EBU

OPERATING EUROVISION AND EURORADIO

TR 047

TESTING HDR PICTURE MONITORS

TECHNICAL REPORT

Geneva

January 2019

Contents

Testing HDR Picture Monitors

EBU Committee	First Issued	Revised	Re-issued
TC	2019		

Keywords: HDR, High Dynamic Range, Reference picture monitor, test

1. Introduction

In April 2018 the Swiss EBU Member, RTS, hosted a group of experts from the BBC, CBC, EBU, IRT, NRK, RAI and ZDF to perform a series of tests on 6 professional, first generation HDR-capable video monitors.

This report summarises a number of key findings from the measurements, both in terms of overall monitor performance and on what can be improved in terms of the measurement methodology used.

2. Scope

The main goals of the testing exercise were to verify the parameters in EBU Tech 3320 [1] and the test methodology described in EBU Tech 3325 [2]. The tests also provided an impression of how well first generation high-end HDR picture monitors perform.

It must be emphasized that the tests were NOT planned to be an exhaustive validation of the HDR picture monitors according to EBU Tech 3320, but rather to 'test the testing'.

3. Test set-up

3.1 Overview

The tests took place in the RTS 'Hackerspace', which includes the "Fred Neumann" small production studio. To guarantee a suitably dark testing environment, black stage curtains were used to create a 'black box'. In this box one display at a time could be measured.

Non-light critical measurements were performed in the normal studio environment, where the monitors were arranged on tables. For most tests outside the black box, the measuring equipment was moved to the display under test, instead of the inverse.

Some specifics of the measurement set-up used are explained below. It should be noted that for this exercise the tests were performed without a frustum, which normally should be used for several of the tests.

Figure 1: Impression of the studio set-up showing five displays outside the 'black box'.
Mid-grey backdrop curtains were lit by D65 illumination to assist the subjective quality evaluation.

3.2 Spectroradiometer alignment

A *Leica 3D Disto* laser measurement and 3D mapping device was used to provide extremely precise measurements in the optical and colorimetric field. Special software for the 3D Disto that was developed by RAI R&D made it easy to achieve the necessarily accurate positioning of the optic axis of the spectroradiometer; guaranteeing a given angle of the radiometer relative to the centre of the screen plane of the display under test.

Leica 3D Disto

Picture monitor under test

CS 2000 spectroradiometer on wheeled tripod with led laser to measure height

Figure 2: Main components of the test set-up

The system is simpler to use than the goniometer systems that are normally used in optical benches for colorimetric display measurements.

3.2.1 Accuracy

The *Leica 3D Disto* is very accurate; the Tie Accuracy (combined distance and angle measurement) is approximately 1mm @ 10m. The angle accuracy is 5" (1.2mm @ 50m). Using the RAI set-up, the Minolta CS 2000 spot spectroradiometer could be positioned at any angle between ±60° relative to the plane of the screen.

3.2.2 Alignment procedure

1. The 3D Disto is freely positioned on a tripod and after switch-on it automatically adjusts itself parallel to the surface of the floor.

2. Using the specially developed RAI "Pointing Software" (running on a USB or WiFi connected PC) the 3D Disto is driven so that its laser light hits a series of points on the display under test: top right (A), bottom right (B) and bottom left (C).

