|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  21st Meeting: Warsaw, PL, 19–26 Jun. 2015 | Document: JCTVC-U1003 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Interlaced video and format range extensions verification test results** | | |
| *Status:* | Output document | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | C. Rosewarne,  V. Baroncini,  G. Barroux  G. J. Sullivan,  A. Tourapis,  M. Naccari | Email: | [chris.rosewarne@cisra.canon.com.au](mailto:chris.rosewarne@cisra.canon.com.au)  [vittorio@fub.it](mailto:vittorio@fub.it)  [guillaume.b@jp.fujitsu.com](mailto:guillaume.b@jp.fujitsu.com)  [garysull@microsoft.com](mailto:garysull@microsoft.com)  [atourapis@apple.com](mailto:atourapis@apple.com)  [matteo.naccari@bbc.co.uk](mailto:matteo.naccari@bbc.co.uk) |
| *Source:* | JCT-VC | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document provides the results of the subjective evaluation activity performed to compare the compression performance of the HEVC and AVC video coding standards on interlaced-scan and format range extensions (RExt) video test material. In particular, compression tests were performed on 8 bit 4:2:0 interlaced-scan video, 10 bit 4:2:2 progressive-scan video, and 10 bit 4:4:4 progressive-scan video test material. For HEVC the Main, Main 4:2:2 10, and Main 4:4:4 10 profiles were used, respectively. Similarly, for AVC, the High and High 4:4:4 Predictive profiles were used (for 4:2:2 content, the High 4:4:4 Predictive profile was used with the chroma format adjusted). The HEVC and AVC bitstreams were produced using recent available versions of the HM and JM reference software, respectively.

The interlaced-scan verification tests were conducted using random access (RA) and low delay (LD) configurations. The format range extensions tests were conducted using the RA configuration only. In all cases, the HEVC bitstreams were coded at half the bit rate of the AVC bitstreams.

The subjective test results show that in no case were the half-bit-rate HEVC bitstreams providing statistically significantly worse quality than the corresponding AVC bitstreams, i.e., the HEVC bitstreams achieved subjective quality scores that were either equivalent to or higher than the subjective quality scores of the corresponding AVC bitstreams, even though the HEVC bit rates were half of the AVC bit rates.

The tests were conducted in accordance with the previously-issued “Verification test plan for HEVC RExt profiles, and Main profile usage for interlaced video” (JCTVC-T1003). However, an adjustment of some of the bit rates for specific range extensions test sequences was performed in order to better facilitate the subjective test.

# Introduction

The verification test for the coding performance of HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2) on interlaced-scan and format range extensions (RExt) video material was performed between the 20th and 21st JCT-VC meetings. This test was the culmination of test planning and test sequence acquisition over several preceding meetings. The reference used for comparison with HEVC was AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10), which is the previous jointly-developed video coding standard and is the current dominant video coding technology used world-wide.

Since the relevant video coding standards specify only a syntax and decoding process, without prescribing how to perform encoding, it is not possible in general to directly compare the capabilities of such standards. Instead, some particular encoding method must be selected as a proxy to represent the coding capabilities of a standard, as encoding techniques and encoded video quality may vary widely from implementation to implementation. For purposes of this test, reference software codebases developed during the standardization process were used to represent each standard – i.e., the HEVC Model (HM) in the case of HEVC and the Joint Model (JM) in the case of AVC. These two software codebases were developed for similar purposes, use similar encoding techniques such as rate-distortion optimization decision-making processing, and were configured in a very similar way for these tests, i.e., in terms of hierarchical picture referencing structures, random-access refresh periods, quantization control settings, etc.

The verification tests involved a subjective evaluation that was performed at a test lab in Rome. During this subjective evaluation, viewing subjects were asked to rank the relative degradation introduced by compression on a set of video test sequences at several bit rates. Ranking was performed by assigning a mean opinion score (MOS) to each playback of a sequence, with the original sequence also played back and identified as the reference.

This document reports on the experiments and results obtained for the verification tests. The remainder of the text is organized as follows. Section 2 describes the test sequences and acknowledges the companies who made these sequences available for this test. Section 3 describes encoder configurations and target bit rates used for the selected sequences. Section 4 describes the subjective test conditions and methodology while Section 5 provides a report of the results obtained from conducting the subjective test. A discussion of the results obtained is presented in Section 6. Section 7 gives a brief conclusion.

# Test sequences

## Selected sequences

The test sequences used for the HEVC interlaced-scan verification test are listed in Table 1, while the test sequences used for the HEVC Range Extensions verification test are listed in Table 2.

Table 1 – Test sequences used in the HEVC interlaced-scan verification testing

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **Test sequence** | **Resolution** | **Frame Rate** |
| YCbCr 4:2:0i | CBS Tennis | 1920×1080 | 60i |
| CBS Basketball | 1920×1080 | 60i |
| ParkJoy | 1920×1080 | 50i[[1]](#footnote-1) |
| CheersHD4\* | 1920×1080 | 60i |
| MusicHD1 | 1920×1080 | 60i |

\* Note that the sequence NewMobCal, which was originally specified in [1], was replaced by CheersHD4 due to an assessment by the test coordinator that this would result in more useful results from the subjective test.

