Official Video Quality Experts Group Minutes
Boston, April 24-28 2006
Hosted at Verizon Labs

Philip Corriveau Co-Chair VQEG - Intel Corporation
Arthur Webster Co-Chair VQEG – NTIA/ITS

Participants List

1	Balasubrawmanian	Vivaik	INTEL
2	Adams	Bruce	Telchemy
3	Bourret	Alex	BT
4	Brunnström	Kjell	Acreo
5	Cermak	Greg	Verizon
6	Cotanis	Irina	Ericsson
7	Cooppan	Dayan	Verizon
8	Corriveau	Philip	INTEL
9	Dhondt	Yves	Ghent Univ.
10	Douglass	Jack	Spirent
11	Fenimore	Charles	NIST
12	Fernald	Royce	INTEL
13	Ford	Carolyn	NTIA/ITS
14	Hands	Dave	BT
15	Huynh-Thu	Quan	Psytechnics
16	Kurita	Takaaki	NTT
17	Kush	Patricia	Verizon
18	Layman	David	INTEL
19	Le Callet	Patrick	IRCCyN
20	Lee	Chulhee	Yonsei Univ.
21	Okamoto	Jun	NTT
22	Pinson	Margaret	NTIA/ITS
23	Rahrer	Tim	Nortel
24	Reckwerdt	Bill	Video Clarity
25	Rodel	Eugen	SwissQual
26	Schmidmer	Christian	Opticom
27	Speranza	Filippo	CRC
28	Sugimoto	Osamu	KDDI R&D
29	Webster	Arthur	NTIA/ITS
30	Widmer	Rene	SwissQual
31	Yamada	Toru	NEC
32	Zarrabizadeh	Mohamamd	Lucent

www.vqeg.org

(Note: The ITU-T JRG-MMQA will meet coincident with the Multimedia Sessions of the VQEG Meeting)

Agenda

Monday, April 24

8:00	Start / Designate Note Taker (Webster/Corriveau) Introductions Meeting Logistics Updates (Maximum 15 minutes - each Group) Independent Lab Group (ILG)(Corriveau/ Speranza) RRNR-TV (Bourret/ Lee) Multimedia (Hands/ Brunnstrom) HDTV (Wolf/ Balasubrawmanian) Tools and Subjective Labs Setup Group (TBD) POC for Source and HRC Sequence collection (Lee)
10:00	Begin MM TestPlan Discussions

- final review and clean-up of the testplan
- review new SRC / HRC content
- review test condition matrix (covers conditions we will use in the testing, we need to specify codecs to be used, bit-rates, frame rates etc)
- fees (once we agree on the matrix we will have a good idea of how many subjective tests are required, this should allow the ILG to be more precise on fees)
- player/software packages (review alternative packages and agree on what to use)
- data fitting method (logistic or cubic polynomial monotonic)
- threshold for rejecting subjects
- review schedule

11:30-1:00	Lunch
1:00	MM TestPlan Discussions Source/PVS Material Review
5:00	End of Day

Dinner arranged for Downtown Waltham (not free)

Tuesday, April 25

8:00	Start / Designate Note Taker Review Monday's Decisions (Webster/Corriveau) MM TestPlan Discussions Source/PVS Material Review
11:30-1:00	Lunch
12:00-1:00	Verizon Lab Tour
5:00	End of Day

Dinner expedition to Downtown Waltham (not free)

Wednesday, April 26

8:00	Start / Designate Note Taker Review Tuesday's Decisions (Webster/Corriveau) MM TestPlan Discussions Source/PVS Material Review Final Review of Decisions
5:00	End of Day

Thursday, April 27

8:00	Start / Designate Note Taker Review Wednesday's Decisions (Webster/Corriveau)
8:15 – 9:15	Presentation: Gross Error Detector - Royce Fernald (INTEL)
9:15 – 10:00	Presentation: Network impairment generator for audio and video - Jack Douglas of Spirent
10:00-10:15	Telchemy Document
10:15- 11:30	RRNR TV (TBD) Source and PVS availability IPTV industry needs
11:30-12:30	Lunch
12:30 – 1:30	HDTV - Vivaik B. Source and PVS availability Current testing Patrick Le Callet Presentation on Monitor study on Motion Blur
1:30 – 5:00	MM Resumes (Hands/Brunnstrom)
5:00	End of Day

Friday April 28

8:00	Start / Designate Note Taker Review Thursday's Decisions (Webster/Corriveau) Final Review of MM test video clips (if needed)
10:00	Review and Documentation of VQEG Decisions Review action items
11:00	Other Business Next Meeting
11:30-1:00	Lunch

1:00 Close VQEG meeting

Decisions from the Meeting

Throughout the notes below – all decisions from the meeting are in bold. Please read carefully through each section to find the bolded areas.

