

**Committee T1 Performance
Standards Contribution**

.....
Document Number: T1A1.5/98- 100
TIBBS File:
.....
DATE: March 16, 1998
.....
STANDARDS PROJECT: Analog Interface Performance Specifications for Digital Video
Teleconferencing/Video Telephony Service (T1Q1-12)
.....
SUBJECT: Summary and meeting reports for Questions 10 and 11 from
ITU-T SG12: Objective and subjective methods for evaluating
audiovisual quality in multimedia services
.....
SOURCE: NTIA/ITS
.....
CONTACT: Arthur Webster (303-497-3567)
.....
KEY WORDS: video quality, picture quality, subjective, objective,
performance assessment, multimedia, JRG, ITU VQEG.
.....
DISTRIBUTION: Working Group T1A1.5
.....

Summary of Video Quality Assessment Work at ITU-T Study Group 12 Meeting Geneva, February 1998

Q11/12

P.DEL (Multimedia communication delay, synchronization, and frame rate measurement) was determined at this meeting and was given the number P.931. Al Morton is the editor of P.DEL and he utilized the text of T1.801.04 with modifications to include additional video formats and to address comments. T1.801.04 -1997

Cooperative work with ITU-T SG9 and ITU-R WP11E (and other standards bodies such as T1A1.5 and IEEE G-2.1.6) will continue through the ITU VQEG. The main focus of this group through the summer of 1998 will be the execution of the validation test for objective measures of video quality over bit rates between 768 kbits/s and 36 Mbits/s. P.OVQ will utilize the results of this test as will other Recommendations in SG9 (J.ovq) and WP11E.

Further validation tests of objective measures of video quality will be planned after the initial test is complete. These future tests will probably begin with at least one for lower bit rates and one which will more fully investigate the effects of transmission errors.

Comments from SG12 on the VQEG documents submitted for review can be found in the liaison statement on page 9 of this contribution.

Q10/12

Draft New Recommendation P.AVQ (Subjective audiovisual quality assessment methods for multimedia applications) was determined and given the number P.911. There were many modifications made at this meeting. The inclusion in the Recommendation of the revised Video classes table and a new Audio classes table will be of particular interest to IEEE G-2.1.6 and T1A1.5.

Also of interest is the inclusion of a model for estimating overall audiovisual quality from the individual audio and video subjective evaluations. Recent research by NTIA/ITS and KPN Research (Netherlands) confirms the previous research done by Bellcore in this area. It was noted that the primary model weight is surprisingly close in each of the three experiments. It was further noted that the video quality rating dominated the overall audiovisual quality for the cases tested. The conditions for which this applies are limited to one way tests when the audio and video are synchronized. This model was included in an Annex of P.AVQ.

Some modifications to the text of P.AVQ are possible till May 15, 1998. The draft text is available on the VQEG/JRG anonymous ftp site:

[ftp.its.bldrdoc.gov](ftp://ftp.its.bldrdoc.gov)

Web address: <ftp://ftp.its.bldrdoc.gov>

Geneva, 18 - 25 February 1998

Question: 11/12

SOURCE: Rapporteur for Q.11/12 (Arthur Webster)

TITLE: Progress Report of Question
"Objective methods for evaluating audiovisual quality in multimedia services"

1. Introduction

During last study period (1992-1996) SG12 studied objective methods for evaluating audiovisual quality as part of Question 22/12. Due to the large amount of work covered by Question 22/12 the question was divided into two new questions (Questions 10/12 and 11/12). Question 10/12 studies subjective methods and Question 11/12 studies objective methods. The text of these Questions can be found in COM 12-1 (p.22-23).

The output from Question 11/12 will consist of new Recommendations covering topics such as objective assessment of one-way video quality (P.OVQ), objective assessment of one-way audiovisual quality (P.OAV), the measurement of multimedia communication delay, synchronization, and frame rate (P.DEL), as well as objective assessment methods for two-way interactive audiovisual services. These Recommendations are expected to devote particular attention to systems used for videoconferencing/vidotelephony and other multimedia services.

Work on aspects of this Question is being undertaken with the cooperation and participation of experts from outside the ITU and from other Study Groups including ITU-T SG9 and ITU-R WP11E.

2. New documents available

- *COM 12-19 Relations between audio, video and audiovisual quality (KPN Research)*
This contribution quantifies the mutual influence of audio and video quality on the combined audiovisual quality and gives a data model for mapping the individual audio and video qualities onto an overall audiovisual quality.
- *COM 12-22 Towards a multi-sensory perceptual model (BT)*
Experimental evidence of the interaction between hearing and vision is presented and a proposal made for a multi-sensory model architecture.
- *COM 12-23 Hypothetical reference circuit selection—proposal to include MPEG-2 with transmission impairments (USA)*
This contribution proposes that the first round of tests conducted by the ITU VQEG include a limited testing of all objective measurement methods to investigate their applicability in systems that exhibit transmission impairments. A minimum of 10% of the clips including impairments is proposed.
- *COM 12-29 Draft new Recommendation on multimedia communication delay, synchronization, and frame rate measurement (Rapporteur)*
This draft New Recommendation describes measurement for the time related aspects of Multimedia Communications Systems (with Audio, Video and Data channels). This second draft of P.DEL adds sections on References, Definitions, Abbreviations, Data Channel Measurements, and Timer Stability and Synchronization. It has been modified to include additional video formats and to address comments. It is felt that the draft is stable and is proposed for determination.

- *COM 12-30 Proposal for the validation of objective models for automatic video quality evaluation (CSELT, EBU)*

This document is intended to inform all the Administrations about the VQEG's proposal concerning the validation of objective measurements of video quality and to solicit submissions of objective models to be included in the ITU Verification Process leading to one or more ITU Recommendations.

- *COM 12-39 Video quality assessment using objective parameters based on image segmentation (Brazil)*

This contribution presents a methodology for video quality assessment using objective parameters based on image segmentation. This methodology is applied to natural scenes and MPEG-2 video codecs. The results presented show that the use of region-based objective measurements provides more accurate predictions compared to predictions based on global parameters.

- *D.50 Examination of agreement between laboratories conducting identical subjective video quality experiments (AT&T)*

This contribution describes the lab-to-lab comparison of subjective test data collected in a large video performance research project. Comparisons of identical subjective tests conducted at different laboratories will be analyzed to assist in the validation of objective measures of video quality. These results are applicable to the current testing planned by the VQEG

- *D.51 Proposed definitions for video classes and video terms from the VQEG (Rapporteur)*

This contribution represents the output of the VQEG ad hoc group on Video Classes and Definitions.

These definitions are offered for review and comment by SG12 and are expected to be included in Draft New Recommendations.

- *D.52 Evaluation of new methods for objective testing of video quality: subjective test plan (Rapporteur)*

This contribution represents the output of the VQEG subjective test plan ad hoc group. This subjective test will be used to evaluate objective measures of video quality over a range of bit rates between 768 kb/s to 36 Mb/s.