Table 2 – Test sequences used in the HEVC Range Extensions verification testing

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Class** | **Test sequence** | **Resolution** | **BitDepth** | **Chroma Format** | **Frame Rate** |
| YCbCr 4:2:2 | EBUHorse | 1920×1080 | 10 | 4:2:2 | 50p |
| EBUKidsSoccer | 1920×1080 | 10 | 4:2:2 | 50p |
| EBUWaterRocksClose | 1920×1080 | 10 | 4:2:2 | 50p |
| YCbCr 4:4:4 | BirdsInCage | 1920×1080 | 10 | 4:4:4 | 60p[[2]](#footnote-2) |
| CrowdRun | 1920×1080 | 10 | 4:4:4 | 50p |
| EBULupoCandlelight | 1920×1080 | 10 | 4:4:4 | 50p |

## Availability of test sequences

The verification tests used test sequences and test conditions that were designed to represent typical applications of the tested technology. Professional material and conditions were also evaluated. As a consequence, some usage restrictions were imposed on some of the sequences used in this activity. It is important to note that although some actions are needed to access the content, all content will remain accessible for an extended period for cross-verification purposes.

### CBS sequences

The “CBS Tennis” and “CBS Basketball” sequences were kindly provided by CBS television. They can be obtained from the authors of this document or from CBS television after sending, with the appropriate company name, the following message to Mr. Greg Coppa ([gmcoppa@cbs.com](mailto:gmcoppa@cbs.com)):

\*Company name\* agrees to not use the test clips, "CBS Basketball" and "CBS Tennis" in any public display, conventions (e.g. IBC), demonstrations, etc., and that they are for MPEG internal research and development for MPEG/JCT-VC tests and interlace AHG experiments exclusively.

CBS Television particularly insists on the following two points:

(1) The CBS video clips, original or processed, will not be made public per the usage agreement and

(2) a single video frame “identifying” the CBS clips may be used in a public description of the tests.

### NBC Universal sequences

The “CheersHD4” and “MusicHD1” sequences were kindly provided by NBC Universal. They can be obtained from the authors of this document or from NBC Universal after a signed usage agreement designed by NBC Universal has been sent and accepted.

NBC Universal can be reached through Mr. Ian Slotin ([Ian.Slotin@nbcuni.com](mailto:Ian.Slotin@nbcuni.com)) and Mr. Thomas Bause Mason ([Thomas1.BauseMason@nbcuni.com](mailto:Thomas1.BauseMason@nbcuni.com)).

### SVT sequences

The “ParkJoy” and the “CrowdRun” sequences were originally provided by SVT Play for various HEVC tests. The clips and a description of the usage agreement and the video features can be found on the VQEG ftp server at <ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat>

### Planet Inc. sequences

The “BirdsInCage” sequence was provided by Planet Inc. for the HEVC Range Extensions standardization process. The corresponding copyright statement is:

The test sequence and all intellectual property rights therein remain the property of the owner below. This material can only be used for the purpose of developing HEVC standards. This material cannot be distributed with charge. The owner makes no warranties with respect to the material and expressly disclaims any warranties regarding its fitness for any purpose.

Owner of these sequences:

Owner: Plannet inc.

Production: Plannet inc. and IMAGICA Corp.

### EBU sequences

The “EBUHorse”, “EBUKidsSoccer”, “EBUWaterRocksClose”, and “EBULupoCandlelight” sequences were provided by EBU for the HEVC Range Extensions standardization process. The corresponding copyright statement is:

This material has been provided by the EBU under the Creative Commons

Attribution-NonCommercial-NoDerivs 3.0 licence:

http://creativecommons.org/licenses/by-nc-nd/3.0/

Please see http://tech.ebu.ch/testsequences for further information

# Test conditions

## General description

The comparison of the subjective performance of HEVC and AVC was achieved using recent versions of the reference software implementation for each standard, i.e., HM-16.2 for HEVC and JM-18.6 for AVC. These implementations were chosen due to easy access and to facilitate reproduction of the test results. However, it is noted that commercial encoders may differ substantially in behaviour when compared to the reference software implementations. Therefore, results and potentially conclusions using different encoders may differ from this evaluation considerably.

In general terms, the test followed the test plan described in the “Verification test plan for HEVC RExt profiles, and Main profile usage for interlaced video” (JCTVC-T1003) [1]. However, as mentioned earlier, one interlaced-scan test sequence had to be changed (see Section 2.1), and several of the originally defined bit rates had to be adjusted to better facilitate the subjective evaluation process. These changes are further described in Section 3.3.