Daily Minutes:

Minutes for Monday April 24th 2006 - Quan

Opening of the meeting (by AW), attendees' self-introduction, agenda and meeting logistics (by GC)

Updates of ad-hoc groups:

- ILG (by PS):
  - Fees – no consensus yet. PS said there's no list of ILG labs. AW said he would have a list.
  - Video material for RRNR-TV: still problematic to find suitable material. FRTVII material not available anymore for copyright reason.
- RRNR-TV: no real progress, still no material available. AB reminded that decision was made at the last meeting to allow original HDTV material to be down-sampled to SDTV.
- MM (by DH):
  - DH reviewed the last MM activities: audio-calls that lead to the completion of the MM test plan (status: approved) but some annexes need to be completed
  - Remaining issues to be discussed during the Boston meeting:
    - Data fitting
    - Schedule of the test plan needs review
    - Selection of video playback systems for running subjective tests
    - Method/threshold for post-experiment screening of subjects votes
    - Clear matrix of HRCs
- HDTV (by VB):
  - Little progress since last meeting
  - Most work done towards gathering HDTV material
- Tools and lab set-up group:
  - No update
  - VB named new co-chair with SW
- POC for SRC (by CL):
  - VQEG Phase1 – open for research
  - SwissQual – open for research
  - NTIA – open for research
  - CRC – open for research
  - KBS – needs NDA
  - Psytechnics – needs NDA
  - NTT – needs NDA
  - FT - ?

- Yonsei - ?

MM discussion (current version of test plan v.1.12)

Review of test plan annexes:

- Annex 1: completed although DH has some text that needs to be incorporated
- Annex 2: completed
- Annex 3: text from JG that was discussed during MM audio call. Text is being updated.
- Annex 4: in progress. Fee structure and mechanism to be included in this Annex. Need for decision on whether or not to allow a proponent to contribute something else than a subjective test
- Annex 5: in progress - will be discussed later in the meeting
- Annex 6:
  - vote on Method1 vs. Method2
    - 6 votes for Method 1: Psytechnics, NTIA, Verizon, NIST, KDDI, Intel
    - 3 votes for Method 2: BT, NEC, Acreo

    DECISION: Method 1 will be applied
  - Rejection thresholds for Method 1:
    - Proposal to change r1 to 0.75 and r2 to 0.8

    DECISION: r1=0.75 and r2=0.8
- Annex 7: completed
- Annex 8: completed

Matrix of HRC:
AW presented a proposal for a number of combinations of SRC x codec x bit rate x frame rate x errors (4 codecs x 5 bit rates x 4 frame rates x 5 additional error types)

CL showed some examples of VGA SRCs that obtained a MOS in the fair-good category in his subjective tests. He also showed some example of PVS.

The group reviewed SRCs that some participants brought to the meeting. Amongst all the SRCs presented, the following ones were rejected:

- NTT:
  - count_* at all resolutions rejected because of un-natural content.
  - Cropping of remaining VGA sources need to be checked and re-reviewed
- FT: none rejected
- SwissQual: none rejected
- Yonsei: ILG will review the 17-min footage
- NTIA:
  - NTIA_SRC_flwrfade
  - NTIA_SRC_hcuff
  - NTIA_SRC_pathsong

End of the day

ACTION: DH to send out an email on the reflector with the following statement: every proponent is required to make a contribution to the MM work in the form of a

subjective test. If a proponent is unable to run a subjective test, they must put in a comparable contribution and declare before fee payment their proposal for a comparable contribution. This will be reviewed by the group.

SUMMARY OF MONDAY'S DECISIONS

- ANNEX 6: Method 1 used for post-experimental screening
- ANNEX 6: rejection thresholds will be r1=0.75 and r2=0.80

Minutes for Tuesday April 25th 2006 - Alex

Review of Monday's minutes, by Alex Bourret of BT.
No comments on the minutes. Minutes were approved.