- *D.53 Evaluation of new methods for objective testing of video quality: objective test plan (Rapporteur)*

This contribution represents the output of the VQEG objective test plan ad hoc group. This objective test plan will be used in the evaluations of objective measures of video quality over a range of bit rates between 768 kb/s to 36 Mb/s.

- *D.55 Report of the Experts meeting on subjective and objective video quality assessment, Turin (October 1997) (Rapporteurs)*

This contribution contains the report of the first meeting of the ITU VQEG (Video Quality Experts Group) which met at CSELT in Turin, Italy in October, 1997. The group's participants include video and audiovisual quality experts from ITU-T Study Groups 9 and 12 and ITU-R WP 11E. This report discusses both objective and subjective methods and contains a useful flow diagram outlining the current validation testing effort which is detailed in D.52 and D.53.

- *D.59 Detection and discrimination of blur in edges and lines (AT&T)*

This contribution summarizes the results of a psychophysical experiment to determine viewer sensitivity to blur in edges and lines. These results add to our knowledge of the human visual system and are applicable to the study of the assessment of video quality.

- *TD 001 (GEN/12) Multimedia Coordination--Liaison Statement from ITU-T SG13 (WP1/13) (For information and comment.)*

This contribution offers an updated "Activities status report" on multimedia Recommendations in progress in the ITU-T.

- *TD 007 (GEN/12) Determined SG13 Multimedia Recommendations --Liaison Statement from ITU-T SG13 (WP 1/13) (For information.)*

This contribution advises that two new Recommendations were determined at the September 1997 meeting of SG13. The titles are: "Network Capabilities to Support Multimedia Services - General Aspects" and "Network Capabilities to Support Multimedia Services - Example of Multimedia Retrieval Service Class - Video-on-Demand Service using ATM-based Network".

- *TD 00 2 (WP2/12) Liaison statement to ITU-R WPs 11B and 11E, ITU-T SGs 9, 12, and 15 and ISO/IEC/CIE on Adaptive Image Quality Control in Future TV Systems. From ITU-R, WP 11A*

This contribution invites SG12 to examine their possible participation in the studies covered in a new question ITU-R [AM/11A]. Adaptive image control will probably utilize objective video quality parameters and is therefore of possible interest to the work of Question 11/12.

- *TD 003 (WP2/12) Liaison statement to ITU-R WPs 11B and 11E, ITU-T SGs 9, 12, and 15 and ISO/IEC/CIE on Assessment and Optimization of Quality of Colour Reproduction in Television. From ITU-R WP 11A.*

This contribution invites SG12 to examine their possible participation in the studies covered in a new question ITU-R [AL/11A]. This subject is of possible interest to the work of Question 11/12.

- *TD 007 (WP2/12) Reply to liaison statement from SG12 concerning Joint Rapporteurs Group on picture quality assessment- From ITU-T SG. (For information).*

This contribution acknowledges the formation of the JRG between Study Groups 9 and 12. It also attaches two draft New Recommendations J.svq and J.ovq and requests comments.

- *TD 008 (WP2/12) Liaison statement from ISO/IEC JTC 1/SC 29 Coding of Audio, Picture, Multimedia and Hypermedia Information-- Audio Tests*

This contribution announces MPEG's plans to carry out verification tests on the MPEG-4 audio coding standard. They ask for suggestions and/or objections to their published plans. They also invite collaboration and/or assistance in these tests.

- *TD 009 (WP2/12) Liaison statement from ISO/IEC JTC 1/SC 29 Coding of Audio, Picture, Multimedia and Hypermedia Information-- Video Tests*

This contribution announces MPEG's plans to carry out error resilience verification tests on the MPEG-4 video coding standard. They ask for suggestions and/or objections to their published plans. They also invite collaboration and/or assistance in these tests.

- *TD 010 (WP2/12) Liaison statement from ISO/IEC JTC 1/SC 29 Coding of Audio, Picture, Multimedia and Hypermedia Information-- Video testing material*

This contribution announces MPEG's willingness to provide SG12 with source material, coded material, and corresponding subjective evaluations to be used in our work. It stipulates that the copyright owners must be contacted for their permission. A list of available material is attached.

- *TD 012 (WP2/12) Reply to RAG TD/124 on coordination activities on potential overlapping between sound and video services questions. From the TSAG Chairman.*

This liaison presents the TSAG's decision (and the rationale behind it) concerning the RAG's proposal to establish a Joint Committee on Sound and Video Services. The TSAG determined that there is no need to establish a Joint Committee to resolve overlap between the ITU-T and the ITU-R. The TSAG goes on to encourage the continuing work of the video quality experts group, which includes members of the ITU-T Study Groups 9 and 12 and the ITU-R WP11E. Any problems concerning overlap of work should be resolved by the Study Group Chairs.

- *TD 013 (WP2/12) Responsibilities for development of standards for multimedia communications. (For information)*

This contribution from SG16 is for information only and outlines how it views its responsibilities as lead study group for the standardization of multimedia communications services and systems.

-

3. Status of Studies

The work on objective measures of video quality is proceeding rapidly. Draft New Recommendation P.DEL (COM 12-29) is ready for determination at this meeting. It has undergone a period of review and we consider it a mature document.

The Rapporteurs' meeting scheduled to be held in Turin became a broader experts meeting, encompassing participants from ITU-T Study Groups 9 and 12 and ITU-R WP11E. This was the first meeting of the ITU VQEG (Video Quality Experts Group). The report can be found in D.55.

With the combined expertise of the participants of VQEG, an ambitious schedule for the performing of validation tests for objective measures of video quality (over bit rates between 768 kb/s and 36 Mb/s) is well under way. No fewer than eight contributions at this meeting are related to this test. The subjective and objective test plans (D.52 and D.53) are outputs of the VQEG and are input to the participating Study Groups for review and comment. COM 12-30 officially announces the testing process and invites interested organizations to announce their intended participation in the process by May 1, 1998. A computer program (object code) is due by June 1 if they wish to participate in this first round of testing. If the work continues according to schedule, draft New Recommendation P.OVQ should be available at the November meeting.

A first draft of P.OAV (Objective assessment of audiovisual quality) may be available in November, but it is possible that the schedule for this draft New Recommendation will slip a meeting.

