1080P/50: THE THIRD-GENERATION HDTV SYSTEM. IS IT A FEASIBLE OPTION FOR PRODUCTION AND EMISSION?

Hoffmann H., Itagaki T., Wood D.

1080P/50: СИСТЕМА ТБВЧ ТРЕТЬОГО ПОКОЛІННЯ. ЧИ ЦЕ Є РЕАЛЬНИЙ ВИБІР ДЛЯ ВИРОБНИЦТВА І МОВЛЕННЯ ТВ ПРОГРАМ?

EBU Technical Department, Switzerland

Brunel University, School of Engineering and Design, United Kingdom

Recently introduced cameras provide outputs in 1920 × 1080 progressive with frame rates at 50 and 60 Hz (abbreviated 1080p50/60), and new studio interfaces are available that support the required baseband bit rate. In addition, the developments in mezzanine studio compression for-mats for 1080p50/60 further support the format in the production environment. One of the critical questions for the future of HDTV is whether the 1080p/50 HDTV format would provide a significant image-quality improvement when delivered to consumers, to justify its use, compared to the existing 720p/50 or 1080i/25 HDTV formats. This paper presents the results of initial subjective tests to begin to answer this question. The European Broadcasting Union (EBU) tests conditions, sequence selection, and coding of the sequences with MPEG-4 AVC, and subjective test are described. The results of the tests are presented and discussed, and a conclusion is drawn. First indications are that the "quality efficiency" of AVC compressed 1080p/50 may actually be higher than that of AVC compressed 1080i/25. This may sound like a surprising result, but some suggestions for the reasons are given.

Similar to telecommunication or computer technologies, HDTV has developed over time in cycles or generations of technology. The original generation of HDTV was based on analog technology and included the MUSE system in Japan and the work conducted in Europe in the Eureka 95 Project (HD-MAC). Practical digital systems of what is considered the first-generation digital HDTV used the MPEG-2 compression system in emission. However, HDTV in these two generations suffered from high receiver costs and were unattractive, because of the size and weight of cathode ray tube (CRT)-based HDTV receivers.

Today, there is a dramatic change in this situation because of the availability of mass-market HDTV displays based on non-CRT technology such as liquid crystal displays (LCDs), plasma display panels (PDPs), or digital light processing (DLP) projector systems. In Europe, the displays are so attractive that consumers have been investing in HD- ready displays [1] without having significant free-to-air HDTV broadcasts available. The second-generation HDTV being made avail- able will still use the HDTV baseband formats devised 30 years ago (1080i) and 10 years ago (720p). However, it does apply advanced compression (MPEG-4 AVC or VC1) and uses new delivery mechanisms to the consumer, such as DVB-S2 or IP.

For the purpose of this paper, the second- and third-generation HDTV baseband formats are defined with the following abbreviations:

Second-Generation digital HDTV:

720p50/60: An HDTV format with 1280 horizontal and 720 pixel vertical resolution, progressive scanning at 50 frames/sec, as specified in SMPTE 296M-2001 [2]

1080i25/30: An HDTV format with 1920 horizontal and 1080 pixel vertical resolution, interlaced scanning at 25 frames or 50 fields/sec, as specified in SMPTE 274M [3] or ITU-R BT.709-5 [4]

Third-Generation HDTV:

1080p50/60: An HDTV format with 1920 horizontal and 1080 vertical resolution, progressive scanning at 50 or 60 frames/sec, as specified in SMPTE 274M or ITU-R BT.709-5.