An accompanying spreadsheet includes a comparison of the objective performance of the HM and JM software under the defined test conditions. Note that target bit rate constraints for the HEVC and AVC bitstreams were met without the use of a rate control scheme. Instead, fixed Quantization Parameters (QPs), as are commonly used in video codec standardization development, were used.

Details of the defined test conditions are further described below.

## Test scope

For the interlaced-scan verification tests, both the Random Access (RA) and Low Delay (LD) coding configurations were tested. However, only RA coding configurations were used for the Range Extensions verification tests. The Range Extensions verification tests, on the other hand, were performed using both 4:2:2 and 4:4:4 sequences.

The test set was therefore defined as follows:

* Interlaced-scan material:
  + 4 bit rates \* 1 config (RA) \* 5 (4:2:0 8-bit sequences) = 20 streams
  + 4 bit rates \* 1 config (LD) \* 5 (4:2:0 8-bit sequences) = 20 streams
* Range Extensions material:
  + 4 bit rates \* 1 config (RA) \* 3 (4:2:2 10-bit sequences) = 12 streams
  + 4 bit rates \* 1 config (RA) \* 3 (4:4:4 10-bit sequences) = 12 streams

In total, 64 HEVC bitstreams and 64 AVC bitstreams were encoded for use in the subjective viewing tests.

## Bit-rate targeting

For each configuration used in the Interlaced-scan and the Range Extensions verification tests, bit rates for encoding were selected based upon the anticipated operating point for each respective configuration.

In order to closely reach, but not exceed, the targeted bit rate constraint, bitstreams were permitted to include one QP increment. The position at which the QP increment occurs is dependent upon a ‘fractional QP’ setting, supported by the JM and HM software.

Using the fractional QP encoder functionality to produce the bitstreams, it was found that the HM could provide very precise control of the resulting bit rate. However, the JM tended to introduce step changes in bit rate at particular fractional QPs. Thus, the achieved bit rates tend to be closer to the target bit rates for the HM than for the JM.

### Interlaced-scan video verification

For the interlaced-scan verification tests, the bit-rate targets are shown in Table 3.

Table 3 – Interlaced-scan verification test target bit rates (Mbps)

|  |  |  |
| --- | --- | --- |
|  | **HEVC** | **AVC** |
| **RA 4:2:0i 8-bit** | 1.5, 3, 5, 8 | 3, 6, 10, 16 (2× the HEVC bit rates) |
| **LD: 4:2:0i 8-bit** | 3, 6, 10, 16 | 6, 12, 20, 32 (2× the HEVC bit rates) |

Table 4 and Table 5 list the QP values applied when encoding each sequence for the interlaced-scan test using the random access and low delay configurations respectively. The resulting actual measured bit rates for each test are also provided.

Table 4 – QP values and actual bit rate used for interlaced-scan testing in RA mode

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Class** | **Test sequence** | **HEVC** | | **AVC** | |
| **QP** | **Bit rate (kbps)** | **QP** | **Bit rate (kbps)** |
| YCbCr 4:2:0i | CBS Tennis | 23.9 | 7828.57 | 21.4 | 13921.19 |
| 25.8 | 4932.40 | 23.7 | 8761.98 |
| 28.5 | 2926.12 | 25.1 | 5707.36 |
| 32.6 | 1486.05 | 29.5 | 2762.07 |
| CBS Basketball | 27.6 | 7954.18 | 25 | 15177.65 |
| 30.1 | 4964.00 | 28 | 8713.20 |
| 33.4 | 2931.00 | 31.7 | 5387.76 |
| 38.3 | 1479.96 | 36.5 | 2645.91 |
| ParkJoy | 31.7 | 7977.96 | 30.1 | 13338.01 |
| 34.2 | 4981.78 | 32.6 | 8769.24 |
| 36.8 | 2973.97 | 34.7 | 5748.69 |
| 40.3 | 1497.77 | 37.4 | 2997.91 |
| CheersHD4 | 33 | 7862.43 | 31.7 | 13962.10 |
| 36.2 | 4991.11 | 34 | 8423.86 |
| 39.4 | 2977.18 | 36.5 | 5892.52 |
| 43.3 | 1491.73 | 40.2 | 2698.64 |
| MusicHD1 | 26.9 | 7833.83 | 24 | 14484.89 |
| 29.8 | 4868.41 | 27.4 | 8942.49 |
| 32.9 | 2957.67 | 30 | 5653.67 |
| 37.8 | 1490.71 | 35.8 | 2685.80 |

