

Title: Report on the VQEG test
Source: ILSC
Editors: L. Contin, P. Corriveau
Status: Draft

Table of Contents

1. Introduction	1
2. Overview of the work done	2
3. Laboratory Set up for subjective evaluations	3
3.1. NHK Science & Technical Research Laboratories	3
3.2. Communications Research Centre Canada.....	4
3.3. CSELT.....	10
3.4. RAI.....	13
3.5. DCITA Communications LAB.....	14
4. Statistical analysis	18
4.1. Subjective test results	18
4.2. Correlation between subjective and objective data	18
5. Conclusions	18

1. Introduction

This document is intended to summarize the ILSC activities since last VQEG meeting (Gaithersburg, May '98). Due to the short time available, it was not possible to have it approved by the ILSC before sending it to the VQEG. The editors think that it faithfully report the activities of the committee and hope it can be approved by the beginning of the second VQEG meeting.

The mandate of the ILSC included:

1. Selection of source sequences
2. Encoding of sequences
3. Verification of Objective models
4. Subjective testing
5. Statistical analysis

Sec. 2 of this report provides an overview of the work done to benchmark the submitted proposals. Sec. 3 provides detailed information about the set up of subjective test laboratories (at least those received by the 3rd of September). Sec. 4 and Sec. 5 summarize the statistical analysis done, while the results will be given in a separate document prepared by NIST. Finally Sec. 6 reports some of the decisions taken at the ILSC audio-conference of the 31st of August .

2. Overview of the work done

All the tasks assigned to ILSC were completed.

The table below provides information about who did what, by when.

Subjective tests were carried out according to the Subjective test plan, while the objective tests were carried out according to the Objective test plan.

What	Who	When
Selection of sequences	ILSC	May '98
Sending patterns for the normalization to CCETT & CRC	Tektronix	
Adding patterns to the source sequences and sending them on D1 tapes to HRC processing sites	CCETT & CRC	
Choice of coding parameters, encoding the material and sending the HRC processed sequences and 'patterned' source sequences on D1 tapes to Tektronix	IRT, RAI	September '98
Sending the 'normalized' encoded material and source material on D1 tapes to FUB	Tektronix	
Sending the 'normalized' encoded material and source material on Exabyte tapes to the proponents and objective sites	Tektronix	
Sending the 'normalized' encoded material and source material on DAT tapes to some of the proponents and objective sites	NTIA	
Sending objective models to CRC, CSELT and objective test sites	All proponents	August 98
Editing the test tapes and sending them to the subjective test sites	FUB	60 Hz: April 99 50 Hz: July 99
Subjective test	50 Hz format: CSELT, RAI, CCETT, DoCatA 60 Hz format: Berkom, CRC, FUB, NHK	August 99
Calculating Video Quality Rating and sending them to CRC, CSELT and NIST	All proponents	May 99 (most of them)- July 99
Verification of objective models	ATT & NIST(SGI) FUB&CRC (Sun) IRT (PC)	June 99
Statistical analysis of subjective test data	NIST	September 99
Analysis of 'correlation' between VQRs and subjective data	NIST	September 99
Audioconference to discuss available results	NIST, CRC, FUB, CSELT, AT&T, IRT	August 99
Meeting in Leidschendam to discuss results and agree on the final report	VQEG	September 99

3. Laboratory Set up for subjective evaluations

3.1. NHK Science & Technical Research Laboratories

Playing system

Monitor specifications in the operational manual:

Make and model	SONY BVM-2010
CRT size (diagonal size of active area)	482mm (19-inch)
Resolution (TVL)	900 (center, luminance level at 30fL)
Dot-pitch (mm)	0.3mm
Phospor chromaticities (x, y) for R, G, and B	R(0.64, 0.33) G(0.29, 0.60) B(0.15, 0.06) tolerance: +/-0.005

Display setup:

Luminance of the inactive screen (in a normal viewing condition)	0.14 cd/m ²
Maximum obtainable peak luminance (in a dark room, measured after black-level adjustment before or during peak white adjustment)	586 cd/m ²
Luminance of the screen for white level (using PLUGE in a dark room)	74 cd/m ²
Luminance of the screen when displaying only black level (in a dark room)	0 cd/m ²
Luminance of the background behind a monitor (in a normal viewing condition)	9 cd/m ²
Chromaticity of background (in a normal viewing condition)	(0.316, 0.355)