In the production environment, new cameras are being made available that provide capture at the sensor 1920 × 1080 pixels or more, at frame rates of 50 and 60 Hz, and, for particular applications, up to 300 Hz. At the interface point between the camera and subsequent studio systems, the captured signal can be converted to 720p/50 or 1080i/25, to support the legacy standard high-definition serial digital interface (HD-SDI)-based studio infrastructure. HD-SDI is specified in SMPTE 292M [5] and the equivalent ITU-R BT.1120-3 [6] and pro- vides today's state-of-the-art HD-SDI for realtime exchange of HDTV signals via single coaxial or fiber links. In order to cope with the bit rate demand of applications such as digital cinema or post-processing, a dual-link specification (SMPTE 372M) [7] of the HD-SDI has been defined. The total bit rate of the single-link interface is 1.485 Gbits/sec, thus the dual-link interface provides 2.97 Gbit/sec. However, dual-link applications may not be convenient for studio-wide installations.

An acceptable 1080p50/60 studio interface for a 4:2:2 sampling structure with 10 bits/sample would require a 3 Gbit/sec single-link interface solution. SMPTE has recently standardized [8, 9] an extension of the HD-SDI interface to 3 Gbits/sec, and first silicon implementations will soon become available on the market. A drawback of this new standard, however, is a limitation in coaxial cable length to approximately 80 to 90 m. Another SMPTE standardization activity covers a 10 Gbit/sec optical fiber link that would also meet the bandwidth requirements of 1080p50/60.

An alternative solution proposed by the British Broadcasting Corporation (BBC) and under standardization in SMPTE is to apply a mild, virtually reversible (mezzanine) compression at the interface points of the HD-SDI infrastructure, to compress a 1080p50/60 signal so that it can be mapped into the available 1.485 Gbit/sec HD-SDI infrastructure.

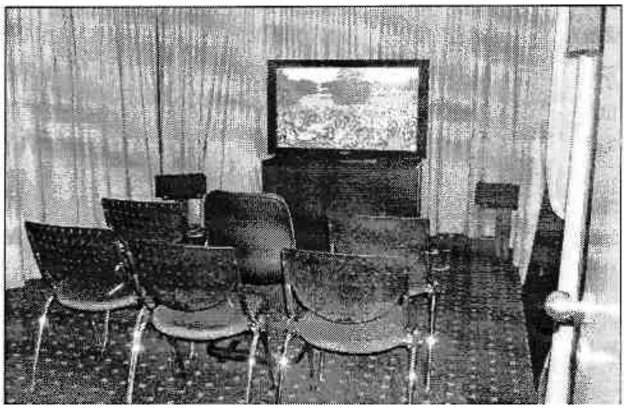

Figure 1 – Room configuration (in bright light for the photo)

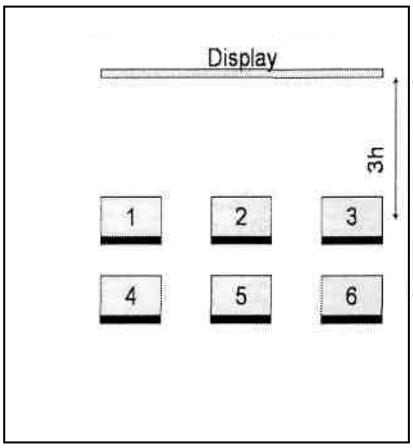

Figure 2 – Seating positions at 3h and 4h

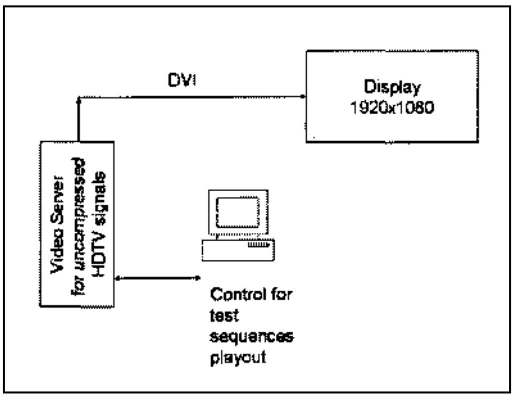

Figure 3 – Infrastructure

These developments show that it is valid to propose an advance on today's first- and second-generation HDTV systems to a third generation, and to study the potential of a progressive HDTV chain in 1920 × 1080 progressive scan with a 50- or 60-Hz frame rate. This format would also have the advantage of being independent of the “political” discussion that surrounds the 1080i/25 or 720p/50 format, and would offer video-quality headroom for a variety of genres and programs, up to some digital cinema applications.

