|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  26th Meeting: Geneva, CH, 12–20 January 2017 | Document: JCTVC-Z0006 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **JCT-VC AHG report: SCC extensions verification testing (AHG6)** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | Haoping Yu Futurewei Technologies  Vittorio Baroncini GBTech  Rajan Joshi Qualcomm  Shan Liu MediaTek  Xiaoyu Xiu InterDigital  Jizheng Xu Microsoft | Email: | haoping.yu@huawei.com  [baroncini@gmx.com](mailto:baroncini@gmx.com)  [rajanj@qti.qualcomm.com](mailto:rajanj@qti.qualcomm.com)  [Shan.Liu@mediatek.com](mailto:Shan.Liu@mediatek.com)  [Xiaoyu.Xiu@InterDigital.com](mailto:Xiaoyu.Xiu@InterDigital.com)  [jzxu@microsoft.com](mailto:jzxu@microsoft.com)    - |
| *Source:* | AHG6 | | |

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

# Abstract

This report summarizes the activities of the JCT-VC ad hoc group on SCC extensions verification testing (AHG6) between the 25th JCT-VC meeting in Chengdu, China, and the 26th JCT-VC meeting in Geneva, Switzerland.

# Mandates

* Study test conditions and coding performance analysis methods for verification of SCC coding performance.
* Finalize the verification test plan for SCC Y1006

# Develop and propose further improvements of the test plan

# Activities

## Email reflector activity

The kick-off message for AHG 6 was sent out on Dec. 9, 2016.

There were more than 100 emails exchanged among the participants of the tests. These emails discussed the subjects relating to test sequences, encoder configurations, bit-stream generation and sharing and selection, performance evaluation and comparison, subjective testing, etc.

## SCC verification test plan

A document detailing the test plan (JCTVC-Y1006) was submitted. It describes a set of test conditions and presents a work plan for test preparation It also provides a DCR-based procedure for subjective evaluation.

The test plan proposes the following test condition:

* Test material

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Resolution** | **Sequence name** | **Category** | **fps** | **Frames to be encoded** |
| 1920x1080 | CircuitLayoutPresentation  ClearTypeSpreadsheet  EnglishDocumentEditing  ChineseDocumentEditing  BigBuckBunnyStudio  KristenAndSaraScreen | TGM  TGM  TGM  TGM  M  M | 30  30  30  30  50\*  60 | 0-239  0-239  0-239  0-239  0-399  0-479 |
| \*Note that this sequence was captured at 60fps but it is tested at 50fps to provide adequate visual duration.  TGM: Text and graphics with motion; M: mixed content; | | | | |

* Software:
* SCM-8.1 is used to generate both HEVC and HEVC-SCC bitstreams. When generating HEVC bitstreams, all the new coding tools adopted in the specifications of HEVC SCC extensions will be disabled. The software is available at <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.10+SCM-8.1>
* JM-19.0: <http://iphome.hhi.de/suehring/tml/download/>
* Coding modes, color space and sampling formats
* Lossy and mathematically lossless
* All Intra (AI), Random Access (RA), and Low-delay (LB)
* RGB, YUV-4:4:4, and YUV-4:2:0
* Test points: a formal subjective evaluation will be conducted by comparing bitstreams from the three encoders at 4 different QP values on all Lossy coding conditions. The final QP values will be selected based on the actual encoding results.

## Bitstreams for SCC verification test

### Status

For the lossy coding conditions described above, thousands of bitstreams have been generated by using all integer QP values between 10 and 47. We even have extra JM and HEVC bitstreams with QP values between 1 and 10, and SCM bitstreams with QP values between 47 and 51. Many bitstreams have been uploaded to the bitstream sharing ftp site.

### Issues with the bitstream selection for subjective testing

* Bit-rates of JM bitstreams and SCC bitstreams are far from each other. As a result, the bit-rate matching approach usually used in subjective testing may not be feasible for some test points in this SCC verification test, which is demonstrated by Figure 1 and 2 below. These two figures show the RD results of the EnglishDocumentEditing RGB sequence coded in AI and RA mode by the three encoders.