Table 5 – QP values and actual bit rate used for interlaced-scan testing in LD mode

| **Class** | **Test sequence** | **HEVC** | | **AVC** | |
| --- | --- | --- | --- | --- | --- |
| **QP** | **Bit rate (kbps)** | **QP** | **Bit rate (kbps)** |
| YCbCr 4:2:0i | CBS Tennis | 21.4 | 15762.96 | 20 | 21087.89 |
| 23.2 | 9690.54 | 21.4 | 17317.23 |
| 25.1 | 5842.63 | 23.4 | 11039.36 |
| 28.1 | 2957.30 | 26 | 5741.45 |
| CBS Basketball | 24.8 | 15577.83 | 24.8 | 23419.15 |
| 26.8 | 9816.53 | 25.1 | 18369.30 |
| 29.3 | 5986.32 | 28 | 10232.68 |
| 33.4 | 2992.20 | 32.1 | 5463.21 |
| ParkJoy | 28.8 | 15985.65 | 28.3 | 22322.81 |
| 31.3 | 9864.78 | 29.4 | 18948.86 |
| 33.8 | 5893.97 | 32.1 | 11051.63 |
| 37.2 | 2993.86 | 36.5 | 5104.96 |
| CheersHD4 | 29.6 | 15822.92 | 29 | 22007.22 |
| 32.4 | 9791.44 | 30 | 19099.11 |
| 35.6 | 5882.89 | 34.9 | 10530.31 |
| 39.7 | 2993.11 | 38 | 5384.44 |
| MusicHD1 | 23.4 | 15719.31 | 23 | 20742.01 |
| 25.7 | 9898.96 | 24.4 | 17507.44 |
| 28.6 | 5877.04 | 27.6 | 10459.69 |
| 32.5 | 2993.24 | 31.3 | 5709.11 |

### Range Extensions verification

For the Range Extensions verification test, the bit-rate targets are defined in Table 6.

Table 6 – Range Extensions test target bit rates (Mbps)

|  |  |  |
| --- | --- | --- |
|  | **HEVC** | **AVC** |
| **RA 4:2:2 10-bit** | 5, 8, 12, 16 | 10, 16, 24, 32 (2× the HEVC bit rates) |
| **RA 4:4:4 10-bit** | 5, 8, 12, 16 | 10, 16, 24, 32 (2× the HEVC bit rates) |

However, after a first preview of the decoded bitstreams performed by the MPEG Test Chair, the difference in quality that was observed in the bitstreams, for almost all of the encoded test sequences, was deemed too small to conduct a formal subjective assessment test.

On one hand, one could argue that staying with the originally assigned QP values seemed to have been meeting the original goal of this verification test. That is that HEVC could provide a similar visual quality as AVC at half the data rate.

On the other hand, it should be noted that a good subjective test experiment requires and should be designed in a manner that the quality range of the video under test is as wide as possible, i.e., ranging from low quality to high quality. However, staying with the original set of QPs did not provide such a wide range of quality. The majority of the decoded video clips were considered to be of very high quality. If a test subject viewed content where the quality is assessed as always being excellent, then this may result in a certain type of frustration during the execution of their task of ranking video on the basis of video quality, resulting in unstable or otherwise unusable results.

As a consequence, the EBULupoCandleLight, EBUHorse and EBUKidsSoccer sequences were encoded at lower bit rates than were originally anticipated in [1], resulting in bitstreams with more visually significant artefacts. This issue is further discussed in Section 6.

Table 7 and Table 8 list the QP values applied when encoding sequences for the Range Extensions tests for the 4:2:2 and 4:4:4 test sequences respectively. The resulting bit rates for each test are also provided.

Table 7 – QP values and actual bit rate used for Range Extensions testing with 4:2:2 sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Class** | **Test sequence** | **HEVC** | | **AVC** | |
| **QP** | **Bit rate (kbps)** | **QP** | **Bit rate (kbps)** |
| YCbCr 4:2:2 | EBUHorse | 23.6 | 15623.02 | 20 | 28010.95 |
| 30\* | 3948.02 | 25.1\* | 9081.52 |
| 34\* | 2159.66 | 28\* | 5569.07 |
| 37\* | 1382.33 | 33\* | 2862.78 |
| EBUKidsSoccer | 25.5 | 15933.20 | 22 | 26828.59 |
| 31.4\* | 4915.98 | 28.1\* | 8759.13 |
| 35\* | 2266.72 | 32\* | 4024.30 |
| 40\* | 700.42 | 36\* | 1870.42 |
| EBUWaterRocksClose | 35.5 | 15558.53 | 32.1 | 28590.01 |
| 37.5 | 11743.93 | 34 | 22114.97 |
| 40.1 | 7914.68 | 37.8 | 14366.56 |
| 42.9 | 4947.23 | 40 | 9032.38 |

*\* Bitstream with different QP than specified in [1].*

Table 8 – QP values and actual bit rate used for Range Extensions testing with 4:4:4 sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Class** | **Test sequence** | **HEVC** | | **AVC** | |
| **QP** | **Bit rate (kbps)** | **QP** | **Bit rate (kbps)** |
| YCbCr 4:4:4 | BirdsInCage[[3]](#footnote-3) | 22.9 | 15490.27 | 19.8 | 28491.51 |
| 23.6 | 11872.27 | 20 | 20012.83 |
| 24.5 | 7909.32 | 21.5 | 12990.51 |
| 25.5 | 4872.12 | 22 | 9017.09 |
| CrowdRun | 30.7 | 15887.64 | 27.1 | 29275.22 |
| 32.5 | 11945.58 | 29.5 | 21869.54 |
| 35.2 | 7942.26 | 32.1 | 14275.25 |
| 38.4 | 4865.37 | 35 | 9266.48 |
| EBULupoCandlelight | 17.2 | 15884.10 | 14.4 | 31834.35 |
| 21.6\* | 4909.76 | 18\* | 9860.45 |
| 24\* | 3049.31 | 22\* | 4361.41 |
| 30\* | 1177.81 | 26\* | 2440.36 |