Review of the SRCs.
SRCs from Opticom :
3 sets of sequences, restricted use for vqeg. D1 -> Digibeta
discovery1_cif (6 minutes 54s) Acceptable exept for some politically charged sections of it. Also, corporate name appears from time to time (Lufthansa).
discovery_part1 - 6 are the original files.
fhg-driving-force. Advertising for Fraunhofer. Accepted.
fhg-driving-force_3425f_1b.
fhg-driving-force_3425f_2a
fhg-driving-force_3425f_2b contains medical images.
fhg-driving-force_3425f_3a
fhg-driving-force_3425f_3b
DH. Subjective test should be targeted to specific applications ? (conferencing, medical...)
ifa_komplett.avi Content taken with a high speed camera. Original content in HD format.
*Christian would like to have a similar NDA.

First pass -> pool. Second screening by the ILG.

Content from Psytechnics:
Made from sequences bought from a company. License agreement has to be signed to used them (similar to KBS agreement). Most shot done in 720p.
PSY_SRC_ccski01_vga
PSY_SRC_ccski02_vga camera motion
PSY_SRC_fest01
PSY_SRC_fest02
PSY_SRC_footb01
PSY_SRC_skidh01
PSY_SRC_skidh02
PSY_SRC_skidh03
in CIF and QCIF only :
PSY_SRC_drink01
PSY_SRC_inter01
PSY_SRC_movie01
PSY_SRC_movie02

Issue of deleting files, as specified in the Psytechnics' agreement. Means the sequences can't be used to improve the models later on. There is a time limit (end of current study period).

Proponants opposed to used the psytechnics content on the basis of the license agreement : NTIA
No ILG is opposed.
4 are in favour of puting these sequences in the pool.
6 don't have a preference.
Decision: Will use Psytechnics for the Pool.
Yonsei SRCs :
KBS_wanggun Original in digibeta, 29,97fps
KBS_soccer
KBS_news
KBS_musiccamp
KBS_morning
KBS_leepark
KBS_gayo

KDDI source material: upconverted from HD to D1 using hardware converter
Total duration : 18 minutes 5 seconds
Summerday (D2 format) 24 minutes 42s in total. (6G)
KDDI_D1 1 minute 39s total.
56G in total.
License term will discuss with KDDI about using the template. The whole content can't be distributed, KDDI will do editing first.
It is possible to provide the whole sequence to use of ILG only.

Arthur : Content provider have to fill the template with their information, send it to the reflector so people can sign them.

Will split in 3 groups : 2 for viewing contents, 1 group for populating the matrix.
Proposal : need to produce a distribution with dates and points of contacts for each part of world (3 regions).

Agreement : The points of contact are : Filippo & Phil for North America, Kjell for Europe and Chulhee for Asia.
Other solution : use of bittorent ?

Agreement was reached on the distribution procedure : test materials sent to ILG, who makes the selection, sends the selected materials to Chulhee, then Chulhee distributes to all regional points of contact.

Quan : Matrix
Grouping tests by aspect to be tested (for example: 2 tests for compression error, 2 for transmission errors, 1 on material, 1 on network condition and 1 on large variety of impairments = 7 tests per resolution)
Advantage : easier in test design. Each lab can specialise on an aspect of the test. Overall size still manageable.
Number of SRCs per HRCs too small ? Should be at least doubled ?
Range of quality on tests such as transmission errors are likely to be poor.
Improvement : the different tests could be shuffled across labs so that viewers have a more varied set of sequences / hrcs to assess.

Presentation from NTT:
On the issue of independently created PVSs by a number of organisations. In 3 steps : agreement on the type of degrations, agreement on the degrees on degration, creation of the PVSs.
Degrations : at least 4 types of degrations : spatial coding distortion, temporal, spatial packet loss, temporal packet loss.
By using the worst and best quality in every test, we obtain anchors. Worst quality becomes bottom anchor, reformated source the top anchor.

Agreement: We will use anchor-like PVS examples as guidelines for proponents and ILG in the production of PVSs as discussed in the NTT contribution. These anchor-like PVSs will not be included in the subjective tests.

Pool of PVSs can be created by populating a matrix of conditions(initial proposal), or by creating tests for each condition (Quan and Margaret's proposal).
In favor of template approach : Verizon, CRC, Ericson, NIST, Opticom, Psytechnics, SwissQual, NTIA, NEC, NTT, KDDI, Intel, BT.
Matrix approach : no-one

Agreement was reached to use the template approach to creation of the PVSs. (Details to be provided Wednesday by MP and QH.