Although there are advantages of 1080p50/60 in the production environment, the question of whether it would also bring a significant quality improvement to the consumer in the home is still open. In particular, the bit rate/quality requirements for emission compared to existing second-generation HDTV formats are important to establish. In order to answer this question, a first round of subjective test to assess identical 720p/50, 1080i/25, and 1080p/50 sequences were performed at the European Broadcasting Union (EBU) headquarters in Geneva.

Setup of the Subjective Test

Subjective tests according to the following setup were established.

Viewing Room

The viewing room was setup according to the guide- lines of ITU-R BT.500-11 [10] with D6500 ambient lighting conditions. Before each session on subsequent days, backlight conditions in the room were rechecked with a Minolta CS100 photospot meter to 10 % of the peak brightness (100 cd/m^2) of the display. D65 florescent lights were used for backlight illumination, and care was taken that no light or reflections appeared on the display (Figure 1).

Display Selection

A critical part of performing subjective image-quality assessment for HDTV is the selection of the display. The professional broadcasting community is aware that consumers will no longer view television on CRT-based displays, but rather FPDs such as LCDs, PDPs, projection devices (e.g., Micro-mirror based), and so on; therefore, these need to be used for today's subjective evaluations. These types of displays generally mask picture impairments to a lesser extent than CRTs, and thus, compared to CRT displays, can be apparent magnifiers of impairments. However, a significant difficulty with subjective tests is that none of the available FPDs can be considered equivalent to a grade 1 reference CRT monitor [11, 12] in terms of color fidelity and other elements. There are ongoing initiatives by ITU-R, SMPTE, and EBU to define criteria's for using FPD as a grade 1 equivalent; however, results are still pending. Among other factors, there are clear difficulties in specifying the typical FPD parameters for video signal processing such as deinterlacing or image scaling. The chosen display represents a compromise of these factors.

For the purpose of the subjective evaluations, a prototype PDP with 50-in. diagonal size and a resolution of 1920 x 1080 pixels was used. This was essentially the most transparent FPD monitor that could be identified.

The display was set to the following conditions:

- Display standard-settings were activated by reset- ting the display.
- Pure cinema: Off
- Frame repetition: 100 Hz
- DVI RGB range: 16-235
- Color temperature setting: Mid ("natural tone")
- Gamma characteristic: 2
- Color transient improvement: Off

- Digital noise reduction: Activated
- MPEG Noise Reduction: Off

Dynamic contrast: Off

- Automatic contrast level: Off
- Interlaces 2 (standard setting)

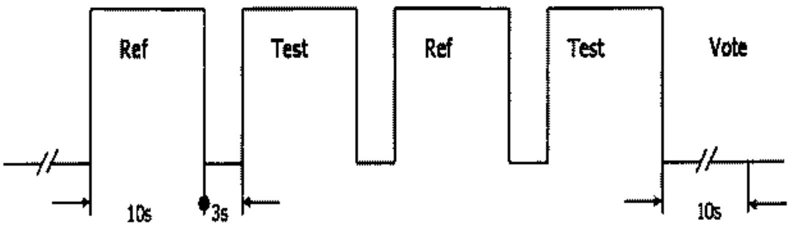

Figure 4 – Test sequence

Brightness and Contrast

The display warmup time was 30 min. First, bright- ness was adjusted until picture line-up generation equipment (PLUGE) was just visible, then contrast was adjusted until peak brightness was set to 100 cd/m^2 (measured with a Minolta CS100, in slow response mode at 1 m distance and perpendicular to the screen). This procedure was repeated until a good compromise between disappearance and just visibility of the PLUGE and brightness and luminance of 100 cd/m^2 was found.