Figure 1 EnglishDocumentEditing RGB sequence coded in AI mode

Figure 2 EnglishDocumentEditing RGB sequence coded in RA mode

* Visual quality impairment due to compression is hard to see even at relative low PSNR values. Figure 3 below shows the ClearTypeSpreadsheet RGB sequence compressed by SCM-8.1 in LB mode with QP=37 at PSNR=36dB. The visual quality of the decoded video looks rather acceptable. Figure 4 shows the result from QP=41 at PSNR=30dB, where the compressed artifacts appear in the red circles.
* Note: the 10-dB PSNR jump from QP=15 to QP=16 in Figure 1 may indicate some issues in JM-19.0.
* Suggestion: we may use PSNR matching approach in selecting bitstreams for subjective testing. Basically, instead of selecting AVC and HEVC and HEVC-SCC bitstreams with close bitrates, we can select the bitstreams by matching the PSNR values of the decoded videos from the three compression technologies. The viewers in the subjective test should give close scores for the three bitstreams in each set. However, the bit-rates of the three bitstreams reflect the relative compression performance (e.g. bit-rate reduction) in the sense of subjective measurement.

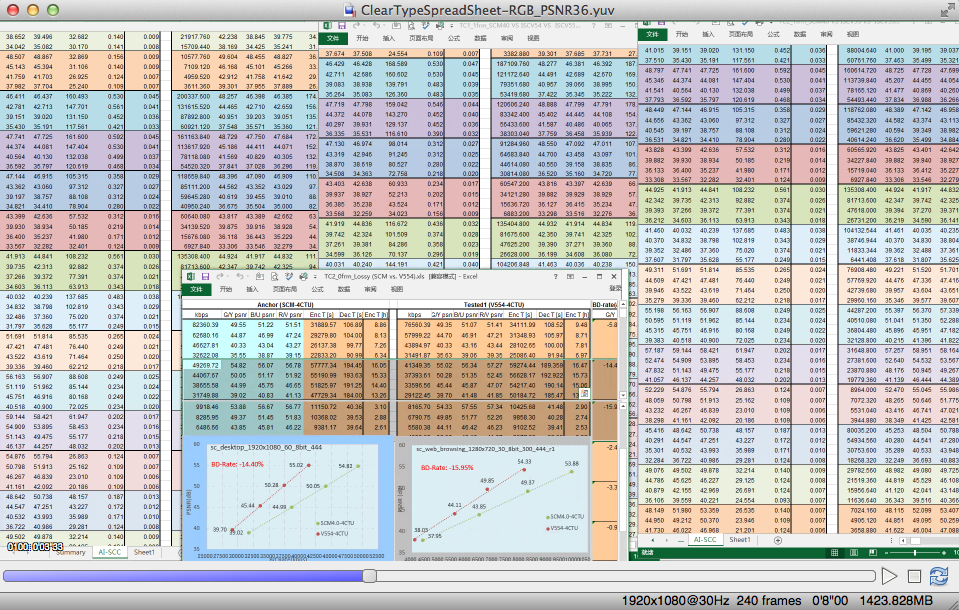


Figure 3 ClearTypeSpreadSheet RGB sequence coded by SCM-8.1 at QP=36, PSNR=36dB

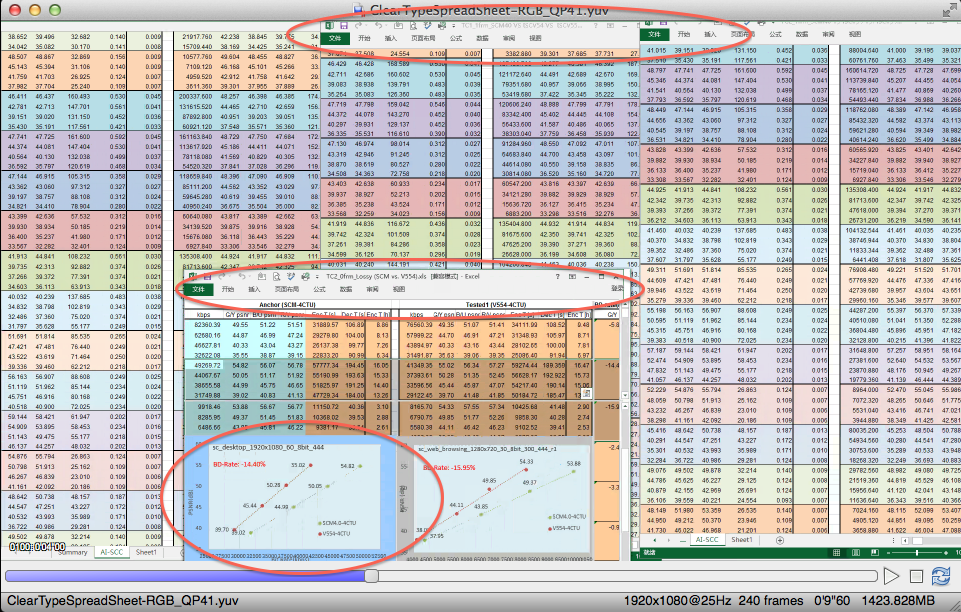


Figure 4 ClearTypeSpreadSheet RGB sequence coded by SCM-8.1 at QP=41, PSNR=30dB

## Subjective Testing

The subjective testing was done in Rome at the GBTech laboratory, during the week before the Geneva 2017 JCT-VC meeting. Originally the schedule was to complete the test by January 2017 5th, but an additional up-load of bit-stream was required to try to optimize the visual assessment.

This led also to a very long analysis of many additional decoded bit-streams that required much more time than what originally estimated.

Due to the above situation, the Test Chair (in agreement with the members of the AhG) reduced to two the test cases and to four the test sequences considered. The color schemes considered where YUV-444 and RGB, discarding the YUV-420.

The YUV-444 subjective evaluation experiment was done using four test sequences, i.e. ChineseDocumentEditing, EnglishDocumentEditing, ClearTypeSpreadSheet and CircuitLayoutPresentation.

The RGB subjective evaluation experiment was done using other four test sequences, i.e. BigBuckBunnyStudio, EnglishDocumentEditing, ClearTypeSpreadSheet and KristenAndSaraScreen.

In this way two test sequences were used for both the color spaces and result might be crosschecked.