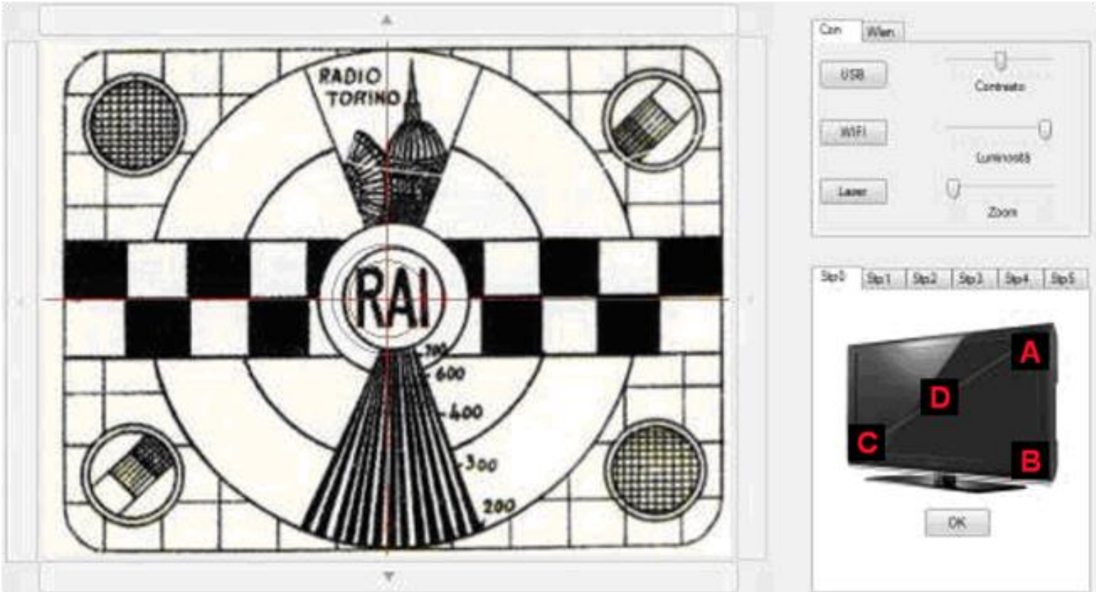


Figure 3: Pointing Software

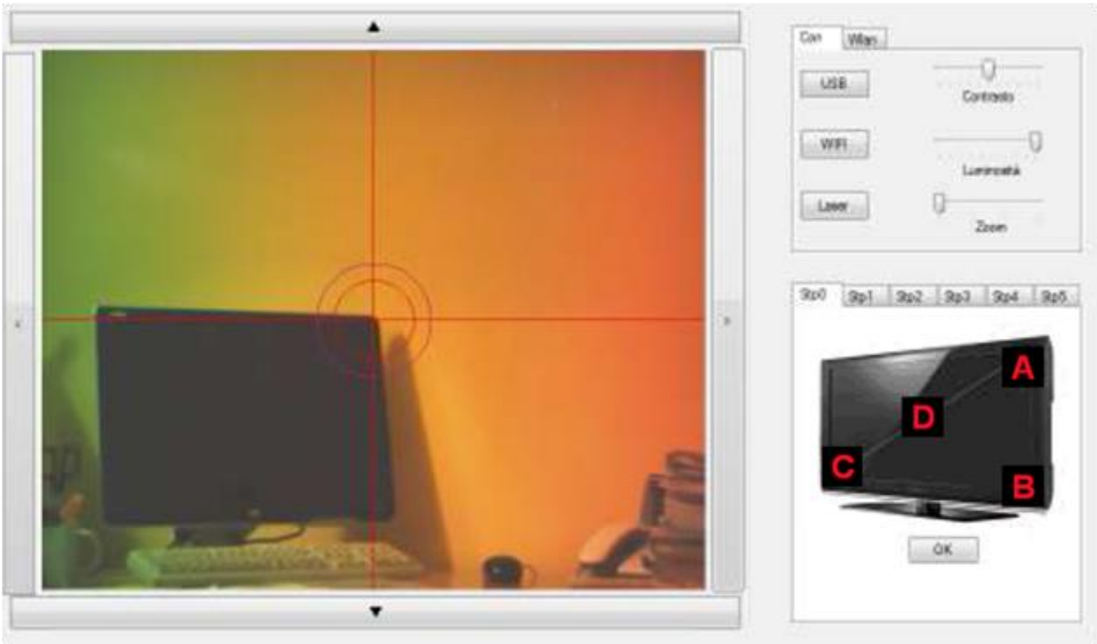


Figure 4: The laser measurement device has a small camera so pointing the beam is easily achieved

3. The laser is pointed to the middle of the display (point D), which should be showing an appropriate centre-marked signal.

4. The 3D Disto measures the height above the floor of point D to a point on the floor (point E) below the screen.

5. The 3D Disto can now calculate and mark a reference point (point F) on the floor where the spectroradiometer should be placed. This point is at exactly 3.5 times picture height away from the centre of the screen.

6. The Minolta CS 2000 analyzer is positioned above point F at a height such that it points directly at point D.The led laser of the Spectroradiometer is used to measure the height

3.3 Delay

To measure the audio-video delay, a 1080p/50 video sequence consisting of a repetition of 1 single white frame followed by 49 black frames (repeated during 2 seconds) was used.