Data Range for Subjective Tests

The video data range setting on the digital video input (DVI) of the display was limited to 16-235 according to the encoding rules of HDTV signals in ITU-R BT.709,4 respectively, SMPTE 274M3 for 1080i/25 and 1080p/50 and SMPTE 296M2 for the 720p/50 signal.

Seating Position

The advantage of a PDP with less critical viewing angle conditions than an LCD, allowed the placement of three observers per row in front of the screen. The distance for the first row was three pictures high, and four pictures high for the second row (Figure 2).

Server and Infrastructure

The setup for the equipment and infrastructure is given in Figure 3.

Interfacing between the signal source and the display was DVI with 4 m length. The 1080p/50, 720p/50, and 1080i/25 source material was played out from a video server in uncompressed form via DVI.

Observers, Test Procedure, and Showing of Sequences

Selection of observers: 21 nonexpert viewers, normal vision tests were performed, male and female, average age. Training sequences and explanation were given before the viewings, and short breaks between the series were offered to relax observers.

The Double Stimulus Impairment Scale (DSIS) method according to ITU-R BT.500-1110 was used in preparation and presentation of the test sequences.

Each sequence was shown twice according to ITU-R BT.500-11 DSIS, Variant II, thus repetition r = 2 (Figure 4).

The uncompressed reference image with a length of 10 sec was followed by mid gray of 3 sec, followed by the impaired test image. After that, the observers had time to vote.

During the voting period, the observers were asked whether they could observe a difference between the reference and the test signal and to mark the corresponding category according to the five in ITU-R BT.500-11 defined terms:

5 imperceptible

4 perceptible, but not annoying

3 slightly annoying

2 annoying

1 very annoying

The tests included sequences, with the reference as test sequence to check reliability and any reference off- set.

The test series included the following HDTV formats:

Series A:

Reference: 1080p/50 uncompressed, 4:2:2, 8-bit resolution.

Test: 1080p/50 at 20 Mbit/sec, 18 Mbit/sec, 16 Mbit/sec, 13 Mbit/sec, 10 Mbit/sec, and 8 Mbit/sec uncompressed.

Series B:

Reference: 1080i/25 uncompressed, 4:2:2, 8-bit resolution.

Test: 1080i/25, 4:2:2, 8-bit resolution at 18 Mbits/sec, 16 Mbits/sec, 13 Mbits/sec, 10 Mbits/sec, 8 Mbits/sec, and 6 Mbits/sec uncompressed.

Series C:

Reference: 720p/50 uncompressed, 4:2:2, 8-bit resolution.

Test: 1080i/25, 4:2:2, 8-bit resolution at 18 Mbits/sec, 16 Mbits/sec, 13 Mbits/sec, 10 Mbits/sec, 8 Mbit/sec, and 6 Mbit/sec uncompressed.

Each test sequence was presented in random order (i.e., ref/13 Mbits/sec, ref/18 Mbits/sec, and so on).

A total of 4 consecutive subjective tests were con- ducted over 2 days with 3 times 6 observers and 1 times 3 observers. Observers were not informed as to which HDTV format they were evaluating, and, in addition the presentation of the Series A, B, C was changed between the tests.

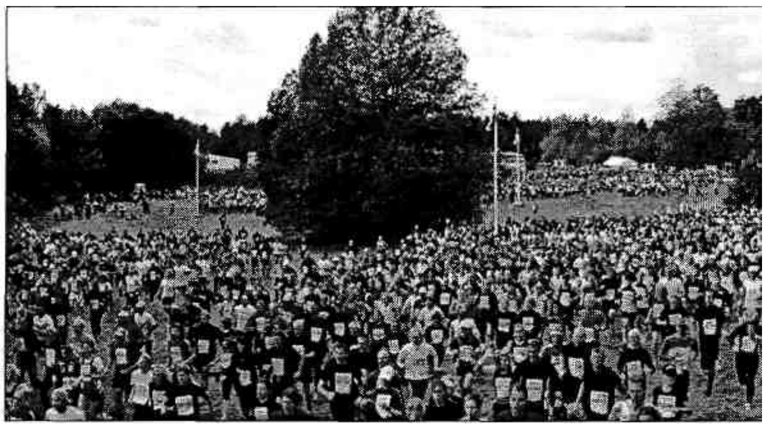

Figure 5 – Crowd Run sequence (uncompressed 720p/50 converted to JPEG in this figure)