*\* Bitstream with different QP than specified in [1].*

# Subjective test

## Viewing equipment

For the interlaced-scan verification tests a 55UG870V LG consumer TV set was used. This is a 55 inch monitor that can support 4k video resolutions. However, the display was used only in its HD (1920×1080) resolution input mode. Adaptive local diming control and other monitor settings were disabled so as to ensure that there would not be any bias introduced through post processing. The video signal was driven through an HDMI 2.0 cable. The decoded bitstreams were deinterlaced prior to display (i.e., by offline deinterlacing) using the ‘yadif’ deinterlacer that is included in the publicly-available ‘ffmpeg’ software. The deinterlacing process produces an output frame rate equal to the field rate of the interlaced video content. To better facilitate deinterlacing, and to avoid possible chroma phase difference issues, an interlace-aware chroma format conversion from 4:2:0 to 4:2:2 was applied prior to invoking the ffmpeg deinterlacing process. This step was performed using the ‘ChromaConvert’ utility that is provided in the ‘HDRTools’ software package available within MPEG.

A professional-grade CG301W 30 inch EIZO monitor, capable of displaying 10 bit sequences in a 4:4:4 chroma format, was used for the Range Extensions verification tests. The CG301W monitor has a native resolution of 2560×1600 and, to use it for this test, a user-defined profile was created with the screen resolution set at 1920×1200 pixels and the frame rate at 50 fps. Letterboxing, with appropriate black mattes introduced above and below each sequence, was used to display the 1920×1080 material.

For both of the above cases, dedicated hardware and software were used to drive the displays in a manner that provides the necessary smooth and homogeneous presentation of the moving images. Special considerations for handling frame rates and colour/bit-depth representation were made, especially for the case of the Range Extension verification test.

The hardware in particular involved a customized PC system that was based on an Intel i7 2011-3 socket based platform. This system used the new X99 chip-set and DDR4 RAM. A storage system consisting of eight SSD disks in a Raid 0 configuration was used to achieve the throughput performance necessary for these tests. In particular, the throughput of this system was measured to be approximately 2.5 Gb/s. Two different video boards were used for the Interlaced and the Range Extensions verification tests: for the Interlaced verification test a NVidia GTX 960 video board was used; for the Range Extensions verification tests a NVidia Quadro K620, capable of supporting 10 bit video output, was used.

For the video playback a custom-made video player called MUP (Multimedia Universal Player) was used. This player is capable of supporting playback in real time even for the most demanding video sequences included in this test, i.e., 4:4:4 10 bit video clips. No loss in bit-depth precision was introduced at any stage of the playback process.

For the interlaced verification testing, the display refresh rate was set to 60 fps. For the test sequence that was originally produced at 50 fps (ParkJoy), the MUP was set to perform a simple up-conversion to 60 fps by repeating some of the deinterlaced frames during playback. This permitted the playback of the entire sequence at its intended speed and without any frame drops. No temporal artefacts were observed on this material through this frame rate up-conversion process.

For the Range Extensions verification testing, the limit of the display capabilities did not allow a 60 fps display refresh rate to be used. Therefore, the 60 fps video clip (BirdsInCage) was slowed down by playing its frames at 50 fps rather than 60 fps.

## Viewing subjects

To conduct the test, 24 naïve viewing subjects were hired. The viewing subjects were selected with the goal of having as equal gender participation as was feasible. The test subjects included university students that were between 18 and 24 years old. Ultimately, 6 female and 18 male viewing subjects were hired.

The viewing subjects were pre-screened for visual acuity and colour blindness using the Snellen chart and Ishihara table methodologies. A post-screening of the viewing subjects was also performed by analyzing their test scores to exclude outliers.

During the pre-screening, one viewer was dismissed due to colour blindness. Two additional viewers were dismissed during the post screening process due to their low response. This brought the total number of valid viewing subjects down to 21 (6 female and 15 male viewing subjects).

## Viewing conditions

For the Range Extensions tests, the viewing subjects were seated in front of the screen, in groups of two people at a time, at a distance of approximately 2.5 times the screen height of the monitor. This distance, although closer than some traditionally specified subjective test arrangements, was chosen to better facilitate the detection of artefacts at the higher quality levels of the bitstreams used for the Range Extensions tests.

