

MEETING RECORD

Video Compression Measurements Subcommittee G-2.1.6

Audio Video Techniques Committee G-2.1

Broadcast Technology Society

Institute of Electrical and Electronics Engineers

Nineteenth Meeting

Charleston Place

205 Meeting Street

Charleston, South Carolina 29401

January 29, 2001

Item 1 - Welcome and Introduction by Interim Chairman, of IEEE G-2.1.6

Interim Chairman Alan Godber called the meeting to order at 11:26 AM. Attendees introduced themselves. See Appendix B for a list of attendees.

Item 2 - Approval of Draft Agenda

A presentation by Alexander Woerner on digital video quality measurement was added as Item 5A after Item 5. A presentation from John Grigg was added as Item 5B. With these additions, the draft agenda was approved.

Item 3 - Review and Approval of Minutes of the Previous Meeting #17, August 7, 2000

The Minutes from Meeting 18, October 30, 2000 were reviewed. The second sentence in Item 5 was modified to read: "He said that five contributions, from Tektronix, NTIA, CRC and Sarnoff (including one duplicated) were submitted to Working Party 1, Study Group 9 (SG9) and said essentially the same thing – none of the objective video quality measurement methods listed in the SG9 recommendation Appendix 1 should be included as normative."

With that correction, the minutes were approved.

Item 4 - Matters Arising from the Minutes

There were no matters arising from the Minutes.

Item 5 - Update Report of ITU Video Quality Experts Group (VQEG) re ongoing testing plans and meeting in Munich, Germany - Arthur Webster, David Fibush, Phil Corriveau, Al Morton, Ann Marie Rohaly and other participants

Arthur Webster reported on the Study Group 9, Working Party 1 (WP 1/9) meeting in Beaverton, Oregon the week of November 6, 2000. WP 1/9 sent a contribution to VQEG recommending VQEG continue its efforts to continue work on Full Reference objective video quality measurements. A draft of recommendation J.144, *Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference* was approved at the meeting. The document will be submitted for final approval at the Study Group 9 meeting on March 6, 2001. J.144 stated before one or more of the methods listed in it would be promoted to normative status, independent validation of improved measurement quality over other methods would be required. It was noted that the plan is to eventually have one measurement method for Full Reference TV.

The video quality question Q11 from Study Group 12 was moved to Study Group 9 as question 21. Arthur Webster was made Rapporteur for the question. Study Group 9 video quality question Q22, concerning television level quality, is now question 4. Alina Karwowska-Lamparska was made Rapporteur for that question.

At the March 6 meeting Study Group 9 may consider a proposed recommendation on loop latency in conversational television programming. When asked how much delay is acceptable, the response was that 1.5 to 2 seconds is considered acceptable.

There was a VQEG ad-hoc meeting December 12-13, 2000 at IRT in Munich for work on reduced-reference/no-reference (RR/NR) quality measurements. At the meeting, items in the proposed draft test plan were reviewed. Although the review was not completed, several important points were considered and tasks assigned, including acquisition of SRC materials, HRCs, etc.

VQEG hopes to finalize the RR/NR test plan at its May 7-11 meeting in Boulder, Colorado. A few days were allocated for NR/RR meetings, which would be followed by a full VQEG meeting. Full reference (FR) tests should be an overlay to the NR/RR tests, although it was noted that if the FR group wanted to do DSCQE tests instead of using SSCQE, the FR group would have to supply the resources for it. People involved with FR measurements hope to be able to examine the HRC/SRC combinations being considered for the NR/RR tests to be sure they have an appropriate spread of degradations. It was mentioned that in the previous VQEG tests, a high percentage of the degradations were very small and some believe this is why the results were inconclusive.

5.1 Further Discussion and Recommendations from the Subcommittee to the next meeting of VQEG

There was a comment that although letters from CBS, ITU-SG9, NABA and others urged T1A1.1 not to move ahead with writing a standard for objective measurement of the quality of compressed video, this opposition was based on the assumption that ITU and VQEG would move along quickly with their work towards a standard. This may increase support for the VQEG efforts. Everyone was invited to the VQEG meeting in Boulder, Colorado in May. Arthur Webster suggested looking on the VQEG web site or email reflector for tests plans around May 15.

Viewing distances for the tests were discussed. Some tests at 3H were recommended, but it was pointed out that if scan lines are visible at 3H, this could be a problem with non-expert viewers. 4H might be better in that case. It was explained that in the broadcast area, you want to be able to see very small degradations. It was suggested a selected set of HRC/SRC combinations be used to evaluate measurement systems and test methodology at closer viewing distances. Another problem, however, with a 3H viewing distance is that the number of people able to view each run is smaller than with longer viewing distances.

Recommendations to VQEG were discussed. VQEG was encouraged to proceed with its tests, as the inclusive full reference (FR) tests must be resolved before any methods are considered for standardization. It was suggested that we support evaluating the use of closer viewing distances for the tests. The planned viewing distance for the RR/NR tests is 4H. There was a recommendation that tests be done at multiple viewing distances, but it was pointed out that if the FR tests have more than one viewing distance and the RR/NR have

one, it would make it more difficult to compare results. There was agreement that objective data should not be sampled at a frame rate, but filtered over a half second time constant. It was suggested details on this proposal be posted to the VQEG reflector. Andrew Watson said he would discuss this with Harley Myler, who was put in charge of the objective part of the new VQEG test plan.