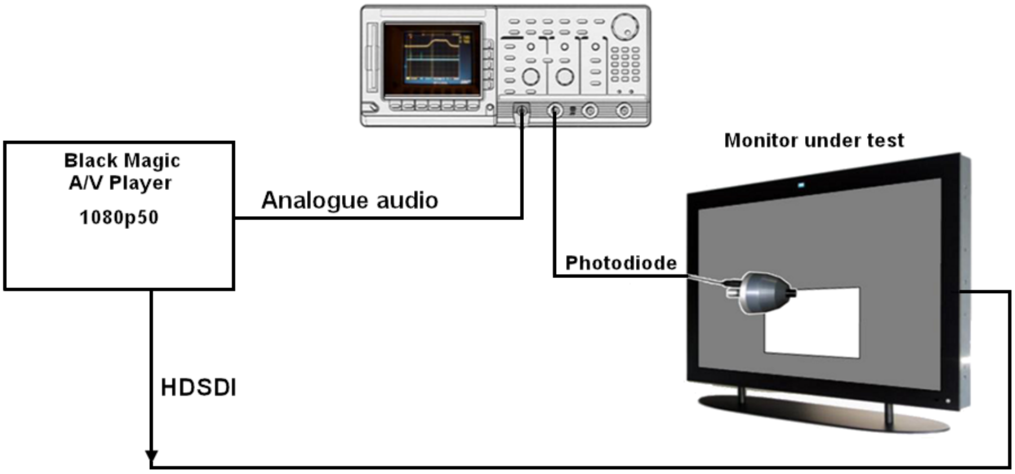


Figure 5: Audio-Video delay measurement

The measurement of the A/V delay of the monitor under test was made using a Black Magic Hyperdeck player as a source. As a test signal, a 1080p50 HD video sequence consisting of a full white frame with a 100 Hz sine wave followed by 49 mute black frames, repeated for a total period of 2 seconds, was used. The photo-diode was connected to a 9V battery via a pull-up resistor of 120 kΩ.

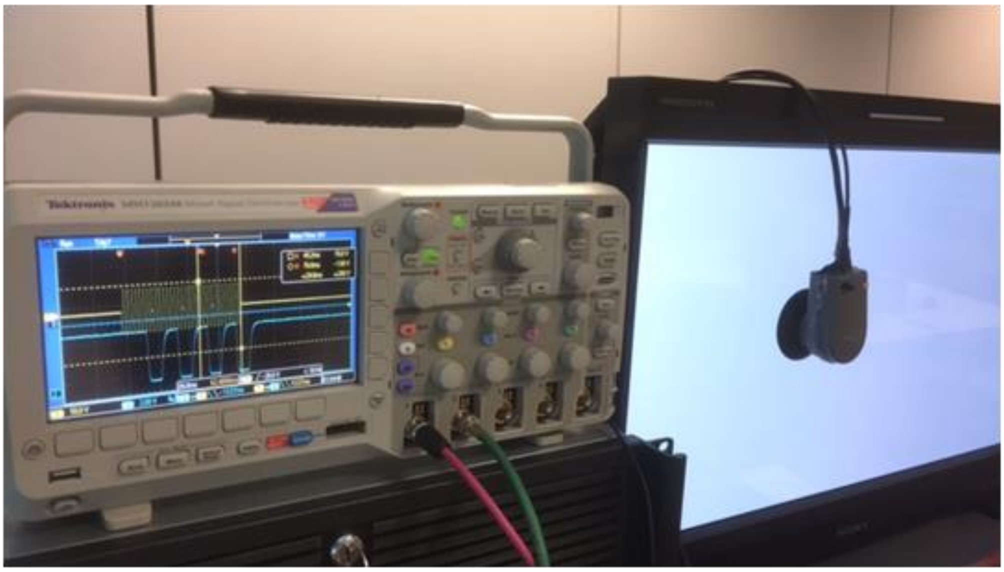


Figure 6: An A/V delay measurement in progress

The analogue audio output of the Hyperdeck was connected to the first channel of the oscilloscope (a TektronixmsO2024B) and the photodiode output was fed to the second input of the oscilloscope.

The overall audio-video delay is seen as a time difference between the rising flanks of the output of the photo diode and the tone burst between the two oscilloscope channels.