Items for discussion at this meeting:

Draft New Recommendation P.DEL – (P.931)

COM 12-29

Audiovisual objective measures -

COM 12-19, COM 12-22, TD 008(WP2/12), TD 009(WP2/12)

Database of subjectively rated video-

TD 010 (WP2/12)

Objective measures of video quality and VQEG validation test for 768 kb/s to 36 Mb/s -

COM 12-39, TD ? (WP2/12), COM 12-23, COM 12-30, D.50, D.51, D.52, D.53, D.55, D.59

Other -

GEN/12: TD 001, TD 007

WP2/12: TD 002, TD 003, TD 007, , TD 012, TD 013

4. Summary of discussion during the meeting

Draft New Recommendation P.DEL -(P.931)

Minor editorial changes were suggested and included in the latest draft. It was agreed to seek determination at this meeting. P.DEL has received circulation at the Turin VQEG meeting in October 1997 and has been available by ftp since before that meeting. A draft was also submitted at the last SG12 meeting. It was suggested that we send P.DEL to ITU-R WP11A and therefore it has been attached to the liaison statement to that group (TDXX).

Audiovisual objective measures -

The workplan shows that the schedule for draft P.OAV will be delayed one meeting. The models found in the KPN document COM 12-19 can be utilized in P.OAV, but suitable objective measures for video and audio must be determined first. The work presented in the BT document COM 12-22 can also be utilized and the future work from that lab is expected to produce more data to add to the knowledge in this area.

Database of subjectively rated video-

The database from the MPEG2 tests are becoming available as the permission of the copyholders is obtained. It is expected that a large database will be available by the next meeting of SG12. It is hoped that the upcoming MPEG-4 tests will yield more material for the database. The database of subjectively rated video from the T1A1 VTC/VT test is also available (from NTIA/ITS) for research on objective measures. The fact that it is only available on D2 and Betacam SP, however, makes it less desirable for high quality testing.

Objective measures of video quality and VQEG validation test for 768 kb/s to 36 Mb/s -

The objective video quality work of CPqD/TELEBRAS (Brazil) was presented by Franca Pessoa (COM 12-39 and TD-25). Their work will be used in the VQEG validation test.

The formal announcement of the VQEG validation test was made in COM 12-30.

Three documents from VQEG were reviewed and the comments agreed to by the ad hoc group can be found in the Liaison to SG9 and WP11E (TDXX). In the work on Q10/12 a table defining audio classes was drafted and it was suggested to include this in the VQEG document on Definitions and Classes.

COM 12-23 was discussed and a video tape was viewed which illustrated potential transmission impairments. It was agreed to mention in the comments on the subjective testplan of VQEG that other transmission impairments besides those utilizing ATM, for example, satellite burst errors, might be included. There was general agreement to include at least 10% transmission errors in the first round of testing. Mr. Vittorio Baroncini (FUB, Italy) disagreed and felt that transmission errors should not be included in the first test, but examined in another, separate test. He feels that viewers might mistake transmission errors for coding errors.

D.50 was discussed and illustrated the high correlation between different labs conducting identical subjective tests. This information contributed to the comments on the subjective test plan found in the liaison to WP11E and SG9.

The report of the Turin meeting was briefly presented.

D.59 was presented and the information underscored the need to utilize aspects of the human visual system in the development of objective measures of video quality.

Other -

The need for an interim meeting of the VQEG was briefly discussed and it is felt that, depending on the comments received from Study Group 9 and WP11E, a meeting might be needed in April or May to resolve the comments. A meeting may also be needed in the September to review the analysis of data from the subjective and objective tests. Work will continue with the VQEG by correspondence.

Four liaisons were drafted:

Title: Liaison to ITU-T SG9 and ITU-R WP11E
Draft P.AVQ will be attached (from the work on Q10/12).

Title: Liaison to ITU-T Study Group 9
This liaison includes comments on draft New Recommendations J.svq

and J.ovq.

Title: Reply to two Liaisons from ITU-R WP11A
A copy of P.DEL will be attached as requested from participants of the ad hoc group.

Title: Reply to Liaison from ITU-T SG13 (WP1/13)
Attached to this liaison is an updated work plan for Questions 10/12 and 11/12, as requested in the liaison.

5. Final conclusions and updated workplan

P.DEL is sent forward for Decision.

Cooperative work with SG9 and ITU-R WP11E will continue through the VQEG. The main focus of this group till the next meeting will be the execution of the validation test for objective measures of video quality including bit rates between 768 kbits/s and 36 Mbits/s. P.OVQ will utilize the results of this test.

Advancement of P.OAV must wait until objective measures of audio and video quality are identified.

Further validation tests of objective measures of video quality will be planned after the initial test is complete. These future tests will probably begin with at least one for lower bit rates and one which will more fully investigate the effects of transmission errors.

Workplan

Schedule

Status

Study item: Multimedia Communication Delay, Synchronization, And Frame Rate Measurement

Deliverable: New Recommendation P.DEL

Milestone

First draft of P.DEL
P.DEL ready for determination

Schedule

April 1998
February 1998

Status

Done
Done

Study item: Objective assessment of video quality

Deliverable: New Recommendation P.OVQ

Milestone

Subjective and Objective Testing Completed
Data Analysis Completed
First draft of P.OVQ
Draft P.OVQ ready for determination

Schedule

August 1998
September 1998
November 1998
August 1999

Status

Study item: Objective assessment of audiovisual quality

Deliverable: New Recommendation P.OAV

Milestone

First draft of P.OAV
Stable draft text for P.OAV

Schedule

August 1999
2000

Status

6. References

Annexes

Report of Turin Meeting (October 1997) D.55
Draft New Recommendation P.DEL

Rapporteurs' Group members for Question 11/12

Name	Company	Tel. Fax	E-mail
Yusuf Ansari	Rockwell International Corporation	+1 714 2215062 +1 714 2213667	ansanys@nb.rockwell.com yusuf.ansari@nb.rockwell.com
J. Antkowiak	Deutsche Telekom Berkom	+49 30 34972490 +49 30 34972491	
D.J. Atkinson	U.S. Dept of Commerce NTIA/ITS	+1 303 497-5281 +1 303 497-5323	dj@bldrdoc.gov
Vittorio Baroncini	FUB Italy	+39 6 54802134 +39 6 54804401	Vittorio@fub.it
John Beerends	KPN Research	+31 70 3325644 +31 70 3326477	j.g.beerends@research.kpn.com
Paul Bowman	BT	+44 1473 644105 +44 1473 644649	paul.bowman@bt-sys.bt.co.uk
Noel Chateau	France Telecom		noel@cnet.francetelecom.fr
Laura Contin	CSELT	+39 11 228 6174 +39 11 228 6190	laura.contin@cselt.stet.it
Jean-Pierre Evain	EBU	+41 22 717 21 11 +41 22 717 27 10	evain@ebu.ch
Mauro Falcone	FUB Italy	+39 6 5480 3352 +39 6 5480 4405	falcone@fub.it
N. Fujino	Fujitsu	+81 44 7543076 +81 44 7543514	nfuji@sig.ts.fujitsu.co.jp
David Fibush	Tektronix	+1 503 6276289 +1 503 6271707	davef@tv.tv.tek.com
Andreas Kamcke	Siemens, AG	+49 89 722 45078 +49 89 722 37078	andreas.kamcke@pn.siemens.de
J.Y. Montfort	FT/CNET	+33 2 96053171 +33 2 96051316	montfortj@lannion.cnet.fr
Al Morton	AT&T	+1 908 9492499	acmorton@att.com
R.A. Osorio	Telemig Brazil	+55 31 2294075 +55 31 2291853	ramo@telemig.gev.br
D. Pascal	CNET/France	+33 2 96051578 +33 2 96051316	pascal@lannion.cnet.fr
Franca Pessoa	CPqD-Telebras Brazil	+55 19 705 6746 +55 19 705 6833	franca@cpqd.com.br
A. Takahashi	NTT	+81 422 59 4447 +81 422 59 2330	aki-tak@nttgn.tnl.jp
Yoshihiro Tomita	Fujitsu		ytomita@sig.ts.fujitsu.co.jp
Arthur Webster	US Dept. of Commerce NTIA/ITS	+1 303 4973567 +1 303 4975323	webster@its.bldrdoc.gov