White balance and gamma (using gray scale test pattern in a dark room)

video level	luminance(cd /m ²)	chromaticity (x, y)	Colour temperature
235 (white)			
208			
176	46.6	(0.308, 0.342)	
144			
112			
80			
48	2.1	(0.309, 0.319)	
16 (black)			

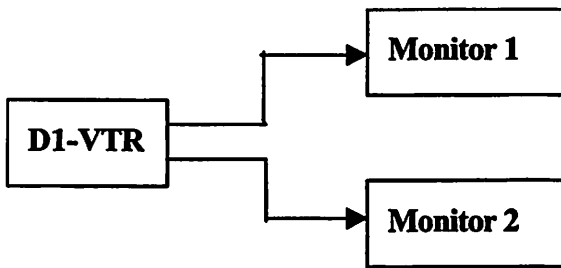
Resolution (using Briggs test pattern in a dark room, report the perceived smallest check-sizes for each luminance level.)

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center	Mid Right	Bottom Left	Bottom Center	Bottom Right
16					0				
48					0				
80					4				
112					4				
144					4				
176					4				
208					3				
235					2				

Note: 0) invisible, 1) 2x2 is visible,, 2) 2x2 and 4x4 are visible, 3) 2x2, 4x4 and 8x8 are visible, 4) 2x2, 4x4, 8x8 and 16x16 are visibl

Video distribution:

Block diagram of distribution system



D1 video out: SDI

Monitor video in: SDI

Data collection: A pair of sliders with 100mm linear scale imitating a DSCQS grading scale.

Viewer characteristics

Number of subjects	17 (HQ), 17 (LQ)
Range of age	22 – 38 (HQ), 22 – 48 (LQ)
Percentage of male	0 (HQ), 0 (LQ)
Comments:	Viewers were employed through a temporary employment agency. All of them were female, but it wasn't intentional.

3.2. Communications Research Centre Canada

PART 1 : Video Monitor Alignment

Two professional grade monitors were used for the subjective evaluations, a Sony BVM-1910 (serial number 2000515) and a Sony BVM-1911 (serial number 2000351).

1.1 Manufacturer's Specifications

Manufacturer's Specifications		
Make and model	Sony BVM-1910	Sony BVM-1911
CRT size (diagonal)	482 mm (19 inch)	482 mm (19 inch)
Resolution (TVL)	>900 TVL (center, at 30fL) ¹	>900 TVL (center, at 103 cd/m ²)
Dot pitch	0.3 mm	0.3 mm
Phosphor chromaticity	R	0.630 , 0.340
	G	0.310 , 0.595
	B	0.155 , 0.070

¹30fL approximately equals 103cd/m²

1.2 Display Setup Characteristics

Display Luminance Setup Characteristics			
Item	BVM-1910	BVM-1911	Notes
Luminance of the inactive screen under ambient light	0.39 cd/m ²	0.33 cd/m ²	In normal viewing conditions with ambient light and lightwall, using non-contact probe (Minolta CS-100 Chroma Meter)
Maximum screen luminance	592 cd/m ²	756 cd/m ²	After black level adjustment, using 700mv white in a 20% center window, using contact

			probe (Minolta CA-100 CRT Color Analyzer)
Screen luminance for peak white level	70.3 cd/m ²	70.2 cd/m ²	Using 700mv white in a 20% center window, using contact probe (CA-100)
Screen luminance for black	0.36 cd/m ²	0.43 cd/m ²	Using a contact probe (CA-100)
Luminance of the monitor surround	10.2 cd/m ²	10.6 cd/m ²	Average of 12 point measurement, using contact probe (CA-100)

Additionally, the monitors were inspected/adjusted for frequency response (9MHz using a 60% sweep), good flat-field purity, convergence, screen geometry, 4:3 aspect ratio, and 5% overscan. No aperture correction was applied.