Session 1: A, B, C

Session 2: C, B, A

Session 3: A, C, B

Session 4: B, C, A

Voting was conducted on paper. Data from each observer comprised name, age, sex, vision (test was conducted), and seating position. Each page of the voting sheets corresponded to one reference-test sequence, and a person providing guidance, made sure that observer did not vote on the wrong pages.

Test Content

A crucial question was the selection of appropriate test content.

For this initial and first series of subjective tests, it was decided to use "critical but not unduly so" content with rich detail and movement (as is usually seen in sports programs). In addition, the prerequisite was that the identical content should be available in all three HDTV formats (1080i/25, 720p/50, and 1080p/50) being tested. Thus the potential options were to use either artificially generated sequences that were shot with three HDTV cameras at the same time, or a single- camera shot with sufficiently high resolution that would allow down converting to the three HDTV formats under test. The latter option seemed the best and was used. Swedish Television generated 65 mm film content with 50 frames/sec. The material "SVT High-Definition Multiformat Test Set,"[13] was digitized and post-processed in a resolution of 3840 × 2160 pixels and then down converted to 1080p/50, 1080i/25, and 720p/50.

The selected scene for the subjective tests was "Crowd Run" (Figure 5) and could be categorized as difficult and demanding, but not unduly so, in the sense that it could still be conceivably part of a television program.

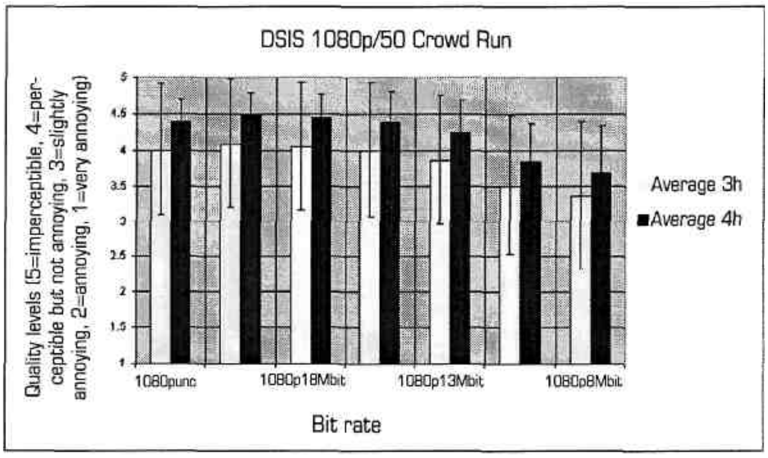

Figure 6 – Mean score of 1080p/50 for various bit rates and at 4h and 3h viewing distance with corresponding error bars

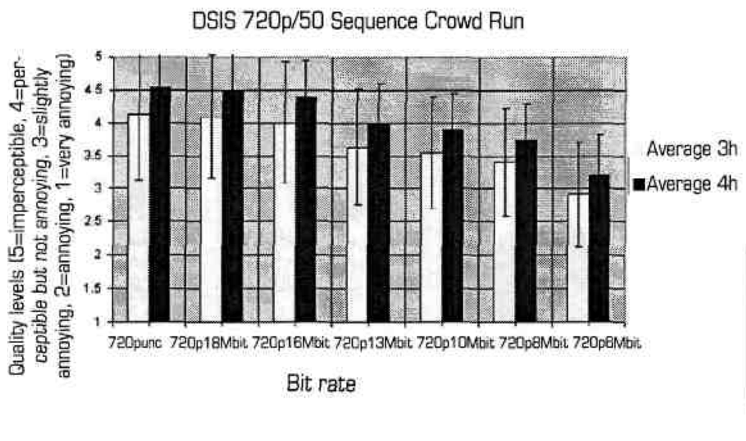

Figure 7 – Mean score of 720p/50 for various bit rates and at 4h and 3h viewing distance with corresponding error bars