Geneva, 17-27 February, 1998

Questions: 10/12, 11/12, 22/9, 23/9, ITU-R WP11E

Source: Rapporteurs of Q 10/12 and 11/12

Title: Liaison to ITU-T SG9 and ITU-R WP11E

LIAISON STATEMENT

TO: ITU-T Study Group 9, ITU-R WP11E, and ITU-VQEG with comments on VQEG documents
APPROVAL: Study Group 12
FOR: Information and action.
DEADLINE: April 1, 1998
CONTACTS: Ms. Laura Contin Tel: +39 11 2286174
CSELT Fax: +39 11 2286190
Torino Italy Email: laura.contin@cse.lt.it

Arthur Webster Tel: +1 303 497-3567
NTIA/ITS Fax: +1 303 497 5323
Boulder, CO, USA Email: webster@its.blrdoc.gov

Concerning VQEG:

Study Group 12 video experts have met and reviewed the documents submitted to them from the VQEG. We have the following comments to the VQEG and Study Groups 9 and ITU-R WP11E which we believe will enhance the proposed documents

Subjective Test Plan:

1. For the test sequence containing text (p.3): We think that a culturally neutral solution is best, however we leave final decision to the ILSC. Possibilities include a sequence including road signs or a sequence of a weather report. Horizontally moving text such as a stock market ticker tap display can also be used.

2. Concerning transmission errors, we believe that additional transmission errors should be considered beyond ATM errors. As an example, satellite burst errors might be included. We welcome follow on tests which can address the issue of transmission errors in more detail.

3. Concerning data format (p.10), please refer to ITU-R BT-500-7 (Annex) for a standard method of formatting the data. *(3 of Rec 500)*

4. We propose that labs use extra viewers in order to ensure that the required number in each lab is obtained. Performance of post-screening according to BT-500-7 (2.3 of Annex2) is recommended. Also the alternative test design (attached) is unanimously recommended by the Q11/12 ad hoc group. The use of 15 viewers per lab per condition is encouraged over the 12 originally proposed.

Objective Test Plan

1. Page 7, we wish to see SGI (IRIX ver tbd) workstation added to the list of computing platforms (3.2) assuming an independent lab has this platform available.
2. Page 10: add sentence to paragraph 4.5: "Complexity will be based on the mathematical analysis of the algorithmic description as was done for MPEG1 and MPEG2". Delete first sentence of last paragraph of 5 on page 11 beginning "If the metrics..."
3. Page 6, section 2.4: Use ITU-R Recommendation BT 601 4:2:2 component format. Full stop. delete from "as described in" .. to the end of sentence.

Proposed Definitions for video classes and video terms.

Table 2 Attributes of Classes: Change note 4 to be consistent with note 2 in D37r1.

"Note 2: Broadcast systems all have constant, but not necessarily low, latency and constant delay variation. For most broadcast applications latency will be low, say between 50 and 1000 milliseconds. For high quality video conferencing, and conversational types of applications in general, latency should be less than around 150 ms. Delay variations are allowed within the given range but should not lead to perceptual disturbing time warping effects.

add bit rates from P.911 for mm5 and mm6

"MM5a: ~0.2 Mbit/s, MM5b: ~0.05 Mbit/s, MM6: < .05, Limit: 0 fps."

Add audio tables to definitions from P.911 e.g. Tables 3 and 4 from P.AVQ.
(attach audio tables)

Concerning Progress in subjective test methods

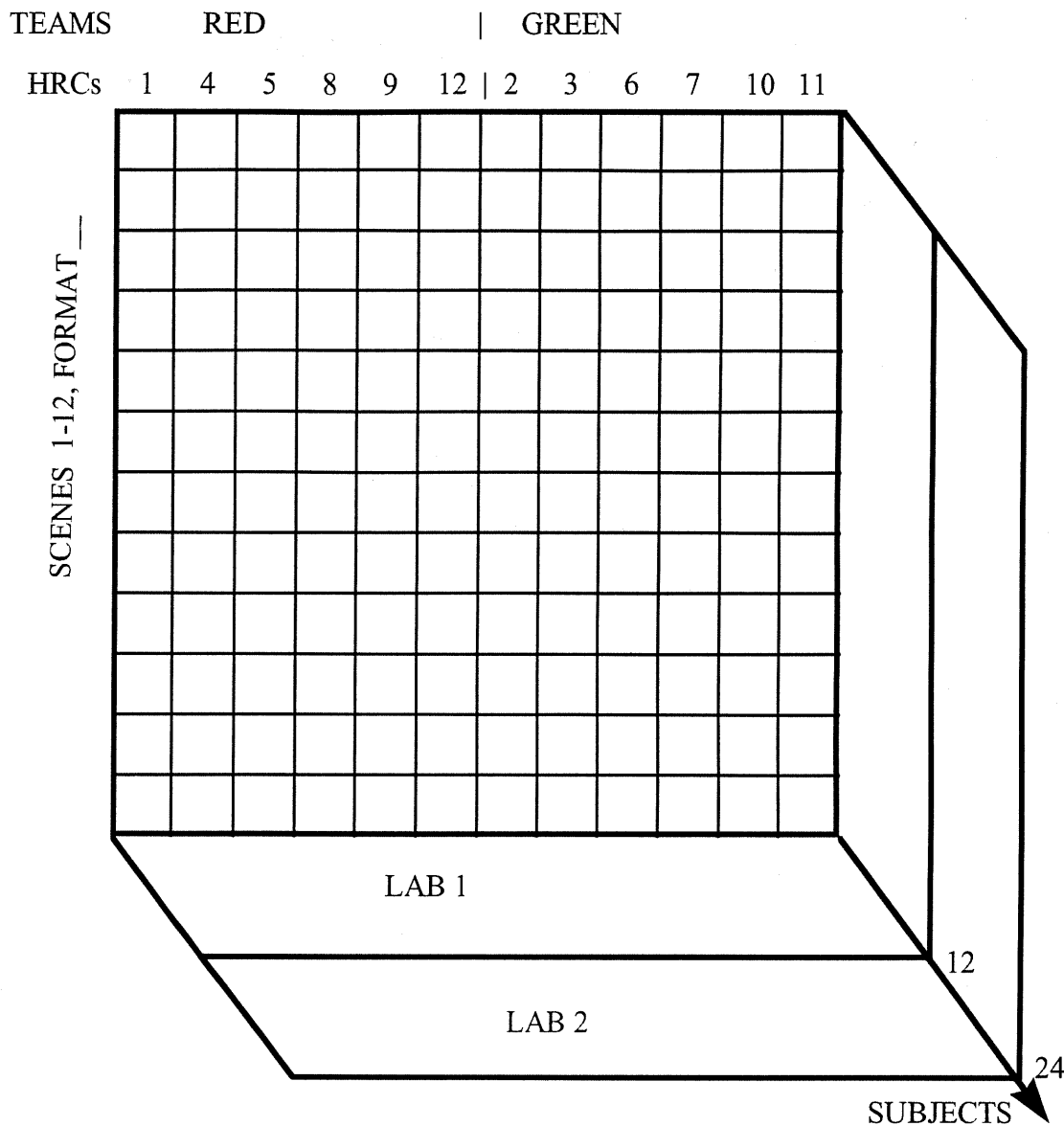
SG 12 wishes to inform WP11E and SG9 that P.AVQ (P.911) has been determined at the last SG12 meeting in Geneva. The text of this Draft New Recommendation is attached.