3.4 Uniformity

The small area uniformity (SAU) test [4] has been shown to provide similar results to the large area uniformity test (LAU) described in EBU Tech 3325 v1 [2] in a fraction of the time. The SAU test uses a commercial digital single-lens reflex (DSLR) camera and lens as the array Light Meter Device, providing a reduced-cost alternative to a scientific CCD sensor or other specialized equipment.

The camera is positioned in front of the display at a distance of 4H (4 times the screen height). This provides an angle of view which is representative of an average viewing situation, while not being too close to minimize the effect of angular dependency in the case of some display technologies.

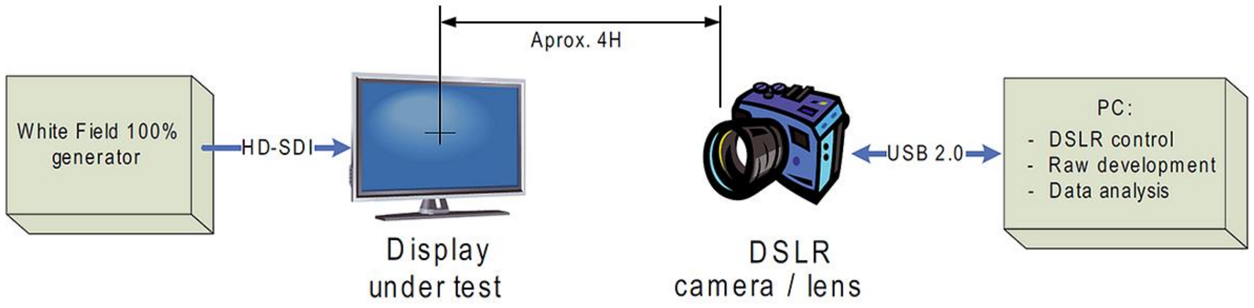


Figure 7: Small Area Uniformity test

A 100% white field test pattern is provided to the display under test (DUT). The brightness of the DUT is adjusted to obtain a certain reference level of operation – e.g. 80 cd/m^2 – at its centre. The room must be dark with no reflective surfaces.

The picture is captured as a raw file and transferred to a computer for raw development and data analysis as described in [4].

4. Main testing results

4.1 Summary of measurements

The tests were principally intended to give experience of measuring HDR monitors and to help the EBU Video Systems group to revise the EBU Tech 3320 and Tech 3325 publications, but they also provided an opportunity to understand the developing state of the art in HDR UHD monitors.

The manufacturers involved took part in the testing in a collaborative fashion; it was about exploring how to test monitors rather than saying which was the best, about seeing how close they were to meeting our (evolving) requirements, rather than confirming whether monitors met a particular specification.

Because the principal interest of the EBU members is in using monitors for general TV production, including live events and studio shows, and because we did not have experience of PQ, the HDR tests were undertaken using HLG.

In addition, it was clear that we would need to create some new test patterns, both for the ITU-R BT.2020 / BT.2100 colour space but also for the HLG signal format. However, at the time of these tests, most available test signals were those defined for SDR, which were subsequently adapted to enable the tests to be undertaken.

4.1.1 Luminance Output

The measured (1% patch) luminance output (for 100% HLG / code level 940) of the monitors ranged from 650 - 1268 cd/m^2, with one low-end HDR monitor achieving 225 cd/m^2.

4.1.2 Black Level

The measured luminance output (for 0% HLG / code level 64) of the monitors ranged from 0.022 - 0.0004 cd/m^2, with the low-end HDR monitor achieving 0.37 cd/m^2.

4.1.3 Colour Gamut

All the displays met the Grade 1B/2 requirement for the Red primary, but not 1A. Two monitors just met the Grade 1A requirement for Blue. The biggest challenge is to meet the Green target (none even met the Grade 1B/2 requirement).

4.1.4 Contrast Ratio

The simultaneous contrast ratio test pattern has not yet been redesigned for HDR, so the contrast measurements taken were the Full Screen (1% patch) measurements.

The contrast ratio between the measured 100% peak (HLG level 940) 1% white patch and an entirely black (level 64) screen varied between 34,000:1 and 2,600,000:1 on the full HDR displays. At least some of those LCD-based displays using dimmable back lights we presume had their backlights effectively "off" during the black measurement. We do not believe this measurement to be particularly meaningful in its present form.