For the interlaced-content verification tests, viewers were seated, in groups of three, on a sofa at a distance of about two times the screen height of the display. This laboratory set-up (see the example in Figure 1) was used in order to better emulate the typical “at home” viewing experience of interlaced content. This was seen as a good compromise between having a comfortable seating arrangement and having the ability to properly assess video quality. It should be noted that the viewing subjects especially appreciated the comfort of this set-up.



Figure 1 – "At home" viewing condition used for interlaced-scan tests

# Test results

The MOS and an associated confidence interval value were computed for each test case, based on the scores collected from the viewers.

Note that the reported confidence intervals are ±1×σ from a given MOS x, rather than the conventional ±2×σ. This approach is taken to facilitate the comparison of two MOS values, e.g. x1 and x2, where non-overlapping confidence intervals require |x1 − x2| > σ1 + σ2. In this case, non-overlapping confidence intervals imply approximately a 95% confidence of a statistically significant difference between x1 and x2 (if σ1 and σ2 differ greatly, this may not be the case).

Table 9 and Table 10 summarize the subjective test results for the interlaced-scan test sequences coded using the random access and low delay configurations, respectively. Results are reported using the MOS and the associated confidence interval for each bit rate. For each sequence, a BD-MOS plot is also provided.

Table 9 – Interlaced-scan Random Access 4:2:0 8 bit sequence subjective test result summary

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S01 CBS Tennis** |  | **Rate** | **MOS** | **CI** | | **AVC** | 13921.2 | 7.60 | 0.24 | | 8761.98 | 7.27 | 0.27 | | 5707.36 | 6.27 | 0.34 | | 2762.07 | 6.27 | 0.36 | | **HEVC** | 7828.57 | 8.00 | 0.23 | | 4932.4 | 7.87 | 0.28 | | 2926.12 | 7.67 | 0.29 | | 1486.05 | 7.27 | 0.28 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S02 CBS Basketball** |  | **Rate** | **MOS** | **CI** | | **AVC** | 15177.7 | 7.33 | 0.19 | | 8713.2 | 7.13 | 0.30 | | 5387.76 | 6.87 | 0.30 | | 2645.91 | 4.20 | 0.22 | | **HEVC** | 7954.18 | 8.53 | 0.22 | | 4964 | 8.27 | 0.25 | | 2931 | 7.07 | 0.22 | | 1479.96 | 4.20 | 0.17 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S03 ParkJoy** |  | **Rate** | **MOS** | **CI** | | **AVC** | 13338 | 6.67 | 0.25 | | 8769.24 | 4.07 | 0.28 | | 5748.69 | 3.00 | 0.19 | | 2997.91 | 1.13 | 0.19 | | **HEVC** | 7977.96 | 6.47 | 0.19 | | 4981.78 | 4.60 | 0.21 | | 2973.97 | 4.53 | 0.27 | | 1497.77 | 1.47 | 0.19 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S04 CheersHD4** |  | **Rate** | **MOS** | **CI** | | **AVC** | 13962.1 | 7.00 | 0.21 | | 8423.86 | 5.47 | 0.23 | | 5892.52 | 3.60 | 0.28 | | 2698.64 | 1.33 | 0.25 | | **HEVC** | 7862.43 | 7.00 | 0.23 | | 4991.11 | 5.87 | 0.23 | | 2977.18 | 4.53 | 0.27 | | 1491.73 | 1.93 | 0.24 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S05 MusicHD1** |  | **Rate** | **MOS** | **CI** | | **AVC** | 14484.9 | 7.07 | 0.28 | | 8942.49 | 6.13 | 0.30 | | 5653.67 | 4.60 | 0.38 | | 2685.8 | 3.00 | 0.29 | | **HEVC** | 7833.83 | 7.60 | 0.25 | | 4868.41 | 6.67 | 0.26 | | 2957.67 | 6.13 | 0.23 | | 1490.71 | 3.33 | 0.26 | |  |

Table 10 – Interlaced-scan low delay 4:2:0 8 bit sequence subjective test result summary