Attachments: Working Document
P.AVQ

Working Document

ALTERNATIVE ASSIGNMENT OF SCENES to LABS and VIEWERS -
Multiple Team Proposal

The figure below is a test design for one HRC range and one video format (12x12). The viewers at each lab are divided into teams, with each team seeing all scenes combined with a subset of HRCs. The HRCs are numbered by increasing quality, and divided among the Teams. A two-team design could be implemented as follows:



(or some other fair assignment of the HRCs to teams. The main point is to illustrate the team concept in our experiment).

Each Lab randomly assigns viewers to teams as they are enrolled for the test, supplying 12 viewers per team. Each viewer would rate 72 scene-HRC combinations. This is close to the 88 combinations given in section 2.4 of the current subjective test plan. This doesn't violate any of the ANOVA design criteria.

This team-lab design could be repeated in each quadrant of the overall 24x24 HRC-Scene Matrix.

The main advantage of this approach is that any bias among ratings collected at a particular lab is distributed evenly over all combinations.

Another advantage is the wide range of combinations over which the lab-to-lab comparisons may be evaluated.

If HRCs are divided among labs instead of teams (as originally proposed), any lab bias affects a subset of the results. Since the common conditions or HRCs used by all labs would be limited, they may not provide an adequate basis to identify or compensate for lab bias.

Administration of two different teams may be slightly more difficult for the labs, and possibly require production of more test tapes, but I believe the advantages outweigh these issues.

The teams could have overlap (e.g., common HRCs), but this may not be necessary.

The division of combinations among teams could also be along the dimension of scenes, instead of HRCs (if someone offers a reason to prefer this).

(Note that there are many ways to extend this proposal of multi-lab teams. For example, there could be 3 teams dividing the 12x12 matrix, each assigned 4 HRCs. The subjects could be supplied by 3 different labs, with 8 viewers per lab-team and 24 total viewers from each lab. This design reduces the number of combinations for each viewer to 48, and further limits the effect of single lab bias in the overall DMOS. However, some of the key features of this one are not evenly divisible by 6.)

ITU - Telecommunications Standardization Sector
STUDY GROUP 12 (WP 2/12)

Temporary Document No.
Original: English

Geneva, 17-27 February, 1998

Questions: 10/12, 11/12/, 22/9, 23/9

Source: Rapporteurs of Q 10/12 and 11/12

Title: Liaison to ITU-T Study Group 9 with comments on J.ovq and J.svq

LIAISON STATEMENT

TO: Study Group 9

APPROVAL: Study Group 12

FOR: information

DEADLINE:

CONTACTS: Ms. Laura Contin Tel: +39 11 2286174
CSELT Fax: +39 11 2286190
Torino Italy Email: laura.contin@cselt.it

Arthur Webster Te: +1 303 497-3567
NTIA/ITS Fax: +1 303 497 5323
Boulder, CO, USA Email: webster@its.bldrdoc.gov

Study Group 12 thanks Study Group 9 for its liaison and acknowledges that the JRG is formed. We have some comments on the draft New Recommendations attached to your liaison.

Concerning the J.ovq;

Please see the classes and definitions document from the VQEG (attached), for an update on the definitions mentioned in J.ovq. Note that bit rate is used here only to give a nominal bit rate range for the given application class.

It would be useful to wait until the results of the VQEG validation test before finalizing the work on J.ovq.

Concerning J.svq

It is important that a viewer training section be included in the Recommendation. It has been found that this is an important part of subjective testing.

Attachments: D.51 (Classes and Definitions Document from VQEG)

Geneva, 17-27 February, 1998

Questions: 10/12, 11/12, SG16, ITU-R WP11A, WP11E

Source: Rapporteurs of Q 10/12 and 11/12

Title: Reply to two Liaisons from ITU-R WP11A

LIAISON STATEMENT

TO: ITU-R WP11A
COPY TO: ITU-T Study Group 16
APPROVAL: Study Group 12
FOR: Information.
DEADLINE:
CONTACTS: Ms. Laura Contin Tel: +39 11 2286174
 CSELT Fax: +39 11 2286190
 Torino Italy Email: laura.contin@cse.lt

 Arthur Webster Te: +1 303 497-3567
 NTIA/ITS Fax: +1 303 497 5323
 Boulder, CO, USA Email: webster@its.blrdoc.gov

Study Group 12 wishes to thank WP11A for its liaison statements on "Adaptive Image Control in Future TV Systems" and "Assessment and optimization of quality of colour reproduction in television". We are pleased to see this work in progress and will follow your progress with interest. If we find that we can contribute to the work of your questions we will keep you informed. Attached is Draft New Recommendation on P.DEL providing a measurement method for multimedia communication delay, synchronization, and frame rate. We hope that it may be of use in your work.

ITU-T Study Group 16 in its role of lead study group for multimedia might also be interested in the progress of your work.