4.1.5 Transfer Function Characteristics

Two displays met the required transfer function tolerance for 1000 cd/m^2 peak brightness, and one for a higher brightness.

4.1.6 Grey Scale Reproduction

The White Reference should ideally have been the 75% HLG signal level defined in ITU-R BT.2408. However, since the original grey-scale test patterns were at 10% intervals, there was no 75% test signal available. The white points have therefore only been analysed at 100% HLG signal, and the grey-scale tracking calculated relative to this.

Only one of the monitors met the Grade 1 HDR requirement.

4.1.7 Colour Reproduction

Four of the six monitors fully met the requirement for test colour reproduction, three of them being exceptionally good.

4.1.8 White Point and Colour Temperature

The white points of the monitors, as provided to be tested, were measured to have Δu*v* errors, compared to D65 (for 100% signal level) of between 3.2 and 9.3, thus none were within the Grade 1 specification, although two met that for Grade 2.

The measured colour temperatures, in the region of 75% HLG signal level, ranged from 6200K to 7100K, with three of the monitors falling within the range 6400K to 6700K

4.1.9 Viewing Angle Dependency

Due to limited facilities, we only tested the viewing angle at ±45 degrees in the horizontal direction, and nothing in the vertical plane. Colour reproduction with viewing angle was not tested, but the grey-scale, white and black levels were.

The Grade 1 and 2 requirements (for contrast to drop by less than 50%) were easily met by the two OLED monitors, and just about met by three of the four LCD monitors.

None of the monitors met the white point deviation spec with viewing angle, and in terms of grey-scale tracking with viewing angle, all but one met the requirement at 50% signal level and all met it at 20%.

4.1.10 Motion Artefacts

Not assessed, since no synthetic sequence for UHD was available.

4.1.11 Screen Resolution

Some monitors were RGBW-based technology. How resolution is measured has yet to be decided.

4.1.12 Image Scaling, De-interlacing and Overscan

One of the monitors included an interlacing (CRT emulation) mode.

4.1.13 Delay Time

Measurements were made of the video processing and presentation delay for a variety of modes. Whilst most of the monitors exhibited delays of between 60 and 90ms, one had a delay below 10ms in SDR modes, and below 25ms for all modes, and another showed delays in the range 30-40ms.

4.1.14 Screen Size

One monitor was 24 inch diagonal, one was 55 inch, and the others were in the range of 30 - 31 inch.

4.1.15 Uniformity

The small area uniformity measure, currently planned to be included in the next revision of EBU Tech 3325, was used, and all monitors achieved above 97% uniformity.

4.1.16 Mura (Imperfections in LCD Panels)

Mura was detected on one display, caused by the bezel.

4.1.17 Streaking (Also Known as Crosstalk, Overspill or Shadowing)

The test results indicate that this is an area for further study, since no monitors met our current specification, although the testing arrangement was not ideal.

4.1.18 Stability and Environmental Conditions

This aspect was not tested.

4.1.19 Pixel Defects

Several dead pixels were found in one OLED display.

4.1.20 Ringing and Handling of Under- and Over-shoots

Not measured.

4.1.21 Treatment of Illegal Signals

Not measured.

4.1.22 Image Sticking (Long-term After-image)

Burn-in was experienced on several displays after applying a test signal for several minutes. On multiple displays the burn-in was of a moving type!

The burn-in either disappeared over time or after going through a cleaning mode.

4.1.23 Supported Standards, and Signal Interfaces

None of the monitors supported the higher frame rates (100/120 fps).

Four of the six monitors supported 12G-SDI

4.1.24 Other Facilities

Some of the monitors included most of the operational features which would be expected in a broadcasting monitor.

4.1.25 Acoustic Noise

Not measured.

4.1.26 Surface Reflectivity (Glare)

Not measured.

4.1.27 Power consumption

The average power consumption (measured using a contrast test pattern, with the monitor in HLG mode), was between 118 W and 370 W.