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S01 CBS Tennis** |  | **Rate** | **MOS** | **CI** | | **AVC** | 21087.9 | 7.60 | 0.18 | | 17317.2 | 7.27 | 0.31 | | 11039.4 | 7.13 | 0.21 | | 5741.45 | 7.20 | 0.31 | | **HEVC** | 15763 | 7.93 | 0.25 | | 9690.54 | 7.87 | 0.23 | | 5842.63 | 7.60 | 0.24 | | 2957.3 | 7.40 | 0.30 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S02 CBS Basketball** |  | **Rate** | **MOS** | **CI** | | **AVC** | 23419.2 | 7.87 | 0.23 | | 18369.3 | 7.53 | 0.23 | | 10232.7 | 7.40 | 0.27 | | 5463.21 | 7.20 | 0.27 | | **HEVC** | 15577.8 | 8.40 | 0.18 | | 9816.53 | 7.73 | 0.25 | | 5986.32 | 7.47 | 0.19 | | 2992.2 | 7.13 | 0.32 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S03 ParkJoy** |  | **Rate** | **MOS** | **CI** | | **AVC** | 22322.8 | 7.33 | 0.38 | | 18948.9 | 6.20 | 0.22 | | 11051.6 | 4.20 | 0.31 | | 5104.96 | 2.00 | 0.29 | | **HEVC** | 15985.7 | 6.73 | 0.35 | | 9864.78 | 6.20 | 0.31 | | 5893.97 | 4.53 | 0.38 | | 2993.86 | 3.07 | 0.26 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S04 CheersHD4** |  | **Rate** | **MOS** | **CI** | | **AVC** | 22007.2 | 7.73 | 0.31 | | 19099.1 | 6.67 | 0.30 | | 10530.3 | 5.20 | 0.31 | | 5384.44 | 2.67 | 0.30 | | **HEVC** | 15822.9 | 7.20 | 0.37 | | 9791.44 | 7.00 | 0.29 | | 5882.89 | 6.07 | 0.29 | | 2993.11 | 3.47 | 0.25 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S05 MusicHD1** |  | **Rate** | **MOS** | **CI** | | **AVC** | 20742 | 7.20 | 0.24 | | 17507.4 | 7.07 | 0.32 | | 10459.7 | 5.80 | 0.27 | | 5709.11 | 4.27 | 0.22 | | **HEVC** | 15719.3 | 7.80 | 0.22 | | 9898.96 | 7.27 | 0.33 | | 5877.04 | 6.93 | 0.24 | | 2993.24 | 6.00 | 0.29 | |  |

Table 11 – Range Extensions random access 4:2:2 10 bit sequence test result summary

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S06 EBUHorse** |  | **Rate** | **MOS** | **CI** | | **AVC** | 28011 | 7.00 | 0.29 | | 9081.52 | 5.54 | 0.26 | | 5569.07 | 5.15 | 0.38 | | 2862.78 | 3.62 | 0.28 | | **HEVC** | 15623 | 8.15 | 0.38 | | 3948.02 | 8.46 | 0.39 | | 2159.66 | 6.85 | 0.37 | | 1382.33 | 5.77 | 0.32 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S07 EBUKidsSoccer** |  | **Rate** | **MOS** | **CI** | | **AVC** | 15177.7 | 7.33 | 0.19 | | 8713.2 | 7.13 | 0.30 | | 5387.76 | 6.87 | 0.30 | | 2645.91 | 4.20 | 0.22 | | **HEVC** | 7954.18 | 8.53 | 0.22 | | 4964 | 8.27 | 0.25 | | 2931 | 7.07 | 0.22 | | 1479.96 | 4.20 | 0.17 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S08 EBUWaterRocks** |  | **Rate** | **MOS** | **CI** | | **AVC** | 13338 | 6.67 | 0.25 | | 8769.24 | 4.07 | 0.28 | | 5748.69 | 3.00 | 0.19 | | 2997.91 | 1.13 | 0.19 | | **HEVC** | 7977.96 | 6.47 | 0.19 | | 4981.78 | 4.60 | 0.21 | | 2973.97 | 4.53 | 0.27 | | 1497.77 | 1.47 | 0.19 | |  |

Table 12 – Range Extensions random access 4:4:4 10 bit sequence test result summary

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S09 BirdsInCage** |  | **Rate** | **MOS** | **CI** | | **AVC** | 21087.89 | 7.60 | 0.18 | | 17317.23 | 7.27 | 0.31 | | 11039.36 | 7.13 | 0.21 | | 5741.45 | 7.20 | 0.31 | | **HEVC** | 15762.96 | 7.93 | 0.25 | | 9690.543 | 7.87 | 0.23 | | 5842.628 | 7.60 | 0.24 | | 2957.302 | 7.40 | 0.30 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S10CrowdRun** |  | **Rate** | **MOS** | **CI** | | **AVC** | 23419.15 | 7.87 | 0.23 | | 18369.3 | 7.53 | 0.23 | | 10232.68 | 7.40 | 0.27 | | 5463.21 | 7.20 | 0.27 | | **HEVC** | 15577.83 | 8.40 | 0.18 | | 9816.526 | 7.73 | 0.25 | | 5986.322 | 7.47 | 0.19 | | 2992.198 | 7.13 | 0.32 | |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **S11 EBULupoCandle** |  | **Rate** | **MOS** | **CI** | | **AVC** | 22322.81 | 7.33 | 0.38 | | 18948.86 | 6.20 | 0.22 | | 11051.63 | 4.20 | 0.31 | | 5104.96 | 2.00 | 0.29 | | **HEVC** | 15985.65 | 6.73 | 0.35 | | 9864.779 | 6.20 | 0.31 | | 5893.974 | 4.53 | 0.38 | | 2993.864 | 3.07 | 0.26 | |  |

From these tests, it can be observed that HEVC, when used at half the bit rate as AVC, can generally provide a better encoding quality than AVC for both the interlaced-scan and Range Extensions sequences.