A request was made for contributions to the VQEG testing. Facilities with subjective test capabilities are needed, as is a volunteer to edit tapes for the tests.

[Action Item] Alan Godber said he would write a letter to VQEG encouraging their work, offering to help and recommending use of a 3H viewing distance in the tests.

Item 5A – A Real Time No-Reference / Reduced-Reference Video Quality Measurement Algorithm – Alexander K.G. Woerner

Alexander Woerner presented *A Real Time NR/RR Video Quality Measurement Algorithm* (Rohde & Schwarz, Inc, IEEE Doc. G-2.1.6/121). The presentation is available on the G-2.1.6 web site. It describes a method for analysis of video sequences compressed using DCT, in real time, either with or without a reference. The system can use either no reference, a reduced reference, or a full reference input signal. It looks at amplitude differences between adjacent pixels and notes peaks at macro block and block boundaries. The results are weighted depending on the high spatial and temporal values. Examples were shown comparing compressed video through a constant data rate multiplex and the same program encoded using a statistical multiplex.

Rohde and Schwarz use the technique in the commercially available DVQ instrument. Advantages include usage of typical video material, no need for short and less representative test sequences and no requirement for a reference signal allowing it to be used at any point where a signal is accessible. Alexander Woerner said tests showed the method had the same correlation to subjective assessments as full reference algorithms. One negative is that the unit can only detect quality degradations from compression artifacts.

5.1 Further Discussion

In response to a question on how the unit handles decoded block sizes that differ from 8 or 16 units due to preprocessing, Alexander Woerner answered that DVQ determines the block size before doing the measurements. He was unable to directly answer a question on how it handled non-blocking artifacts like mosquito noise, but did state that if the artifacts show an MPEG block structure, the DVQ should catch it. There have been no tests on this to date. There was a general impression that this method depended on the block structure and would not work well if the block structure was hidden or if the artifact involved softening. The committee was advised to wait until the results of the VQEG NR/RR tests to see if this was really the case. A standardized set of test tapes with different types of artifacts may also be useful in testing the performance of this method.

The graph in the presentation showing correlation between the DVQL-W reading and subjective scoring was discussed. SSCQE was used to obtain the subjective score, with the results averaged over half a second. It was noted that the graph illustrated that in any half-second period there may be subjective and objective readings that don't agree, but on average they do.

Item 5B – Report on activity in T1A1 on a North American Standard for Full Reference TV Picture Quality Objective Metrics – John Grigg, T1A1

John Grigg summarized the development of picture quality objective measurement methods, starting with the ANSI standard based on PSNR that had limited application and continuing through the introduction of the Sarnoff method in 1997 and the VQEG evaluation of various methods. He noted that the proponents in the VQEG tests ended up in a statistical dead-heat and, so far, there had been no second round of testing and no clear winner. The lack of standard has been a problem for video carriers as there is no way to establish services based on digital video quality.

John Grigg explained that many entities have purchased the Tektronix picture quality measurement instrument based on the Sarnoff method. Comments from users indicate its good enough. Since the process of setting an objective picture quality measurement standard doesn't seem any to be any closer than it was three years ago, T1A1.1 authorized a technical subcommittee to develop a draft standard.

There were differences of opinion on when VQEG would be done with its work. While some say it will only take a year, John Grigg figured it would take 5 years. He stated there is work to be done and an objective standard is needed now. Once the standard is set, he hoped other manufacturers would build measurement equipment to it.

5.1 Further Discussion

There was concern that the standard would be based around Tektronix's proposal to the exclusion of other proposals, when there is no evidence one system is better than the other. Alan Godber noted that if this committee proceeded to do this, the standard board and many people in the industry would stop us. We must study alternatives. It was pointed that the way SMPTE handles standards is if manufacturers and users both want a standard, SMPTE would agree to it unless others opposed it.

The PSNR method received support from some attendees, who felt if a standard were needed now; it would be the best choice. It was pointed out that the only instrument that measures PSNR is the Tektronix instrument. In addition, tests by IRT showed that the method proposed by Tektronix had a higher correlation than indicated in the VQEG tests. It was argued that these tests were not verified in the U.S. and IRT was paid by Tektronix to do the tests.

Several attendees felt it was premature to choose one approach over another as a standard. Several broadcasters sent letters to T1A1 opposing setting a standard at this time. They are the video carriers' customers. It was pointed out, however, that the networks have offered noting to this process recently, so they shouldn't complain.

It was suggested that instead of calling the Tektronix proposal a standard, it be released as an operating practice or technical report. For writing tariffs, a technical report could serve the same function as a standard.

There was no opposition to releasing the T1A1 proposal as a technical report.

Item 6 - Selecting Test Method, Test Material and Test Labs for a Unit of Measure (JND) & a Means of Calibration for Video Impairment, Chair, Leon Stanger

Andrew Watson presented his *Outline of a Standard for the JND Method of Evaluating Video Quality*, IEEE Doc. G-2.1.6/122. Items in the standard were discussed point by point.