4.1.28 Additional remarks

Here are some additional observations made during the tests:

- Dual-cell LCD displays can show a 'double image' effect in high contrast image, when viewed from (non-perpendicular) angles.
- Viewing LCD displays from an angle (e.g. a horizontal angle of 45° or more) results in poorly contrasted images; effectively removing HDR perception.
- Displays with the same panel can show vastly different characteristics, which underlines the importance of the image processing circuits in video monitors.
- All monitors showed zone plate aliasing in HDR mode, but not in ITU-R BT.2020 SDR mode.
- All models showed banding on the luminance ramp on the EBU test pattern in HDR mode
- Super whites and sub-blacks were clipped by one monitor and with another the sub-blacks were only available in SDR mode.
- All models under test did not overscan.
- Blue seemed noisy on all displays.

5. Conclusions

5.1 Current state of HDR monitors

UHD and HDR monitors are in their infancy, and none of those seen so far to quite meet the EBU's published specification, although progress towards that is most encouraging.

Attaining the ITU-R BT.2020 / BT.2100 wide colour primaries was always going to be a challenge, and the measurements bear that out, with no monitor yet even falling within the Grade 1B tolerance boxes, which are considered as interim (i.e. relaxed) primaries, as a first step towards the full Grade 1A specification. However, it is commendable and most encouraging that several of the monitors had excellent grey-scale tracking and test colour reproduction.

That three significantly different technologies (OLED, LCD with local backlight dimming and dual-layer LCD) could each provide high quality monitoring is also most encouraging.

The observed "moving burn in" was one feature we were not expecting to see, which we expect manufacturers will look into, as is the inaccuracy in the absolute white point.

A further matter that we hope will be noted by the relevant manufacturers is the requirement not to clip sub-black and super-white signals, such as in the ITU-R BT.814 PLUGE signal, when displaying video-range signals.

5.2 Suggestions for related specifications and test material

The test results suggest the following modifications of EBU Tech 3320, EBU Tech 3325 and related test patterns. Numbers in brackets indicate paragraphs in the related publication.

Note that at the time of writing EBU Tech 3320 and EBU Tech 3325 are already being reviewed for updating, the points noted below are not exhaustive.

5.2.1 EBU Tech 3320

- (1.5.8) The viewing angle dependency is defined on the contrast ratio. Note that other bodies (e.g. the ICDM) are investigating how to test this. In the tests, it was not the simultaneous contrast which was measured at different angles, but the luminance. The luminance was observed to often drop off near the blacks. It should be considered if the contrast ratio should be replaced by a different measurement and if the drop in luminance should be added.
- (2.3.3) The tests showed current monitors do not meet the colour primary tolerances requirements, especially for the green primary. This should be considered when reviewing EBU Tech 3320.
- (2.3.6) The text on grey scale reproduction currently says deviation from grey should not be visible for luminance below 1 cd/m^2. This is hard to measure (takes a very long time; requires an absolutely dark environment and very expensive equipment). It also is unrealistic in the sense that monitors are never used in such dark environments. It should be considered to relax this requirement.
- (2.3.7) It should be considered to specify the colour reproduction twice: once for reference white (D65), as is currently the case, and once relative to the *measured white point*. The idea behind the latter is that the human eye is more sensitive to colours which are off relative to the actual white point. Note that both measurements are relevant in different ways:
  - the absolute reference is most important for comparisons between different monitors;
  - the relative reference is important for comparisons between the same monitors.
- (2.3.7) It should be clarified that the reference white is 75%, by referencing ITU-R BT.2408.
- (2.3.7) It should be considered to measure the colour reproduction not only at 75%, but at 100% as well. This will provide additional points on the graph and it will confirm the colour production in highlights (e.g. making sure there is no clipping).
- The value for PQ reference white (58%) should be added.
- (2.3.5) It should be considered to add a transfer function tolerance for PQ. Question: how?
- A glossary should be added, including the various terms which are often (mis)used in this area, e.g. luma, luminance, signal level, luma level, peak luminance, headroom, footroom, super white, super black, reference white, …

5.2.2 EBU Tech 3325

- The colour tests should test the 75% and the 100% colour reproduction as well. The reason for including this requirement is that it provides extra points on the graph and confirms the colour reproduction in highlights; thus making sure the monitor is not clipping in highlights.
- More colours lying outside of the ITU-R BT.709 colour triangle should be defined for testing.
- It should be considered to perform HDR measurements with a 1:2 time interleaved mid-grey image to avoid burn-in (mitigation) effects. For SDR a time ratio of 1:5 may be used.
- The simultaneous contrast measurement test pattern should be revised to include HDR.
- Add the small area uniformity measure.