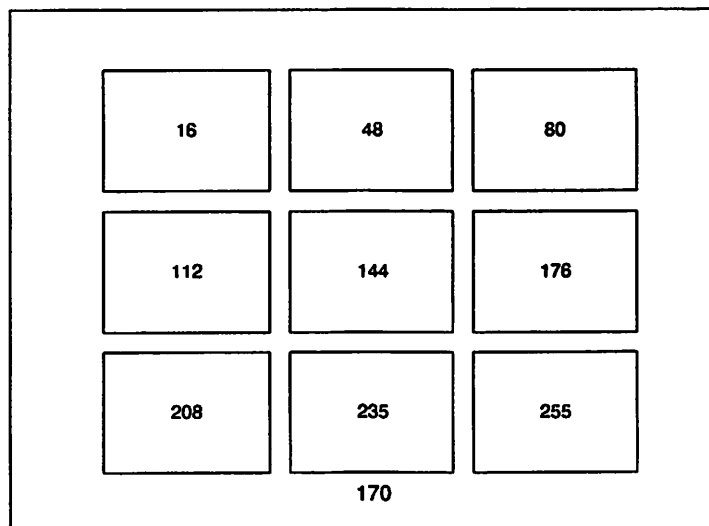
1.3 Phosphor Chromaticities

The next table lists SMPTE/EBU phosphor chromaticities. The measurements were taken using full screen flat-fields, using a contact probe at center screen under dark room conditions.

Chromaticity Verification				
	BVM-1910		BVM-1911	
	x	y	x	Y
RED	0.635	0.335	0.633	0.332
GREEN	0.304	0.602	0.307	0.601
BLUE	0.143	0.058	0.143	0.059

1.4 Grayscale Tracking Measurements

A specialized test pattern was used to characterize the gray-scale tracking. The pattern consisted of nine spatially uniform boxes, each being approximately 1/5th the screen height and 1/5th the screen width. All pixel values within a given box are identical, and all pixel values outside the boxes are set to a count of 170.



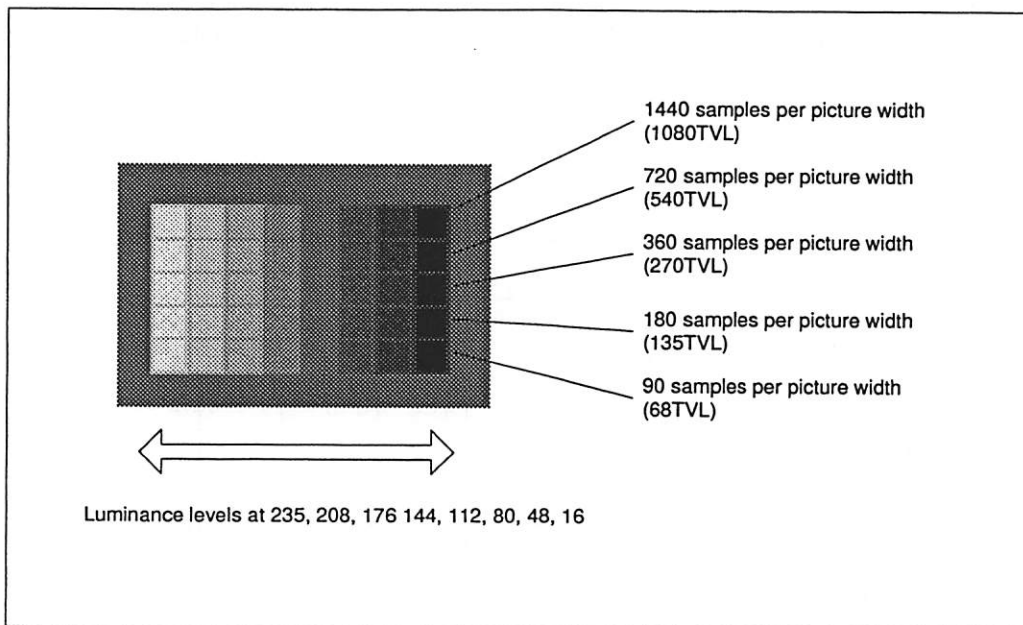
The following measurements were obtained:

Gray Scale Tracking for BVM-1910				
Level count	Degrees K	x	y	cd/m2
16	7600	0.298	0.316	0.62
48	6510	0.313	0.327	1.86
80	6420	0.314	0.328	5.32
112	6630	0.311	0.323	11.6
144	6490	0.313	0.327	21.5
176	6540	0.312	0.325	33.4
208	6830	0.308	0.320	47.5
235	6660	0.311	0.322	65.9
255	6640	0.311	0.322	76.0