We need to clearly outline the scope and purpose of the standard. Section 4, Comparison with Existing Standards, should acknowledge other standards and discuss how this one is different. Rec. 500 was listed as a reference in Section 15. Only items essential to implement the standard should be listed here. This is a normative reference. More definitions could be added to Section 16. HRC was discussed. In the definition section, it was suggested the definition of HRC only cover MPEG codecs and similar devices. Calling 601 an HRC is confusing. The reference would be a null HRC.

Section 17 received considerable discussion. An abstract block diagram, rather than a list of specific equipment, should be satisfactory here. An informative annex could show how to do the tests in the real world. The specifications for the method need to be detailed. There was concern about being so general that every subjective test would become different from the other. It was suggested we limit generalities now, so we can generate a calibrated tape. In the future, we may want to make it more general. Items to be specified include display size and resolution, durations, viewing distance, screening of observers, etc. It was suggested a section be added giving an over-all description of a test session. The severity of the HRC should also be addressed.

Appendices cover how to estimate JND on the scale from the binary comparison data. This is largely mathematical. More figures would be helpful here. In the description of the EASE method, the question was asked whether we wanted to put software in the standard. Pseudocode was recommended. It was noted that in

ITU, they have C code as part of the standard. Andrew Watson said he used Mathematica, which is a form of Pseudocode.

6.1 *Further Discussion and Action*

Activities for the committee were discussed. Andrew Watson said he could use some help on the introduction, scope and purpose. It was suggested we assign someone to do the first three sections, however, after discussion it was agreed that Andrew Watson would handle the whole document. **ACTION ITEM:** Rick Redford will write specifications for the display, screening of observers, etc. Suggestions were solicited for the first three sections. They should be sent by email to Andrew Watson.

It was decided not to post the current document on the web site. **ACTION ITEM:** Doug Lung will set up a private website area for G-2.1.6 for a copy of the standard as it is being worked on. A memo will be posted on the G-2.1.6 email reflector asking members to comment on sections in the standard.

Item 7 - Further Discussion of Compression Measurement Methodologies

Item 7.1 – Discussion of Future Work, Additional Assignments

There was no further discussion.

Item 8 - Any Other Business.

There was no other business.

Item 9 - Date(s) of Future Meeting(s).

The next meeting will be held in conjunction with the May VQEG meeting in Boulder, Colorado. The meeting was tentatively set for May 11, 2001, following the G-2.1.4 meeting.

The subcommittee thanked John Grigg and T1A1 for arranging the room.

The meeting was adjourned at 5:57 PM.

Submitted by:
H. Douglas Lung
Secretary

APPENDIX "A"

List of Documents Distributed

29 January 2001

Draft Agenda - IEEE Compression and Processing Subcommittee G-2.1.6, Nineteenth Meeting, Monday, January 29, 2001, Alan Godber, Chairman, ([216m19an.pdf](#))

Draft Meeting Record, G-2.1.6, Compression and Processing Subcommittee, Eighteenth Meeting, October 30, 2000, San Antonio, TX, Doug Lung, Secretary, [IEEE Doc. G-2.1.6/120](#), January 22, 2001.

A Real Time NR/RR Video Quality Measurement Algorithm, Alexander K. G. Woerner, Rohde & Schwarz, Inc, January 29, 2001, IEEE Doc. G-2.1.6/121. [DVQ-Charleston.zip](#)

Outline of a Standard for the JND Method of Evaluating Video Quality, Andrew B. Watson, NASA Ames Research Center, January 29, 2001, IEEE Doc. G-2.1.6/122.

APPENDIX "B"
ATTENDANCE RECORD
January 29, 2001

Name	Affiliation	Telephone	Fax	E-mail
Chairman: Alan Godber	Consultant	(732) 846-4476	(732) 846-4476	mailto:agodber@earthlink.net
Secretary: Doug Lung	Telemundo	(305) 884-9664	(914) 412-2886	dlung@transmitter.com
David Fibush	Tektronix	(503) 628-3040	(503) 627-4486	dave@fibush.net
John Grigg	Qwest	(303) 707-5604	(303) 707-9041	jjgrigg@uswest.com
Michel Poulin	Leitch	(416) 443-2716	(416) 445-4762	michel.poulin@leitch.com
Rick Redford	ABC	(212) 456-4450	(212) 456-4451	rick.redford@juno.com
Leon Stanger	DirecTV	(310) 726-4676		LStanger@compuserve.com
Dick Streeter	Consult/CBS	(908) 791-9876	(908) 791-9878	rstreeter@att.net
Andrew Watson	NASA	(650) 604-5419	(650) 604-0255	mailto:abwatson@mail.arc.nasa.gov
Arthur Webster	NTIA / ITS	(303) 497-3567	(303) 497-5969	mailto:webster@its.bldrdoc.gov
Alexander K. G. Woerner	Rohde & Schwarz	(410) 910-7836	(410) 910-7801	Alexander.Woerner@rsa.rohde-schwarz.com