The ILG can decide to remove test results when the MOS score of a SRC is lower than 4.0. Perceived by ILG as too much responsibility, because could influence the final result. Also would result in loosing a whole dataset. (Test plan, page 17 point 20, on the schedule). If decision is taken before data are submitted, at least remove pressure on the ILG

Agreement was reach to specify that the decision to remove SRCs will be taken before the objective test results are submitted by the proponents to the ILG.

Minutes for Wednesday April 26th 2006 – Margaret P.

Vote on design of subjective tests. In favor: Opticom, Ericsson, Verizon, NTT, NEC, SwissQual, Lucent, KDDI, NTIA, Intel, IRCCyN; Opposed: none, No Preference: 3

Agreement was reached to start with Quan & Margaret's design proposal, and develop an appendix that provides guidelines to proponents for designing subjective tests.

Draft text reads as follows:

"Guidelines to labs designing subjective tests in step #8, section 5.3,

- Design for compression only: HRCs must not include transmission errors. Maximize a range of bit-rates and frame-rates and pre-processing and post-processing. A wide variety of codecs should be included (e.g., 4).

- Design for transmission errors: Most HRCs should include transmission errors (simulated & live transmission errors, packet loss, freeze frames,

etc.) Maximize a range of visually different transmission errors. Transmission errors must be paired with more than one bit-rate, codec, and frame-rate. Low, medium and high levels of transmission errors must be represented. Live network errors should be included. Preferably, a minimum of 30% HRCs should contain live network errors.

- Design for both compression & transmission errors: An approximately balanced mixture of HRCs with and without transmission errors will be presented. These tests will include transmission errors at many bit-rates.

Each experiment will include one secret HRC, created by the ILG. Each experiment will include a minimum of one secret SRC. Each subjective test should span the entire quality range, as demonstrated by the anchor clips."

Survey of the number of tests each proponent is willing to run

BT – 2 tests per resolution (2 QCIF, 2 CIF, & 2 VGA) → 6 tests
NTT – up to 4 tests per size (3CIF & 3VGA) → up to 12 tests → 6 test limit current
KDDI – up to 6 tests (2 QCIF, 2 CIF, & 2 VGA) → up to 6 tests
SwissQual – 1 test, QCIF or CIF → 1 test
Psytechnics – (1QCIF, 1CIF, 1 VGA) → 3 tests
NTIA – 1 test, VGA preferably → 1 test
Opticom – 1 test, CIF or QCIF → 1 test
Yonsei – 2 tests per resolution (2 QCIF, 2 CIF, & 2 VGA) → 6 tests
Lucent – potential proponent, 1 QCIF test → 1 test

Number of Tests per Resolution, if Opticom & Swiss & NTIA test QCIF:
QCIF 10
CIF 10
VGA 10

An idea was discussed that the ILG could retain some secret content entirely secret, in case future testing was required. This idea will be discussed later.

Agreement was reached that the ILG will perform twelve paid tests: 4 QCIF, 4 CIF & 4 VGA, provided that approximately 10 proponents participate. If a significant number of proponents drop out, then this agreement must be readdressed.

The price of these tests is expected to be approximately $60,000 divided by the number of proponents. Division of these fees among proponents & ILG will be determined later.

Action Item: The following content Providers needing a signed agreement form either must upload their agreement form to the MM1 ftp site, or email to proponents & ILG: Yonsei, KBS (through Yonsei), KDDI, SwissQual, NTT, Opticom, and Psytechnics. (note: add TDF to content receivers for RRNR-TV).

Action Item: Christian will create a mailing list of proponents & ILG for this purpose.

Action Item: deadline for sending out agreements from content providers to proponents & ILG will be 3 weeks – May 22, 2006.

Action Item: deadline for returning signed agreements to content providers – June 5, 2006.

Action Item: Arthur & Christian will keep a master list of who can receive what content.

NTIA & SVT providing content that does not need a signed agreement form but still is for research purposes only.

What to do if a proponent or ILG cannot sign some of the agreements? This will be decided later.

Agreement was reached: Proposed fees will be $6,000 per proponent, plus $500 per additional model past the first.

The precise figures per proponent will depend upon the number of proponents and whether proponents will receive a discount for conducting multiple subjective tests. Strong interest was expressed in favor of discounts but no suitable scheme could be agreed upon.

Agreement was reached: to *not* discount proponent fees when multiple proponent tests are conducted.