Gray Scale Tracking for BVM-1911				
Level count	Degrees K	x	y	cd/m2
16	8500	0.286	0.308	0.67
48	6890	0.306	0.326	1.59
80	6370	0.315	0.326	4.35
112	6690	0.309	0.333	11.5
144	6440	0.313	0.332	20.5
176	6280	0.317	0.329	30.1
208	6860	0.307	0.328	52.9
235	6690	0.310	0.328	71.6
255	6420	0.314	0.327	81.6

1.5 Display Resolution Estimates

To visually estimate the limiting resolution of the displays, a special Briggs test pattern was used. This test pattern is comprised of a 5 row by 8 column grid. Each row contains identical checkerboard patterns at different luminance levels, with different rows containing finer checkerboards. The pattern is repeated at nine different screen locations.



Estimated Resolution in TVLs from visual inspection of the Briggs Pattern for BVM-1910.

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center *	Mid Right	Bottom Left	Bottom Center	Bottom Right
16	0	0	0	0	0	0	0	0	0
48	>540	>540	>540	>540	>540	>540	>540	>270	>270
80	>270	>540	>270	>540	>540	>540	>270	>540	>270
112	>270	>270	>270	>270	>270	>270	>270	>270	>270
144	>270	>270	>270	>270	>270	>270	>270	>270	>270
176	>270	>270	>270	>270	>270	>270	>270	>270	>270
208	>270	>270	>270	>270	>270	>270	>270	>270	>270
235	>135	0	>270	0	>135	0	0	0	0

Estimated Resolution in TVLs from visual inspection of the Briggs Pattern for BVM-1911

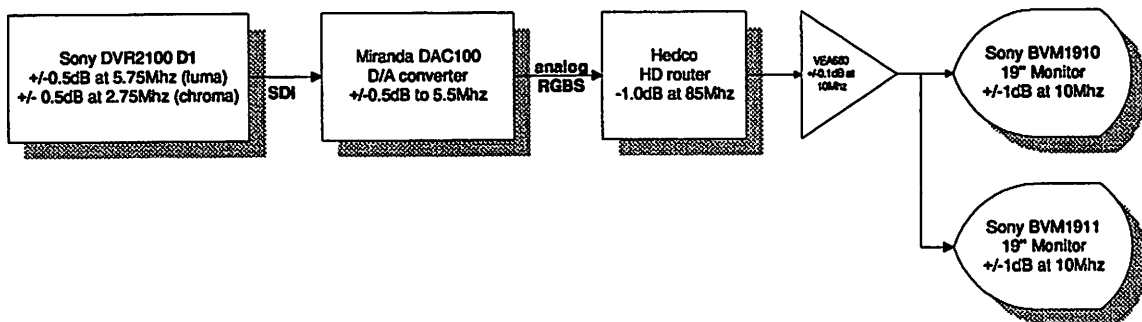
Level	Top Left	Top Center	Top Right	Mid Left	Mid Center *	Mid Right	Bottom Left	Bottom Center	Bottom Right
16	0	0	0	0	0	0	0	0	0
48	>540	>540	>540	>540	>540	>540	>540	>540	>540
80	>540	>540	>270	>270	>540	>540	>540	>540	>540
112	>270	>540	>270	>270	>540	>270	>270	>270	>270
144	>270	>270	>270	>270	>270	>270	>270	>270	>270
176	>270	>270	>270	>270	>270	>270	>270	>270	>270
208	>270	>270	>270	>270	>270	>270	>270	>270	>270
235	0	>270	0	0	>135	0	>135	>135	>270

PART 2: Video Distribution System

2.1 Video Distribution Diagram

The video signal distribution utilized at the ATEL for these subjective test sessions is summarized in the following diagram.