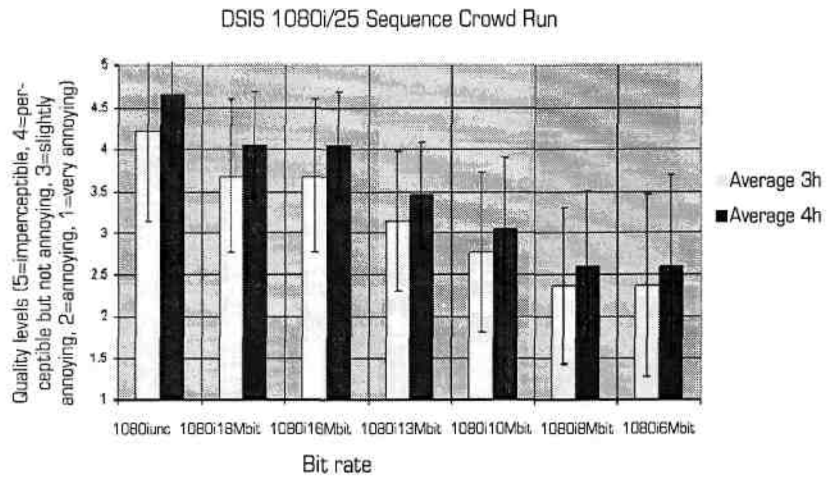

Figure 8 – Mean score of 1080i/25 for various bit rates and at 4h and 3h viewing distance with corresponding error bars

Test Results

Figures 6, 7, and 8 show the mean score *Ujkr* ± *δjkr* for the viewing distance 3h and 4h, for the three HDTV formats: 1080p/50, 720p/50, and 1080i/25.

Interpretation and Discussion of Test Results

The tests reinforced the conclusions of earlier investigations on the use of 720p/50 and 1080i/25, which found that a progressive HDTV format provides a better perceived image quality when compressed with MPEG-4 AVC and viewed on FPD in typical broadcast bit rates between 6 and 18 Mbits/sec. Even as high as 13 Mbits/sec the compressed 1080i/25 format is rated as slightly annoying, compared to the reference 1080i/25 uncompressed.

Although to some extent predictable by theory, the surprise result of these tests was that in practice, the 1080p/50 HDTV format performed well (Fig. 6). At 8 Mbits/sec, the format was still rated above "slightly annoying," compared to the 1080p/50 uncompressed reference.

The 1080p/50 signal has double the pixel rate and baseband bandwidth of the 1080i/25 signal, so if these were the only factors influencing compression efficiency the result would be remarkable. If this were the case, one would have expected 1080p/50 to require a higher bit rate than the 1080i/25 format, to achieve acceptable quality in its failure curve. The impressive performance of the 1080p/50 is due to the following:

- The 1080p/50 starts with a higher quality than the 1080i/25, and the 1080p display is able to reap this benefit.
- The MPEG-4 AVC encoder compresses progressive signals more efficiently than interlaced, which is probably due, mainly to improvements in the accuracy of motion estimation. In fact, the first stages of the compression system substitute for interlacing. They do the same job—bandwidth reduction—more efficiently than interlacing.
- With a progressive 1080p/50 signal, the signal processing in the 1920 x 1080 display is minimized, because there is no need for deinterlacing or scaling to the native resolution of the display.

Conclusion and Further Work

The results of this initial subjective test suggest that it is very worthwhile to continue research and studies on the 1080p50/60 format for future HDTV applications and future delivery to the home, particularly for large displays >50 in. and 1920 x 1080 resolution or higher. Subjective tests have shown that a progressive HDTV format has clear bit rate and image quality advantages. The currently available 720p/50 is therefore a very well- suited emission format, provided that no interlaced was used in content generation.