Attachments: P.DEL (P.931)

Geneva, 17-27 February, 1998

Questions: 10/12, 11/12, 21/13

Source: Rapporteurs of Q 10/12 and 11/12

Title: Reply to Liaison from ITU-T SG13 (WP1/13) and liaison for
ITU-T SG16 and ITU-R SG10

LIAISON STATEMENT

TO: ITU-T SG 13 (WP1/13), ITU-T SG16 and ITU-R SG10

APPROVAL: Study Group 12

FOR: Information

DEADLINE:

CONTACTS: Ms. Laura Contin Tel: +39 11 2286174
CSELT Fax: +39 11 2286190
Torino Italy Email: laura.contin@cse.lt.it

Arthur Webster Te: +1 303 497-3567
NTIA/ITS Fax: +1 303 497 5323
Boulder, CO, USA Email: webster@its.bldrdoc.gov

Please find attached the updated work plans for Q10/12 and 11/12 and
Note that Draft New Recommendations P.AVQ (Subjective audiovisual quality
assessment methods for multimedia applications) and P.DEL (Multimedia communication delay,
synchronization, and frame rate measurement) were determined at this meeting.

Attachments: Workplans of Questions 10 and 11/12

Geneva, 18-25 February 1998

Question: 10/12

SOURCE: Rapporteur for Q.10/12 (Ms Laura CONTIN)

TITLE: Progress Report of Q.10/12

1. Introduction

The text of Question 10/12 "Subjective methods for evaluating audiovisual quality in multimedia services" can be found in COM 12-1. This question is aimed at developing and validating subjective test methods and relative tools to evaluate either video or audiovisual quality in multimedia.

During last study period (1992-1996), SG12 produced three Recommendations on this matter:

- P.910 - "Subjective video quality assessment methods for multimedia applications"
- P.920 - "Interactive test methods for audiovisual communications"
- P.930 - "Principles of a Reference Impairment System for Video"

At last SG12 meeting it was agreed that the main target of this Question is to develop a recommendation on one-way subjective test methods to jointly evaluate audio and video quality. At the same time, the Question will continue to revise the above listed Recommendations, as new knowledge is attained.

2. New documents available

COM 12 - 19 Relation between audio, video and audiovisual quality (KPN Research - NL)

This contribution illustrates a test aimed at evaluating the impact of audio and video quality on the overall audiovisual quality. The experiment was divided in five runs, each of them characterised by a different tasks: 1) listening and evaluating audio only, 2) viewing and evaluating video only, 3) viewing and listening but evaluating audio only, 4) viewing and listening but evaluating video only and finally 5) viewing and listening and evaluating both.

The main conclusion is that there is a significant interaction effect between audio and video quality. Furthermore the video quality dominates the overall perceived audiovisual quality.

COM 12 - 22 Towards a multi-sensory perceptual model (BT)

This contribution provides a proposal for the underlying structure of a multi-sensory perceptual model. Experimental evidence of the interaction between hearing and vision is presented and a proposal made for a multi-sensory model architecture. This architecture is able to take account of the interaction between the individual senses and the nature of the task undertaken.

COM 12 - 45 (FRANCE TELECOM - CNET)

This contribution illustrates a methodological experiment carried out on a videoconference system. Two test methods were used and experimental results compared. The first method was derived from the SSCQE method described in ITU-R Rec. 500-7 and applied to an interactive context. The second method was conforming to ITU-T Rec. P.920. Different audio and video conditions were taken into account.

Del. 38 - Results of an audiovisual desktop video teleconferencing subjective experiment (USA)

This contribution discusses the analysis of an audiovisual desktop video teleconferencing subjective experiment conducted at ITS. It includes: testplan, test results, subjective model of audiovisual quality based upon the individual subjective audio and video scores and a discussion of the results. The results of this experiment have given insight into how audio and video quality relate to audiovisual quality. For this

experiment, video quality was the main component of the overall audiovisual quality. The session ordering effect is also discussed.

Del. 37 - Proposal for draft new Recommendation P.AVQ (CSELT - I)

This contribution provides a preliminary draft of Recommendation P.AVQ, that describes one-way subjective test methods to be used to jointly evaluate audio and video quality. This document should be considered a basis for discussion in the ad hoc group meeting. As in P.910, ACR, DCR and PC methods are proposed.

Del 50 - Examination of agreement between laboratories conducting identical subjective video quality experiments (AT&T - USA)

This contribution describes the lab-to-lab comparison of subjective test data collected as part of the T1A1.5 video performance research project. Each lab recruited at least 10 viewers for each team, permitting the ratings from any lab to be compared with the ratings from two identical experiments conducted at different locations. The method used was the DCR and viewing conditions conformed with ITU-R Rec.500. Low lab-to-lab MOS rms errors and high correlation summarize these comparisons.

Del 59 - Detection and discrimination of blur in edges and lines (AT&T)

This contribution summarizes the results of a psychophysical experiment to determine viewer sensitivity to blur in edges and lines. These results add to our knowledge of the human visual system and are applicable to the study of the assessment of video quality.

TD 55 - Report of the Experts Meeting on Subjective and Objective Video Quality Assessment - Turin, Italy October 1997 (Rapporteur Q.11/12)

This contribution presents the results from the meeting of video quality experts at Turin, Italy in October, 1997. Experts from ITU-T SG12, ITU-T SG9 and ITU-R SG11 participated at that meeting. The main results were the establishment of a group of experts, named Video Quality Expert Group (VQEG), and the definition of its workplan, aimed at validating objective test methods. VQEG intends to make available the outcomes of its work to ITU-T SG12, ITU-T SG9 and ITU-R SG11 to develop Recommendations in this area.

TD 51 - Proposed definitions for Video Classes and Video Terms from the VQEG (Video Quality Expert Group) - (Rapporteur Q.11/12)

This contribution presents a set of definitions that have been drafted by members of the ITU VQEG (Video Quality Expert Group) ad hoc committee on video classes and definitions. These definitions are offered to the participating ITU Study Groups for further review and comment. It is expected that the definitions of terms and video classes will be included in new Draft Recommendations in the area of video quality.

TD 52 - Evaluation of new methods for objective testing of video quality: subjective test plan - (Rapporteur Q.11/12)

This contribution presents a subjective test plan that has been drafted by members of the ITU VQEG ad hoc committee for the subjective test plan. The subjective test plan will be used to evaluate video quality in the bit rate range of 768 kbit/s to 36 Mbit/s. In conjunction with the objective test plan, it will be used to evaluate several proposed methods for objectively assessing video quality by measuring the correlation between subjective and objective assessments. It is expected that this test plan will be included in new Draft Recommendations in the area of video quality, probably as an annex.

TD 53 - Evaluation of new methods for objective testing of video quality: objective test plan - (Rapporteur Q.11/12)

This contribution presents an objective test plan that has been drafted by members of the ITU VQEG ad hoc committee for the objective test plan. The objective test plan will be used to evaluate video quality in the bit rate range of 768 kbit/s to 36 Mbit/s. In conjunction with the subjective test plan, it will be used to evaluate several proposed methods for objectively assessing video quality by measuring the correlation between subjective and objective assessments. It is expected that this test plan will be included in new Draft Recommendations in the area of video quality, probably as an annex.