Simplified Distribution Diagram for VQEG Project Playback



2.2 Video Distribution Characterization

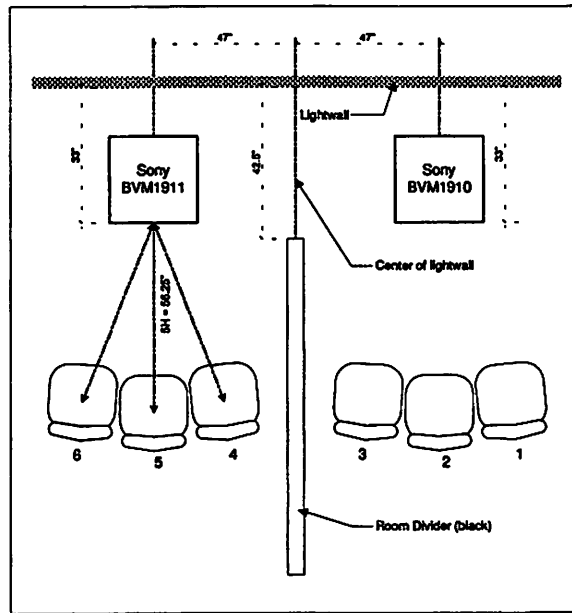
To characterize the video distribution system, a Tektronix TSG1001 test signal generator output was fed to the analog inputs of the Hedco router, using an 1125I/60 signal. A Tektronix 1780WFM was used to obtain measurements at the BVM-1911 input

Characterization of the Distribution System		
Item	Result	Comment
Frequency response	0.5 to 10MHz (+/- 0.1dB)	For each color channel Using fixed frequency horizontal sine wave zoneplates
Interchannel Gain Difference	-2mv on Blue channel -1mv on Red channel	Distributed Green channel as reference Using 2T30 Pulse & Bar and subtractive technique
Nonlinearity	<0.5% worst case on Green channel	Direct output of signal generator as reference (Green channel) Using full amplitude ramp and subtractive technique
Interchannel Timing	Blue channel: 1.75ns delay Red channel: 1.50ns delay	Relative to Green channel output Using HDTV Bowtie pattern

PART 3: The Viewing Environment

The viewer environment is summarized in the following diagram. The ambient light levels were maintained at 6 – 8 lux, and filtered to approximately 6500 degrees Kelvin. The monitor surround was maintained at 10cd/m², also at 6500 degrees. No aural or visual distractions were present during testing.

Theatre Setup for VQEG Tests



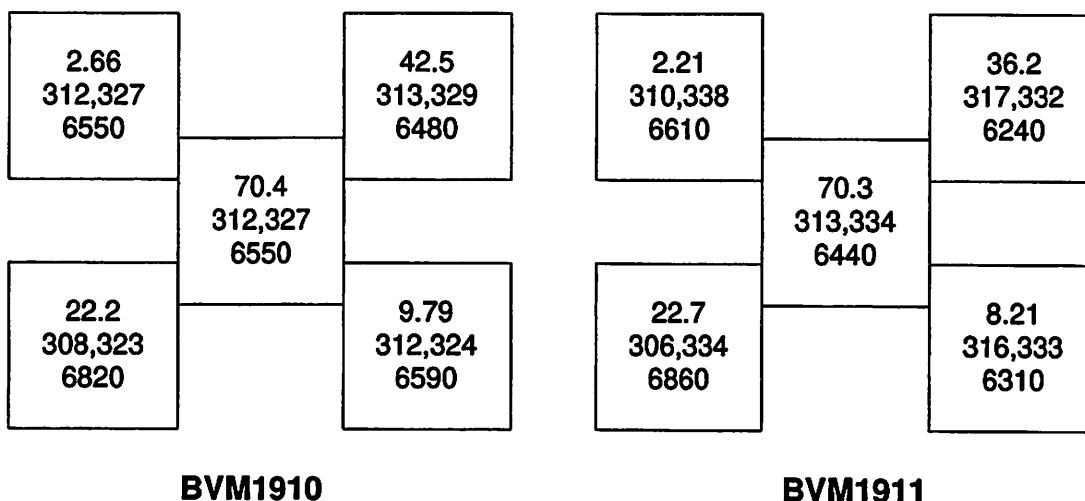
NOTES:
 Monitor control panels and make/model numbers are hidden from view.
 Monitors seated on identical 25" high dollies draped in black cloth.

PART 4 : Monitor Matching

Additional measurements were obtained to ensure adequate color matching of the two monitors used in testing.

