whether models should receive manual calibration data about PVSs as data input. Section 3.2 voted down 9-2.