TD 2 - Liaison statement to ITU-R WPs 11B and 11E, ITU-T SGs 9,12, and 15 and ISO/IEC/CIE on Adaptive Image Quality Control in Future TV Systems. (from ITU-R, WP 11A)

This contribution invites SG12 to examine their possible participation in the studies covered in a new question ITU-R [AM/11A]. Adaptive image control will probably utilize objective video quality parameters and is therefore of possible interest to the work of Q. 11/12.

TD 3 - Liaison statement to ITU-R WPs 11B and 11E, ITU-T SGs 9,12, and 15 and ISO/IEC/CIE on Assessment and Optimization of Quality of Colour Reproduction in Television (from ITU-R, WP 11A)

This contribution invites SG12 to examine their possible participation in the studies covered in a new question ITU-R [AL/11A]. This subject is of possible interest to the work of Q.11/12.

TD 7 - Reply to liaison statement from Study Group 12 concerning Joint Rapporteurs Group on picture quality assessment (from ITU-T SG9)

This contribution acknowledges the formation of the JRG between SG9 and 12. SG9 also provides copy of two draft new Recommendations (J.svq and J.ovq) and solicits comments on them. J.svq deals with subjective picture quality assessment methods for digital cable television applications, while J.ovq deals with objective methods for the assessment of picture quality in digital video systems used in cable television.

TD 8 - Liaison statement from MPEG - Audio Tests

MPEG informs SG12 about the status of MPEG-4 audio coding and provides its preliminary plan for verification test of MPEG-4 error resilience functionality. MPEG is interested in SG12's comments on the test plan for the mentioned verification test.

TD 9 - Liaison from MPEG about MPEG verification tests (SC29/N2181)

MPEG informs SG12 about the status of MPEG-4 visual coding and provides its preliminary plan for verification test of MPEG-4 error resilience functionality. MPEG is interested in SG12's comments on the test plan for the mentioned verification test.

TD10 - Liaison from MPEG about availability of audiovisual sequences (SC29/N2174)

This document is in reply to a previous liaison statement, where SG12 asked for MPEG-2 video library to develop objective models. MPEG will provide the requested material, provided that the sequence copyrights owners allow SG12 to use the relevant material for its purposes.

TD 12 - Reply to RAG TD/124 on coordination activities on potential overlapping between sound and video services questions. (TSAG chairman)

This liaison presents the TSAG's decision (and the rationale behind it) concerning the RAG's proposal to establish a Joint Committee on Sound and Video Services. The TSAG determined that there is no need to establish a Joint Committee to resolve overlap between the ITU-T and the ITU-R. The TSAG goes on to encourage the continuing work of the video quality experts group, which includes members of the ITU-T Study Groups 9 and 12 and the ITU-R WP11E. Any problems concerning overlap of work should be resolved by the Study Group Chairs.

TD 13 - Responsibilities for development of standards for multimedia communications (SG16)

This contribution from SG16 is for information only and outlines how it views its responsibilities as lead study group for the standardization of multimedia communications services and systems

3. Status of Studies

In multimedia services the perceived quality depends on the task to be performed. For example in a free conversation through a videophone the perceived quality primarily depends on delay, lip-synchronization and audio quality, while in a mainly one-way application like remote-teaching the perceived quality could be primarily related to the quality of graph and low motion picture sequences.

Therefore it is important that Q.10/12 will specify assessment criteria and procedures as general as possible in the field of multimedia, but on the other hand it is important to develop test methods tailored on

application-oriented quality requirements. The sections below address some of the possible study items of the Question.

3.1 New evaluation methods

As anticipated in the Introduction, the main goal of Q.10/12 is, at the moment, the development and validation of a methodology for jointly assessing audio and video quality. SG12 has received a number of contributions on this subject. The test results presented in these contributions add to our knowledge in this area and provide further understanding about the relation between audio, video and audiovisual quality. However considerable work is needed to get to the determination of the draft new recommendation P.AVQ.

In addition, contributions on task-based test methods are also solicited. In fact, specific tasks like recognition of objects, recognition of emotions, readability of text and graphics can be suitably applied to evaluate the performance of services like surveillance, videophone and remote teaching respectively. These tasks could be used either on one-way or interactive tests and should provide more meaningful information than the usual global quality evaluation.

Finally, another area that is not completely covered by existing ITU-T Recommendations is the evaluation of error resilience techniques. This area is becoming particularly important for mobile audiovisual communications.

3.2 Updating of existing Recommendations

Recommendation P.910 specifies one-way video test methods, that can be applied to any multimedia application, provided that interaction aspects are not under investigation. Either refinements of the recommended methods or new test methods may be considered and introduced in this recommendation. At this meeting SG12 has received a contribution, that proves the stability of the DCR method. This can be considered a further validation of this method. Similar contributions, proving the stability of either ACR or PC methods, would be highly appreciated.

Recommendation P.920 specifies conversation methods, that are particularly useful to investigate the impact of delay. In this area further studies about tasks and stimuli for conversation are needed.

Recommendation P.930 reflects an early status of research in the specification of reference impairment systems for video. The knowledge attained in this area during last study period was suitable to specify the principles for such reference systems. Further studies are now needed to specify the system itself.

3.3 Test material

A suitable audiovisual library should be set up. A potential starting point for this library is represented by the MPEG library. The rapporteur of Q.10/12 sent a letter to the MPEG sequence copyrights owners asking a formal permission to use these sequences in ITU. Up to now, she received a formal permission only from Microsoft and a number of verbal commitments that should be followed by a formal statement. It is urgent to decide how to manage this library, that is, for example, who will collect and distribute the sequences, how an ITU member can receive the sequences, for what purpose the sequences can be used, etc.

Any further contribution to set up this library is certainly appreciated. In particular sequences with non-English speech and sequences with 'uncopyrighted' music background are needed.

3.4 Test activities

Q.10/12 should continue to support the test activities aimed at validating objective test methods to be included in ITU recommendations, either by contributing to the definition of the test design or by conducting the tests.

3.5 Liaison with other Study Groups

Cooperation with other Study Groups, in particular with SG9, ITU-R SG 11 and ITU-R SG 10 should be encouraged.

4. Summary of discussion during the meeting

An ad hoc meeting was held to discuss technical issues related to Q.10/12. The main focus of this meeting was the methodology for subjective assessment of overall audiovisual quality. A number of contributions on this topic were submitted, among them one was about conversation tests, while the others were dealing with one-way subjective tests and relation between audio, video and audiovisual quality.

The contribution about conversation tests provided the results of an experiment aimed at comparing two different methodologies to evaluate video teleconferencing systems. One method was an extension of the SSCQE method described in ITU-R Rec. BT.500-7, the other one was the method described in ITU-T Rec. P.920. Although the comparison between the results from the two methods is quite difficult, because of the different nature of the methodology, it seems that the test method specified in P.920 is more appropriate in conversation tests. The importance of the task proposed to the subjects was discussed and contributions either about the characterization of the tasks (e.g. duration, level of involvement of subjects, etc.) or about tasks that have been suitably used in experiments were solicited. Some proposals were made, including games and collaborative work in a virtual space.