Christian offered to provide software to the ILG that will the perform data analysis.

Proposal to modify existing data plan on fitting to apply both logistics & cubic polynomial to each model/subjective test pair, and use the fit that works best for each model/subjective test pair. 2/3 vote required for change. Proposal failed, as follows:

In Favor: Opticom, NTT, KDDI, Psytechnics, SwisQual, Intel, Irccyn, Ericsson = 8.
Opposed: Yonsei, BT, NTIA, Verizon, NEC, Acreo. = 6

Agreement was reached to modify existing data plan change to perform the cubic polynomial monotonic fit as the primary fit, with the logistic fit as the backup in case of no convergence. 2/3 vote required for change. Proposal passed, as follows:

In Favor: Psytechnics, Opticom, Verizon, NTIA, NTT, Swissqual, KDDI, BT, Intel, Acreal.
Opposed: NEC

Telchemy presentation followed. See email sent to mmtest reflector from A Webster on this date. Software, provided free, can be used to create simulated transmission errors.

VGA Anchor Candidates:
Drtywin_64_15
Drmfeet_256_15
Fish2_128_15
Tea1_128_15
Brick2_128_15

Minutes for Thursday April 27th 2006 – Greg

Reviewed and approved minutes from Wednesday's session.

Discussion of RR: We have video material now, but still need more. Arthur takes as an action item to contact Universal and Teranex. We cannot do RRNR and MM simultaneously, so we do not need much time for discussing RRNR today.

Presentation by Jack Douglass of Spirent (which recently acquired SwissQual). Topic: G.1050/TIA-921 IP Network Model Overview, recently approved standard. Lists large number of parameters that affect performance. Presumably, the model handles much of this. Based the model on ITU-T Recommendation Y.1541 on network performance levels. Lists packet loss, delay, other parameters for 3 levels of network performance in 1541. List 8 impairment severity levels across the 3 levels of network performance quality. Lists 133 access rate combinations. Refers to 1064 combinations of access rates and impairment levels. Of these, a small number are appropriate to video.

Begins to describe real-life impairment profiles for good, bad, and severe network impairment conditions. Stresses that network packet loss is not described by a Gaussian distribution. Shows "network model coverage" statistic with MOS plotted against it. MOS scores are from the E-model. Anecdote about his own Vonage service degrading over time. One use for model is to predict network/service performance as traffic increases over time. Shows 1064 vectors of parameters for the network model. These parameters are empirical, derived from much input from service providers. Shows plots of simulated network performance for a number of impairment scenarios. Each scenario gives repeatable time series of impairments. So, you could store source video, run it through the model scenarios in real time, and use that as a way of producing impaired video stimuli for tests. Plays video examples with different impairment conditions.

Royce Fernald of Intel presenting on a method for detecting video gross error detection (GED) tool (preso available on VQEG ftp site). Will make tool available free, and is looking for possible VQEG endorsement. Concentrates on frame-level impairments, especially impairments over 802.11 broadband wireless. Use video clips with color blocks added. Play video through system being tested. Check to see if frames have made it through the system by looking at color blocks. This makes the method content-independent. Can be used for many transport and display conditions. GED coding adds color blocks to video, then looks for them later as video is played through system. Can detect dropped or repeated frames this way. The markers do not get destroyed by normal encoding, remain on fames after encoding/compressing or D/A conversion. A trick is to use very distinctive colors. Have tested this on all the major coders, and at many bit rates, down as low as 64 kbs. Requires a way of capturing video files as they come out of the system under test.

The GED metric is the sum of dropped, repeated, and out-of-sequence frames. So, how does GED metric correlate with subjective data? Experiments with humans show very high correlation. Digression on method of sampling the frame identifier color patches. Uses colors that are all 8 possible combinations of R,G, B values that are either 0 or 255. Q: Does method work with loss-concealment methods? Not clear.

Also can use composite marker sequences to uniquely mark each frame. Composite is a 3x3 grid with one of the 8 colors in each cell. Also can be used to keep track of concatenated video sequences and to mark start of sequences. Also can use markers to crop capture files with added frames. Also can use it to align original and processed video files for full-reference VQM. "Goal of GED is to take temporal element out of product evaluation." GED is used to reject wireless video products that are impairing large percentages of their frames.

Mentions GED application that is being made available to everyone, patent-free.