Displaying Full Field Colorbars									
	Yellow			Cyan			Green		
Monitor	x	y	Y	x	y	Y	x	y	Y
1910	0.422	0.502	59.8	0.219	0.317	51.8	0.303	0.596	47.6
1911	0.411	0.511	65.7	0.225	0.331	58.2	0.306	0.594	52.6
	Magenta			Red			Blue		
	x	y	Y	x	y	Y	x	y	Y
1910	0.319	0.158	20.8	0.626	0.331	15.3	0.145	0.060	4.66
1911	0.319	0.158	19.2	0.623	0.327	13.6	0.146	0.062	4.04

The following grayscale measurements utilize a 5 box pattern, with luminance values set to 100%, 80%, 60%, 40% and 20%. Each box contains values for luminance in cd/m², x and y co-ordinates, and color temperature in degrees Kelvin.



PART 5 : Schedule of Technical Verification

- Complete monitor alignment and verification is conducted prior to the start of the test program.
- Distribution system verification is performed prior to, and following completion of, the test program.
- Start of test day checks include verification of monitor focus/sharpness, purity, geometry, aspect ratio, black level, peak luminance, grayscale, and optical cleanliness. In addition, the room illumination and monitor surround levels are verified.
- Prior to the start of each test session, monitors are checked for black level, grayscale and convergence. Additionally, the VTR video levels are verified.
- During each test session, the video playback is also carefully monitored for any possible playback anomalies.

3.3. CSELT

Monitor specifications in the operational manual:

Specification	Value
Make and model	SONY BVM20F1E
CRT size (diagonal size of active area)	20"
Resolution (TVL)	900
Dot-pitch (mm)	0.3
Phospor chromaticities (x, y) for R, G, and B	R → 0.640, 0.330 G → 0.290, 0.600 B → 0.150, 0.060
Measuring Equipment:	TEKTRONIX LUMACOLOR2 HEAD J1810 Calibrated 13/1/99

Display setup:

General configuration parameters

Parameter	Value
Aperture correction	NONE
Focus should be adjusted for maximum visibility of high-spatial-frequency information	MULTIBURST O.K.
Geometry adjusted to minimize errors & provide desired overscan. The non-active video region is defined as: - the bottom 14 frame lines - the top 14 frame lines - the right 14 pixels - the left 14 pixels	OVERSCAN SET
Purity (spatial uniformity of white field) should be optimized	
Convergence optimized	Done
Video signal distribution system	SDI DIRECT FROM D-1 TO MONITOR

Monitor setup

Measurement	Value
Luminance of the inactive screen (in a normal viewing condition)	0 [Cd/m ²]
Maximum obtainable peak luminance (in a dark room, measured after black-level adjustment before or during peak white adjustment)	500 [Cd/m ²]
Luminance of the screen for white level (using PLUGE in a dark room)	70 [Cd/m ²]
Luminance of the screen when displaying only black level (in a dark room)	0
Luminance of the background behind a monitor (in a normal viewing condition)	13 [Cd/m ²]
Chromaticity of background (in a normal viewing condition)	6400-6500 °K

White balance and gamma (using gray scale test pattern in a dark room)

Video level	Luminance (cd/m ²)	Chromaticity (x, y)	Color Temperature [°K]
255	85.1	317,316	6350
235 (white)	70.2	314,314	6550
208	52.2	312,312	6800
176	37.3	311,319	6700
144	22.8	307,319	6900
112	12.2	298,317	7750
80	5.18	268,323	9300
48	1.05	Not Measurable	Not Measurable
16 (black)	0	Not Measurable	Not Measurable

Gamma, evaluated by means of linear regression (excluding lowest point): 2.584

Resolution (using Briggs test pattern in a dark room, report the perceived smallest check-sizes for each luminance level.)

The observation results are reported here (1 subject)). For the interpretation of the data refer to the Picture in Sec. 3.2 Part 1.5.

1. Conditions:

ILSC report

- Dark room
- Viewing distance \approx 1H. (center screen)

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center	Mid Right	Bottom Left	Bottom Center	Bottom Right
16	0	0	0	0	0	0	0	0	0
48	540	540	540	540	540	540	540	540	540
80	540	540	540	540	540	270	540	540	540
112	270	540	270	270	540	270	270	270	270
144	270	270	270	135	270	135	135	135	0
176	135	135	135 ^(*)	0	135	0	0	0	270
208	135 ^(*)	0	135 ^(*)	0	0	0	0	0	135 ^(*)
235	0	0	0	0	0	0	0	0	0