An Excel file accompanying this report provides the MOS BD-rate computation for each sequence under each testing condition. This file uses a template provided by the BBC with contribution JCTVC-U0183 [2].

Note that for the MOS BD-rate computation, if any two points have the same MOS value, the derivation of the MOS BD-rate fails. One such case exists amongst the results of the subjective test. For the “CBS Tennis” sequence in the RA configuration, two different bit rates resulted in the same subjective score of 6.267. As a consequence, the MOS BD-rate result could not be computed for this case.

Notwithstanding the above, the trend of improved subjective quality is clearly evident from the plots contained in the above tables.

The MOS BD-rate calculations indicate that for the interlaced-scan verification tests, nine MOS BD rate values out of 10 indicate a bit-rate reduction of greater than 50%. Similarly, for the Range Extensions verification tests, five MOS BD rate values out of six show a bit-rate reduction of greater than 50%.

# Discussion

Throughout the verification test preparation, there were some discussions about how to determine the target bit rates for each sequence used for the subjective testing. On one hand, bit rates similar to those used in typical industrial applications were desired. However, typical subjective test methods require common coding artefacts to be properly noticeable to provide meaningful subjective scores. As a consequence, it was necessary to reduce the target bit rates in some cases.

It can be expected that the coding efficiency can vary substantially between the low and high bit-rate coding cases. For example, some Range Extensions coding tools may become less efficient when QP values are increased due to having additional bit-rate overhead requirements. At lower bit rates (higher QP values), coding tools with lower bit-rate overhead are generally preferred. Such behavior may also be altered by the presence of a rate control scheme.

In the case of the presented results, a compromise was found by using bit rates that were as close as possible to those used in practical applications that would still showcase coding artefacts. However, it is expected that as video compression standards keep on improving, the typical application bit rates and those required for optimal subjective viewing may tend to differ even more. The video coding and testing community would benefit from further studies in this domain. For example, a methodology to identify the point at which a bitstream ceases to be ‘visually lossless’ could be beneficial when testing high bit-rate operating points.

# Conclusion

The results reported herein show that for HM encoded HEVC bitstreams, for all sequences tested in this activity, targeting a bit rate that was half that of the corresponding JM encoded AVC bitstreams in no case resulted in a subjective quality worse than the subjective quality of the JM bitstreams. That is to say, the subjective quality of the HM either matches or exceeds the subjective quality of the JM, despite the fact that the HEVC bitstreams were encoded at approximately half the bit rate of the corresponding AVC bitstreams. This conclusion applies to both interlaced-scan and Range Extensions (10 bit 4:2:2/4:4:4) test sequences.

It is also interesting to note that for the Range Extensions results, the viewers were able to more easily differentiate the high-bit-rate encoded streams than the lower-bit-rate ones.

# Acknowledgements

The JCT-VC thanks CBS Television and NBCU for providing professional interlaced-scan test sequences for these verification tests.

The JCT-VC also thanks Canon for encoding all bitstreams for this activity.

The JCT-VC would also like to thank the MPEG Test Chair, Vittorio Baroncini, for designing, preparing, and performing the subjective viewing tests in a dedicated laboratory environment.

The JCT-VC also thanks the six companies whose financial support allowed the completion of the verification tests of interlaced-scan coding capability and Range Extensions coding capability:

* + Cable Television Laboratories, Inc. (CableLabs)
  + CBS Corporation
  + Fraunhofer Heinrich Hertz Institute (HHI)
  + Fujitsu Laboratories Ltd.
  + Microsoft Corporation
  + Sony Corporation

# References

1. C. Rosewarne, V. Baroncini, A. Tourapis, G. Barroux, M. Naccari “Verification test plan for interlaced video and format range extensions” JCTVC-T1003, 20th JCT-VC Meeting, Geneva, CH, February 2015.
2. A. S. Dias, M. Mrak, “Visualisation of BD-rate interpolation curves for MOS and PSNR results” JCTVC-U0183, 21st JCT-VC Meeting, Warsaw, PL, June 2015.

1. *All the interlaced video verification viewing sessions were run with a display refresh rate of 60 fps; for this reason some deinterlaced frames of the sequence “ParkJoy” were repeated on the display so that the average displayed frame rate of this test sequence was kept at 50 fps.* [↑](#footnote-ref-1)
2. *All the Range Extensions verification viewing sessions were run at 50 fps due to limitations of the display technology; for this reason it was decided to play the sequence “BirdsInCage” at the frame rate of 50 fps, about 17% slower than its original frame rate.* [↑](#footnote-ref-2)
3. *For the “BirdsInCage” test sequence, the listed bit rates are based on the source frame rate of 60 fps rather than the displayed frame rate of 50 fps that was used in the test, as previously noted herein. If desired, the reported bit rates can be multiplied by 5/6 to adjust for the difference in the displayed frame rate.* [↑](#footnote-ref-3)