Subjective assessment vs. GED scores: Produced 100 impaired video clips with 10 levels of impairment. Tested with 50 subjects; ratings on 5-point impairment scale. Single presentation per video sequence. Impairments were dropping or repeating frames. Shows plot of ratings vs. GED metric; clear monotonic relationship. Also shows similar plots broken out by SRC. Plots are different for different SRCs. Shows nice plot of MOS vs. number of GED errors, which is declining log curve. MOS of 4 or above is possible for 1.3% GED frames or less.

Current standardization efforts: IEEE 802.11 Task Group T will accept methodology only if VQEG endorses. Will VQEG endorse? Royce will distribute software, then seek feedback. Arthur says VQEG does not do official endorsements. Filippo suggests that Task Group T send a liaison to VQEG asking for an opinion. Arthur notes that KDDI has a somewhat similar watermarking method that has been standardized. Tim Rahrer of Nortel strongly endorses. Arthur asks for a sort of vote:

DECISION

Does anyone object to VQEG verbally endorsing? No objections. The installer will be available on the VQEG ftp site.

Now on to RR/NR (A. Bourret & C. Lee co-chairs): Does VQEG want to tune the RR/NR work to IPTV? What that means is just using more packet loss impairments. Tim Rahrer says Nortel can do packet loss impairments. Also, we need to include MPEG4. If the license agreements for the video material brought to this meeting go through, then we will have enough video for RR/NR – especially if we can also get Universal and Teranex material. Filippo notes that since we are not going to run MM and RR/NR in parallel, we have time to change the Test Plan and/or create new HRCs for RR/NR. Some talk about mechanics of transforming video in one format to other formats for RR/NR. Apparently, there is no technical barrier. Kjell mentions 6.5 minutes of video that could be used for any of the programs, including HD.

ACTION ITEM (Chulhee see below)

Who will do an inventory of material suitable for RR/NR? Discussions about possibly changing the Test Plan a little to amend the list of HRCs.

Question: Should RR/NR be delayed to accommodate MPEG4 and H.264? Answer: No real need to delay? Do proponents need more time to accommodate H.264 in their models? Or, is the bottleneck the ILGs? Since only 3 of the 6 RR/NR proponents are here at the meeting, we can't say anything here about adding H.264. D. Hands says that it would be a disaster to delay RR/NR. **NOTE: BT, Intel, NTIA,**

Nortel all want H.264 in the test, 50% of HRCs.Audio call to decide (to be scheduled latter part of May) CRC says they cannot do RR/NR immediately after MM. Intel cannot either. BT, Nortel, and NTIA can make H.264 HRCs.

ACTION ITEMS: 1. Chairs will schedule conference call to include other (absent) proponents to discuss H.264 and new schedule for RR/NR. 2. Arthur will get some sort of resolution regarding Universal and Teranex. 3. Provide list of current source material and identify which of it is suitable for RR/NR (C. Lee).

HD Test Plan (Vivaik B. of Intel): **Action ITEM: Can Chulhee also indicate suitability for HD of video material that he is cataloging for MM and RRNR? Yes.** Reviewing test plan version from Stockholm meeting, Sept., 2005. Much of the Test Plan is copied from FRTV plans. Proposal is to put TP on reflector for comment, then to arrange a conference call. Details of TP: What kind of monitor? Answer: Not specified, but it should be good. Margaret notes that there should be greater specification of content to be used. **Scheduling conference call for Thursday, June 8, 6:00 am Pacific** time. Prospective proponents include KDDI, NTT, NTIA, Yonsei, BT. Prospective ILG: IRCCYN, Nortel, Verizon, Intel, FUB, CRC.

Patrick Le Callet (IRCCyN) on motion blur and CRT vs. LCD differences: This is a report of a subjective test of uncompressed HD (1080 50i) at 7 bitrates of H.264 from 2.5 to 10 Mbs. Compares studio monitor CRT vs. Philips LCD. Shows plots of MOS vs bit rate for LCD, CRT. There's a more or less constant offset, with the CRT having the advantage. LCD and CRT have different color response. Motion blur is greater for LCD. Response time does not account for motion blur. Because image in CRT is pulsed, blur from one image to the next is smaller than with LCD's more continuous display. They have a simulated demo of how blur occurs with LCD. Do experiment with human Ss to determine subjective blur width, which increases with picture velocity (pixels per frame). Their results agree with theoretical results of Pan et al. which show that 75% of blur is attributable to hold-type. These results also suggest ways to reduce motion blur in LCDs such as black-data insertion and back-light flashing. These results might be applied to displaying HD over LCD. The preso will be available on the VQEG ftp site. **Therefore, we may need to use CRTs for HD testing**. For the same reasons, LCDs are better for displays where there is not a lot of motion, such as PCs.