^(*) checkerboard is visible only on upper line

Viewer characteristics

Number of subjects	17 (HQ) , 18 (LQ)
Range of age	18-64 (HQ), 17 – 52 (LQ)
Percentage of male	40 (HQ), 55 (LQ)
Comments:	One subject did not show

3.4. RAI

Playing system

Monitor specifications in the operational manual:

Make and model	SONY BVM2010P
CRT size (diagonal size of active area)	20"
Resolution (TVL)	900
Dot-pitch (mm)	0.3
Phospor chromaticities (x, y) for R, G, and B	R(0.64,0.33) G(0.29,0.6) B(0.15,0.06)

Display setup:

Luminance of the inactive screen (in a normal viewing condition)	0.02 cd/m ²
Maximum obtainable peak luminance (in a dark room, measured after black-level adjustment before or during peak white adjustment)	508 cd/m ²
Luminance of the screen for white level (using PLUGE in a dark room)	70.2 cd/m ²
Luminance of the screen when displaying only black level (in a dark room)	0.012 cd/m ²
Luminance of the background behind a monitor (in a normal viewing condition)	3.5 cd/m ²
Chromaticity of background (in a normal viewing condition)	5500 °k

White balance and gamma (using gray scale test pattern in a dark room)

video level	luminance(cd /m ²)	chromaticity (x, y)
235 (white)		
208		
176	32.8	(0.3, 0.332)
144		
112		
80		
48	1.6	(0.309, 0.331)
16 (black)		

Resolution (using Briggs test pattern in a dark room, report the perceived smallest check-sizes for each luminance level.)

The observation results are reported here (2 subject). For the interpretation of the data refer to the Picture in Sec. 3.2 Part 1.5.

2. Conditions:

- Dark room
- Viewing distance ≈ 1H. (center screen)

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center ^(#)	Mid Right	Bottom Left	Bottom Center	Bottom Right
16					0				
48					540				
80					540				

112					540				
144					540				
176					540				
208					270				
235					270				

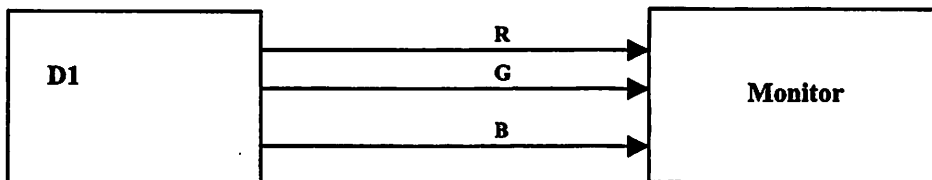
• Viewing distance $\approx 4H$. (center screen)

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center ^(#)	Mid Right	Bottom Left	Bottom Center	Bottom Right
16					0				
48					270				
80					270				
112					270				
144					270				
176					270				
208					270				
235					135				

• Viewing distance $\approx 5H$. (center screen)

Level	Top Left	Top Center	Top Right	Mid Left	Mid Center ^(#)	Mid Right	Bottom Left	Bottom Center	Bottom Right
16					0				
48					135				
80					135				
112					135				
144					135				
176					135				
208					135				
235					135				

Video distribution



3.5. DCITA Communications LAB

Monitor specifications in the operational manual:

Specification	Value
Make and model	SONY BVM2010PD
CRT size (diagonal size of active area)	19"
Resolution (TVL)	900
Dot-pitch (mm)	0.3

Phosphor chromaticities (x, y) for R, G, and B	R → 0.640, 0.330 G → 0.290, 0.600 B → 0.150, 0.060
Measuring Equipment:	MINOLTA TV-2150 TV Colour Analyser II

Display setup:

General configuration parameters

Parameter	Value
Aperture correction	NONE
Focus should be adjusted for maximum visibility of high-spatial-frequency information	MULTIBURST OK
Geometry adjusted to minimise errors & provide desired overscan. The non-active video region is defined as: - the bottom 14 frame lines - the top 14 frame lines - the right 14 pixels - the left 14 pixels	OVERSCAN SET
Purity (spatial uniformity of white field) should be optimised	Checked
Convergence optimised	Done
Video signal distribution system	Parallel 601 DIRECT FROM D-1 TO DDA TO MONITOR

