



Assessment of Video

Teresa Hang Nguyen

May 18 2005



Introduction

- Double stimulus continuous quality scale (DSCOS)
 - Context effects occur when subjective ratings are influenced by severity and ordering of impairments within the test session.
- Double stimulus comparison scale (DSCS)
 - Context effects are minimized since viewers are shown pairs of video sequences in a randomized order.
- Single stimulus continuous quality evaluation (SSCQE)
 - Allows viewers to dynamically rate the quality of an arbitrarily long video sequence using a slider mechanism with an associated quality scale.

Find answer to the questions that have been raised.

- Do SSCQE and DSCQS yield inherently different subjective scores?
- Are contextual effects and the lack of an immediate reference picture problematic for SSCQE?
- To what extent does previously viewed video impact the viewer's current opinion score in SSCQE?

Subjective testing

- Methodologies
 - SSCQE, DSCQS and DSCS
- Original subjective experiments
- Meta-experiment

Original subjective experiments

Video sequences for the SSCQE meta-experiment:

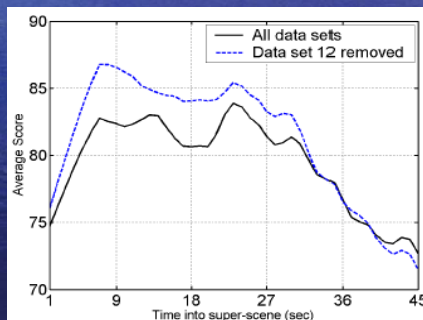
- Television systems
- A particular emphasis on MPEG-2 coding impairments
- Some of the video systems included MPEG-1 coding
- VHS record/playback
- Multiple-generation dubbing with 1/2 inch professional tape recorders
- MPEG-2 bit-streams corrupted with digital errors.

Table 1 Original Subjective Tests

Data Set	Method	Viewers	Video Clips	ITU-R BT.500-
One	DSCS	32	42	BT.500-3
Two	DSCS	32	105	BT.500-3
Three	DSCS	32	112	BT.500-3
Four	DSCQS	67	90	BT.500-8
Six	DSCQS	80	90	BT.500-8
Twelve	SSCQE	32	40	BT.500-10

Secondary subjective experiment

- Performed by NTIA/ using SSCQE with hidden reference removal.
- Video clips were chosen from the six original data sets.
- This subjective meta-experiment was designed to last 30 minutes.
- A panel of 20 viewers was split into two groups, with each group seeing one of the two possible orderings.



Reproducibility of subjective ratings

- The SSCQE method produces a subjective rating every half second.
- The DSCQS and DSCS methods produce one subjective rating for each 8-10 second long video clip.
- One study examined both a 10 second averaging process of the SSCQE ratings and a non-linear averaging process.
- Another study found that the DSCQS subjective rating of a 30 second video clip is different than the average SSCQE subjective rating for those 30 seconds.

References

1. M. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies", <http://www.its.bldrdoc.gov/n3/video/pdf/spie03subj.pdf>
2. ITU-r recommendation BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," Geneva, 2002 (available at www.itu.org).
3. Video Quality Experts Group (VQEG), "RRNR-TV Test Plan" (available at www.vqeg.org).
4. ITU-T COM 9-80-E, "Final report from the video quality experts group (VQEG) on the validation of objective models of video quality assessment," approved Mar. 2000, Ottawa, Canada (available at www.vqeg.org).
5. S. Wolf and M. Pinson, "Video Quality Measurement techniques," NTIA report 02-392, Jun. 2002 (available www.its.bldrdoc.gov/nc/video/documents.htm).
6. R. Hamberg and H. Ridder, "Time-varying Image Quality: Modeling the relation between Instantaneous and Overall Quality," SMPTE Journal, Nov. 1999, p. 802-811.
7. R. Aldridge, D. Hands, D. Pearson and N. Lodge, "Continuous quality assessment of digitally-coded television pictures," IEEE Proceedings online no. 19981843, Oct. 1997.