Kjell shows 6.5 minutes of HD video that is available, 1.3+ terabytes of data. How does anyone download that??? **(Up to 1080p is probably adequate for the HDTV test) Kjell is contact point.**

Back to MM and choice of bottom anchor for video content:
VGA:
NTIA_SRC_drtywnd_VGA_64_x264_15_.avi (Alex, restricted to P+ILG)
NTIA_SRC_drmfeet_VGA_256_x264_15_.avi (Alex, restricted to P+ILG)
FB_oki_4096_kbps_30fps_0plsc_7pltp (NTT, unrestricted)
football_VGA_H264_0320k_30fps (CRC?, unrestricted)
CIF:
football_cif_h264_128k_30fps (CRC?, unrestricted)
QCIF:
Swissqual pvs_39 (unrestricted)

Swissqual pvs_78 (unrestricted)
Ericsson 7-12_testclp_QCIF_64kbps_MPEG4_10hz_5percBLER
(segments: football, flowergarden) (unrestricted)
Yonsei lp13_m038_15.yuv (restricted to those signing KBS agreement)
Yonsei lp19_m038_15.yuv (restricted to those signing KBS agreement)
Yonsei mu11_m038_15.yuv (restricted to those signing KBS agreement)

Now on to MM player software: 1. Acreo, 2. Yonsei, 3. NTIA's is sick. Acreo's only runs RGB. We see demos of Acreo's and Yonsei's. Trying to determine whether there is any reason to choose one player over another. Answer seems to be "no." Also, the amount of effort required to make players available to users is about the same. Voting on players:

Player: Acreo, must make modifications.
Acreo: VZ, CRC, SwissQual, KDDI, IRCCyN, Intel
Yonsei: VZ

Intel agrees to perform de-interlacing for the MM test. Phil and Vivaik contacts. Should Intel, in principle, do the de-interlacing if it does a better job than the 3 current methods? Answer: Yes. If so, shall we give them test sequences so we can do a visual test? Answer: Yes. It would be best if the troublesome test sequences not require licensing agreement with Intel.

Much later in meeting: We find that de-interlacing of all source material would take 1.5 months by Intel. Margaret suggests not de-interlacing before model submission. De-interlacing also may depend on Intel signing licensing agreements. It's all very tangled. Chulhee suggests that proponents do their own de-interlacing and bring Intel in for de-interlacing only if proponents have problems. Meanwhile, Intel will de-interlace some test sequences to see if it is significantly better than other methods. **Margaret suggests that de-interlacing be done only on the source material that the ILG chooses for the test. ILG/Intel will de-interlace and re-size. Agreed by group.**

Review of MM schedule by items: Item 3, fee payment, not on schedule (now on schedule 19July2006). Item 4, source video, requires revision; this has to be linked to licensing agreements. Therefore, 2 weeks after signed agreements gives us 19 June for item 4. Item 4 renumbered to be Item 3. Step #3, fees, to be submitted by 19 July; renumbered to be Item 4. Discussion about how billing actually happens. Therefore, we need a deadline for ILGs to send invoices to proponents; that date is 31 May, and Filippo will tell ILGs who they bill. ILGs send invoices to proponents by 19 June. Item 7 (former item 5) on delivery of source video: Wording change, but no date change (roughly 20 July). New item 8 added to cover possible de-interlacing artifacts; date is Step 7 plus one month (20 Aug). Item 9 (former item 6) on when models are to be submitted: [Group is becoming addled and FUBAR.] It's getting very difficult to keep up with all the revisions to the schedule, especially where steps/ items are added or deleted. The step of de-interlacing complicates the specification of the schedule. Model submission is Step 5 (source acknowledgment) plus 3 months.

Item 12 (formerly item 7), no change. Item 13 (formerly #8), no change. Item 14 (formerly #9)

Minutes for Friday April 28th 2006 – Vivaik

Thursday's meeting minutes review (Arthur)
Reviewed and Approved.

Next Meeting:

KDDI willing to host in Japan depending on dates.
BT can host in London or Ipswich
BT and IRCCyN can co-host in France
Intel and NTIA can host in the US as well.