The laboratory set-up was done using an improved version of the SW player MUP (version 7) that was able to play also RGB planar video files (not used in visual assessment so far), running on two identical PC equipped with high speed SSD drives in Raid 0 configuration and last generation X99motherboards with i7-6850 Intel CPUs and 64G of DDR-4 RAM .

Such configuration allowed a quick decoding of the bit-streams and a smooth presentation of the video clips.

Two TV sets where used as monitors. i.e. the LG OLED B6 (55” plane) and Samsung 55KS7500 (44” curve); these TV sets were selected due to their ability to present the images with a real low black level; all local post processing features were disabled to avoid an non faithful presentation of the images; it has to be noted that it was necessary to decrease the value of the backlight level of the Samsung TV set, to avoid visual stress in the viewing subjects, in particular when the sequence CircuitLayoutPresentation was presented.

Three subjects were seated in front of each display and they run the eight test sequences changing any time the kind of display; in other worlds a group of three subjects watched four test sessions on the OLED and the other four test sessions on the Samsung.

To control the level of the stress and fatigue each group of three subjects did not worked in total more than half a day, this means that a total of 12 subjects worked every half day of test: while six subjects where working, the other six where having rest. The complete testing of the four YUV444 and of the four RGB test sessions required the participation of a total of 48 subjects to four days of test.

Each subject was pre-screened for visual acuity and color blindness (Snellen Charts and Ishihara tables) and post-screened for consistency of their individual results to the general data set.

The tables in annex I provided results of testing experiments in form of tables and graphs.

## Related contributions

**JCTVC-Y1006: Verification test plan for HEVC screen content coding extensions [H. Yu, V. Baroncini, R. Joshi, S. Liu, X. Xiu, J. Xu]**

This contribution provides a test plan for verification of the coding performance of HEVC screen content coding extensions. It describes a set of test conditions and presents a work plan for test preparation.

# Recommendations

It is recommended to accomplish the following tasks during the 26th JCTVC meeting:

* select bitstreams and finalize the timeline for subjective testing
* discuss remaining issues