The question about the importance of the task was discussed also for one-way test method and contributions were solicited.

Concerning the one-way methods to evaluate overall audiovisual quality, maybe the most interesting result is about the relation between audio, video and audiovisual quality when audio and video are synchronized. The results of tests, performed on different kinds of material and different ranges of quality, by using the nine-level scale ACR, showed that variance in audiovisual quality is dominated by video quality. Moreover audiovisual quality can be predicted quite well from the one-way audio and one-way video quality as derived from audio only and video only subjective experiments. The most stable mapping from the separate audio and video quality to the overall audiovisual quality includes only the interaction factor between audio and video quality and the parameter values derived from different experiments are very similar. This can lead to the conclusion that a unique model giving the mapping from the separate audio and video quality to the overall audiovisual quality can be derived at least in the case of one-way tests evaluating synchronized audio and video signals. It was agreed to include this model in an annex of P.AVQ.

Draft P.AVQ were extensively discussed and revised. The new draft of this recommendation is attached. The problem of the management of the ITU audiovisual library was discussed. It was reported that in the case of the ITU video library, sequences are distributed by SMPTE. Thus, also in the case of the audiovisual library, a similar procedure could be used. But at the moment there are no volunteers to distribute the sequences.

Finally, it was agreed to send liaisons to ITU-T SG 9, SG13 and SG16 and ITU-R SG10 and SG11 to inform them about status of the work in Q. 10/12 and Q. 11/12. The text of these liaisons is attached to the rapport of Q.11/12 (TD 24r2). A reply to liaisons from MPEG was also prepared and its text is attached to this report.

5. Final conclusions and updated workplan

P.AVQ is considered mature for determination. Thus, the workplan agreed at the last meeting is revised as follows:

Study item: Joint evaluation of audiovisual quality

Target: Approval of P.AVQ

<u>Milestone</u>	<u>Schedule</u>	<u>Status</u>
Agreement on scope of P.AVQ	<i>Apr. '97</i>	done
Determination of P.AVQ	Feb. '98	done
Decision on P.AVQ	Nov. '98	

Study item: Library of audiovisual sequences

Target: Set up of a library of audiovisual sequences

<u>Milestone</u>	<u>Schedule</u>	<u>Status</u>
First set of audiovisual sequences available	Nov. '98	
Definition of library management procedures	Nov. '98	

Study item: Revision of P.910, P.920 and P.930

Target: Revise and up-to-date ITU-T recommendations about subjective video and audiovisual test method

<u>Milestone</u>	<u>Schedule</u>	<u>Status</u>
Proposals for revisions	Nov. '98	

Annex I

Lists of documents

- *D24 Impact of One Way Delay on Video Conferencing Quality (Bellcore)*
- *D25 Effect of Delay on Multipoint Video Conferencing Systems (Bellcore)*
- *D.019 Performance assessment of audio, video and multimedia quality (USA)*
- *D.021 Objective and subjective measures of MPEG video quality: summary of experimental results (USA)*

Annex II

{P.AVQ (Document TD 32r2) should be attached here}

Annex III

Source: SG12
Title: Liaison to ISO-IEC/JTC1/SC29/WG11 (MPEG)
Approval: Agreed to at SG12
For: Action
Deadline: December '98

Contacts: Mr. Arthur Webster Phone: +1 303 497-3567
NTIA/ITS, 325 Broadway FAX: +1 303 497-5323
Boulder, CO 80303, USA e-mail: webster@its.bldrdoc.gov

SG12 of ITU-T wishes to thank MPEG for their liaisons about MPEG-4 audio and video verification tests and MPEG sequence library.

SG12 is certainly interested in MPEG-4 video verification activities and will appreciate to receive reports of test conducted by MPEG. Moreover, SG12 is wondering if MPEG could provide the test material used in MPEG-4 video verification tests, provided that SG12 obtains a formal permission from the copyrights owners for using MPEG-4 sequences. This material will be used to develop and optimize ITU recommendations about objective quality evaluation methods.

Annex IV

List of Collaborators for Question 10/12

Name	Company	Tel. Fax	E-mail
Yusuf Ansari	Rockwell International Corporation	+1 714 2215062 +1 714 2213667	ansanys@nb.rockwell.com yusuf.ansari@nb.rockwell.com
J. Antkowiak	Deutsche Telekom Berkom	+49 30 34972490 +49 30 34972491	
D.J. Atkinson	U.S. Dept of Commerce NTIA/ITS	+1 303 497-5281 +1 303 497-5323	dj@bldrdoc.gov
Vittorio Baroncini	FUB Italy	+39 6 54802134 +39 6 54804401	Vittorio@fub.it
John Beerends	KPN Research	+31 70 3325644 +31 70 3326477	j.g.beerends@research.kpn.com
Paul Bowman	BT	+44 1473 644105 +44 1473 644649	paul.bowman@bt-sys.bt.co.uk
Noel Chateau	France Telecom		noel@cnet.francetelecom.fr
Laura Contin	CSELT	+39 11 228 6174 +39 11 228 6190	laura.contin@cse.lt.stet.it
Jean-Pierre Evain	EBU	+41 22 717 21 11 +41 22 717 27 10	evain@ebu.ch
Mauro Falcone	FUB Italy		falcone@fub.it
N. Fujino	Fujitsu	+81 44 7543076	nfuji@sig.ts.fujitsu.co.jp

		+81 44 7543514	
David Fibush	Tektronix	+1 503 6276289 +1 503 6271707	davef@tv.tv.tek.com
Andreas Kamcke	Siemens, AG	+49 89 722 45078 +49 89 722 37078	andreas.kamcke@pn.siemens.de
J.Y. Montfort	FT/CNET	+33 2 96053171 +33 2 96051316	montfortj@lannion.cnet.fr
Al Morton	AT&T	+1 908 9492499	acmorton@att.com
R.A. Osorio	Telemig Brazil	+55 31 2294075 +55 31 2291853	ramo@telemig.gev.br
D. Pascal	CNET/France	+33 2 96051578 +33 2 96051316	pascal@lannion.cnet.fr
Franca Pessoa	CPqD-Telebras Brazil	+55 19 705 6746 +55 19 705 6833	franca@cpqd.com.br
A. Takahashi	NTT	+81 422 59 4447 +81 422 59 2330	aki-tak@nttgn.tnl.jp
Yoshihiro Tomita	Fujitsu		ytomita@sig.ts.fujitsu.co.jp
Arthur Webster	US Dept. of Commerce NTIA/ITS	+1 303 4973567 +1 303 4975323	webster@its.bldrdoc.gov