Dates

May need a face to face meeting to review the proposed subjective tests and make sure they are balanced. Fillipo suggested that it is possible to put a spreadsheet up on the FTP server with data on all planned tests so that we can have a continuous check.

A vote was called to decide on the review process for each subjective test being planned.

Option 1: Only the test lab does the checking

Option 2: All test materials are made available to everybody and everybody reviews it. If anybody chooses not to review them, they will not have a second chance to complain later.
Acreo
Psytechnics
CRC
NTT
Nortel
KDDI
BT
Yonsei
Intel Corporation
IRCCyN

Option 3: We have meeting to check each test
No votes

Option 4: ILG does the checking
Verizon
Intel

Option 2 was adopted.

A vote was called to decide on the distribution mechanism for the PVSs

Option 1: They are sent from proponent to test lab directly
No votes
Option 2: All PVSs are sent to a single point of contact
Acreo

Psytechnics
CRC
Nortel
NTIA
NTT
Intel Corporation
Swiss Qua
KDDI
BT
Yonsei
IRCCyn

Option 2 is adopted.

A vote was called to decide whether the no more than 20% of non-secret PVSs per proponent limit should apply to each individual test in addition to the application over all tests.

Votes in favor:
Nortel
Verizon
NTIA
NTT
SwisQuall
Lucent
KDDI
BT
Yonsei
IRCCyN

Votes for having more than 20%
Psytechnics

A decision has been made to set the limit to 20% per individual test.

The step by step process from start to finish

Proponents create the list of HRCs they can create
Proponents create an initial experimental design
Proponents and ILG will create PVSs to cover all test designs
PVS distribution will be created where each of the PVS in the test design, an organization will be assigned to create the PVS. This will be done by ILG
Each organization will be informed of the PVSs they need to create
PVSs are then sent to the point of contact (ILG-Philip-Intel)
Point of contact distributes all the PVSs to everybody
Balanced test review will be performed by each organization if interested
Tests are performed.

Schedule will be decided upon offline

For secret PVSs, the group agreed not to change the test plan and to allow ILG flexibility to select PVSs for each individual experiment. It is also open to the ILG if they want to choose to conduct secret test(s).

Minimum bit rates for PVSs

Option 1: To keep the test plan as it is:
No Votes.

Option 2: To change the bit rate specifications in the test plan.
Psytechnics
Nortel
Verizon
NTIA
NTT
CRC
Swiss Quak
KDDI
BT
Yonsei
IRCCyN
Intel Corporation
Acreo
Video Clarity

A decision was made to change the bit rate specifications

On how to change them:

To specify no minimum or maximum for bit rates and just specify examples.
For:
Psytechnics
KDDI
Intel

Against:
NTT

It was decided to have minimum bit rates specified in the test plan.
Vote on lowest bit rate for QCIF at 16 kbps
Psytechnics
Video Clarity
Nortel
NTIA
NTT
Swiss Qual
KDDI
Yonsei
Intel Corporation

16 kbps minimum is decided for QCIF

To reduce lower bit rate for CIF from 128 kbps:

For:
Psytechnics
Video Clarity
Nortel
NTIA
NTT
Swiss Qual
KDDI
Yonsei
Intel Corporation
Verizon
BT

Against:
None.

To reduce CIF bit rate to 64K

For:
Psytechnics
Video Clarity
Nortel
NTIA
Swiss Qual
KDDI
Yonsei
Intel Corporation
Verizon

Against:
NTT

Decision made to reduce it to 64kbps for CIF

On reducing the VGA minimum bit rate to less than 320K:

For:
NTT
KDDI
BT
Yonsei
NTIA

Against:
None:

For 128K as the minimum bit rate for VGA
NTIA
NTT
NEC
Yonsei

Against: None.

Minimum bit rate for VGA was agreed to be 128 Kbps

Need to have an audio call (Dave Hands or Phil Corriveau to set up)to decide on open issues and schedule.

BT will send an updated test plan by end of May and arrange audio call soon after. Vivaik will help work on testplan to get it out sooner if possible.

Liaison statements: JRG-MMQA drafted by AW and sent to Phil, Dave, Kjell, and Eugen (Pero) to SG12 and SG9.

Liason to SG9 regarding HDTV, RRNR-TV. (drafted by AW)

Vivaik will replace Mylene on Tools, Patrick will be added or replace Wolf on Tools.

Vivaik will be added as Co-Chair of HDTV