Findings that the 1080p50/60 format provided a better failure characteristic than the 1080i/25 HDTV format, at the same bit rate, suggests that the 1080p50/60 HDTV can seriously be considered not only for production but also in emission.

It was noted in the tests, that the DSIS method only shows the failure characteristics of the individual formats; consequently no direct comparison of the three formats in one graph was permissible. The authors have therefore developed a new subjective testing method called Triple Stimulus Continuous Evaluation Scale (TSCES) [14].

The problem of a suitable large FPD reference display is significant and requires urgent work of the standardization bodies. Also, the subjective test conducted in this paper represents only initial research; further tests with different kinds of content and different types of sources (i.e., created with CMOS or CCD cameras with native 1920 × 1080 pixel or high- er resolution and at least 50 Hz frame rates) must now be done, as well as viewing on different kinds of display technologies such as LCD with a CRT equivalent grade-1 characteristic. In addition, the settings for encoding and decoding should be further optimized.

The authors have been generating new sequences of 1080p/50 with a Sony HDC1500 CCD camera and have conducted further testing sessions with the DSIS method and the new TSCES method. These tests were conducted with 50 in. PDP or 52 in. LCD with 1920 × 1080 resolutions. In general, it can be reported that the new tests have confirmed the overall advantages of 50- 60 Hz progressive scanning systems. With higher critical content and at lower bit rates (<16 Mbits/sec), the 720p/50 was rated best, and with low critical content at higher bit rates (>16 Mbits/sec), the 1080p/50 format was rated best (very little compression artifacts and assessors appreciated the higher resolution compared to 720p/50). 1080i/25 was not as good as the other two formats. The authors will report in details about these tests in technical literature in the near future.

Acknowledgments

We would like to thank the Mr. Hinz and Mr. Schwarz from the Heinrich Herz Institut in Berlin for coding the sequences. In addition, we would also like to thank Lars Haglund from SVT for his valuable work in providing the sequences and for his contributions in the preparation phase of these tests.

References

1 European Industry Association EICTA, “Conditions for High Definition Labeling of Display Devices” January 19, 2005.

2 SMPTE 296M-2001, “1280 × 720 Scanning, Analogue and Digital Representation and Analogue Interface” www.smpte.org.

3 SMPTE 274M-1998, “1920 × 1080 Scanning and Analogue and Parallel Digital Interfaces for Multiple Picture Rates” www.smpte.org.

4 ITU-R BT.709-5, “Parameter Values for the HDTV Standards for Production and International Programme Exchange” 2002.

5 SMPTE 292M-1998, “Bit-serial Digital Interface for High- Definition Television Systems” www.smpte.org.

6 ITU-R BT.1120-3, “Digital Interfaces for HDTV Studio Signals” 2000.

7 SMPTE 372M-2001, “Dual Link 292M Interface for 1920 × 1080 Picture Raster” www.smpte.org.

8 SMPTE 424M-2005, “3 Gb/s Signal/Data Serial Interface” www.smpte.org.

9 SMPTE 425M-2005, “3 Gb/s Signal/Data Serial Interface—Source Image Format Mapping,” www.smpte.org.

10 ITU-R BT.500-11, “Methodology for the Subjective Assessment of the Quality of Television Pictures” 2003.

11 EBU-R65, “Grade-1 Colour Picture Monitors” 1997.

12 EBU-Tech 3263 “Specifications of Grade-1 Colour Picture Monitors” European Broadcasting Union, Geneva, July 1996; available at www.ebu.ch.

13 L. Haglund, “The SVT High-definition Multiformat Test Set” Swedish Television, Stockholm, 2006; ftp://vqeg.its.bldrdoc.gov.

14 H. Hoffmann, T. Itagaki, and D. Wood, “New Psycho- physical Method of Television Picture Quality Evaluation (EBU-II)” Electronics Letters, accepted for publication in next edition.