- To measure the transfer function new patterns are needed with more detailed steps around cv 64 and cv 940.
- A description of the crosstalk measurement should be added.
- It should be considered how to measure ‘dynamic burn-in’ (without damaging the displays…).
- A glossary should be added, including the various terms which are often (mis)used in this are, e.g. luma, luminance, signal level, luma level, peak luminance, headroom, footroom, super white, super black, reference white, …

5.2.3 Test patterns

- The simultaneous contrast measurement test pattern should be revised to include HDR
- The colour test patterns should be extended (more colours, outside of ITU-R BT.709).
- The grey scale patterns for transfer function testing should be extended.
- Test patterns for PQ should be included.

6. Acknowledgements

The EBU gratefully acknowledges the contribution of the participating EBU Member organizations and monitor vendors/dealers.

7. References

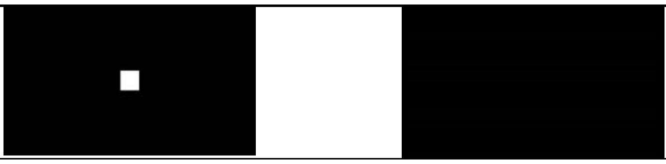
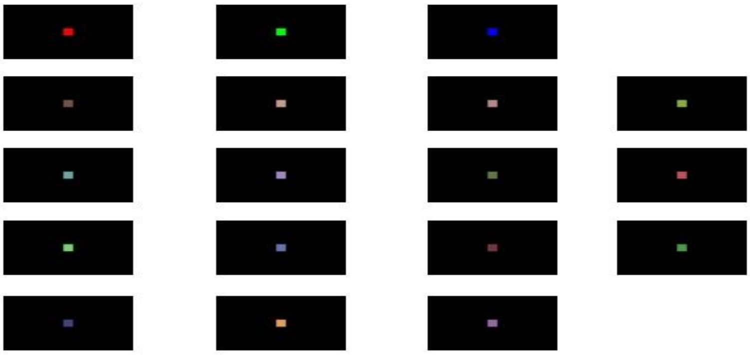
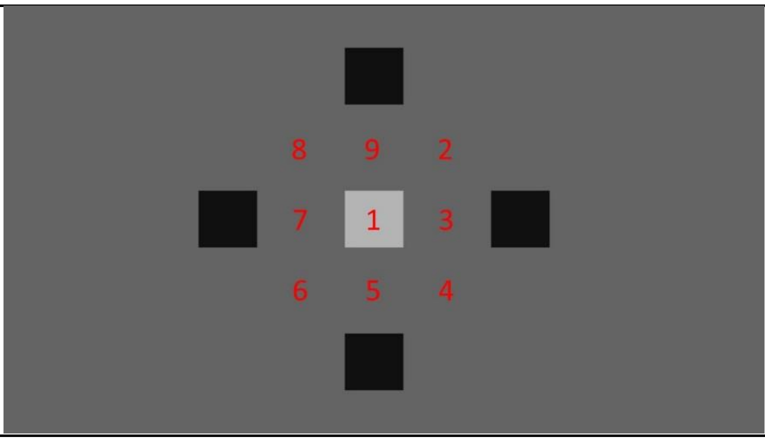
[1] EBU Tech 3320 v4.0 – User Requirements for Video Monitors in Television Production https://tech.ebu.ch/publications/tech3320

[2] EBU Tech 3325 v1.0 - Methods for the Measurement of the performance of Studio Monitors https://tech.ebu.ch/publications/tech3325

[3] Spectroradiometer CS-2000/CS-2000A, Instruction manual, Konica Minolta, https://www.konicaminolta.com/instruments/download/instruction_manual/display/pdf/cs-2000-2000a_instruction_eng.pdf

[4] EBU Tech Review 2009 Q2 – Display Measurement — a simple approach to small-area luminance uniformity testing https://tech.ebu.ch/publications/trev_2009-Q2

Annex 1 Test patterns used

Test pattern	Used for
	E.g. contrast Ratio (code value 940, code value 64)
	Measurement of transfer function as defined in EBU Tech 3325
	EBU colours as defined in EBU Tech 3325
	Cross talk measurement points (average of 2, 4, 6, 8 as reference, deviation from average)