Monitor setup

Measurement	Value
Luminance of the inactive screen (in a normal viewing condition)	0 [Cd/m ²]
Maximum obtainable peak luminance (in a dark room, measured after black-level adjustment before or during peak white adjustment)	165 [Cd/m ²]
Luminance of the screen for white level (using PLUGE in a dark room)	70.2 [Cd/m ²]
Luminance of the screen when displaying only black level (in a dark room)	0.2-0.4 [Cd/m ²]
Luminance of the background behind a monitor (in a normal viewing condition)	9.8 [Cd/m ²]
Chromaticity of background (in a normal viewing condition)	6500 °K

White balance and gamma (using grey scale test pattern in a dark room)

Video level	Luminance (cd/m ²)	Chromaticity (x, y)	Colour Temperature [°K]
255	79.4	316,327	6900
235 (white)	70.2	312,328	6800
208	49.0	312,328	6550
176	33.7	308,325	6450
144	22.3	311,327	6900
112	11.7	313,325	6900
80	6.3	313,333	6350
48	2.7	290,321	6350
16 (black)	1.2	307,302	Not Measurable

Gamma evaluated by means of linear regression (excluding lowest point): 2.076

Resolution (using Briggs test pattern in a dark room, report the perceived smallest check-sizes for each luminance level.)

The observation results are reported here (1 subject)). For the interpretation of the data refer to the Picture in Sec. 3.2 Part 1.5.

Conditions:

- Dark room
- Viewing distance \approx 1H. (centre screen)

Level	Top Left	Top Centre	Top Right	Mid Left	^(#) Mid Centre	Mid Right	Lower Left	Lower Centre	Lower Right
16	540H	540H	540H	540H	540H	540H	540H	540H	540H
48	540H	540H	540H	540H	540H	540H	540H	540H	540H
80	540H	540H	540H	540H	540H	540H	540H	540H	540H
112	540H	540H	540H	540H	540H	540H	540H	540H	540H
144	540H	540H	540H	270	540H	540H	540H	540H	540H
176	270	270	270	270	540H	270	270	540H	270
208	270	270	270	270	270	270	270	540H	270
235	270	270	270	270	270	135	270	270	270

^(#) Mandatory

540H means horizontal pattern only at 540 resolution, in all these cases a full checkerboard is visible at 270 resolution in both H & V

Video distribution:

Parallel Rec-601 direct from Sony DVR-1000 D-1 machine to Abacus Digital Distribution Amplifier then directly connected to monitor via Parallel Rec-601 (27 MHz 8 Bits) 110 ohm twisted pair shielded cable (length 25 Metres).

4. Statistical analysis

NIST collected all the raw data from subjective test laboratories and proponents and produced a common dataset that was distributed through the VQEG reflector.

NIST produced the official report on the statistical analysis of both subjective test results and correlation between subjective data and objective evaluations.

4.1. Subjective test results

The statistical analysis of the subjective test result was performed according to the following steps:

1. Examination of the distribution of Differential Opinion Scores (DOS) grouped by sequence and HRC.
2. Exclusion of outliers
3. Calculation of Mean DOS (MDOS) and CI by picture format and HRC
4. Examination of selected factor effects (i.e. laboratory, quality, picture format, gender and age of viewers).
5. Correlation between test laboratories

4.2. Correlation between subjective and objective data

The statistical analysis about the evaluation of the performance of proposed models was done according to the following steps:

1. Scatter plot of DMOS versus Video Quality Rates (VQR)
2. Pearson correlation by quality and linear fit of DMOS versus VQR
3. Spearman correlation and linear fit of ranked DMOS versus ranked VQR
4. Monotonic Cubic fit

5. Conclusions

Preliminary results were discussed by some of the ILSC members during an audioconference the 31st of August. A number of additional analysis were identified, including:

1. the exclusion of outliers in subjective test results according to the criterion specified in ITU-R Rec. BT.500-8.
2. Metric 1 and Metric 4 specified in the Objective Test Plan Sec. 4.3.
3. Logistic fitting as specified in the Objective Test Plan Sec. 4.2

Due to the lack of time, it will not be possible to conclude the statistical analysis *before* the VQEG meeting. However it was proposed to conclude the analysis *during* the meeting, to get the formal approval of the VQEG on the results of this work.

From the data available, it seems that especially for the high bit rate the performance of the models is not satisfactory and a few of them showed a similar performance. Therefore a number of ILSC members agrees that further work is needed at least in the range of the high quality.