|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extensions of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  13th Meeting: Geneva, CH, 17–21 Oct 2015 | Document: JCT3V-M1001 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-HEVC Verification Test Report** | | |
| *Status:* | Approved | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | Vittorio Baroncini, FUB  Karsten Müller, HHI  Shinya Shimizu, NTT | Email: | [vittorio@fub.it](mailto:vittorio@fub.it)  [karsten.mueller@hhi.fraunhofer.de](mailto:karsten.mueller@hhi.fraunhofer.de)  [shimizu.shinya@lab.ntt.co.jp](file:///C:\Users\tech\AppData\Local\Temp\shimizu.shinya@lab.ntt.co.jp) |
| *Source:* | JCT-3V | | |

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

1. **Introduction**

This document describes the results of the subjective quality assessment test for HEVC-based coding of multiview video and depth. A description of the test conditions and subjective testing methodology is also provided.

1. **Test Conditions**
   1. ***Input and output test sequences***

The multiview test sequences with associated depth data, and corresponding input views that were used for the experiments are specified in the table below. For both coding methods, 3D-HEVC and MV-HEVC, a set of stereo pairs were generated as given in the table below. The stereo pair consists of 2 synthesized views.

|  |  |  |  |
| --- | --- | --- | --- |
| **Seq. ID** | **Test Sequence** | **Input Views** | **Output Stereo #1**  **(2 synthesized views)** |
| S01 | Poznan\_Hall2 | 7-6-5 | (6.25-5.75) |
| S02 | Poznan\_Street | 5-4-3 | (4.25-3.75) |
| S03 | Undo\_Dancer | 1-5-9 | (4-6) |
| S04 | GT\_Fly | 9-5-1 | (6-4) |
| S05 | Kendo | 1-3-5 | (2.5-3.5) |
| S06 | Balloons | 1-3-5 | (2.5-3.5) |
| S08 | Newspaper1 | 2-4-6 | (3.5-4.5) |
| S10 | Shark | 1-5-9 | (4.0-6.0) |

The view synthesis algorithm that was used for generating synthesized views is the “VSRS-1D-Fast” software (<https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware>).

* 1. ***Encoder Configuration***

Two HEVC-based codec configurations were assessed in the subjective test:

* MV-HEVC: anchor codec for multiview and depth without block-level changes to decoding process
* 3D-HEVC: enhanced compression of multiview and depth with modifications to block-level decoding process for dependent texture views

The encoder configuration settings for both codecs are consistent with the CTC which are outlined below:

* Inter-view coding structure
  + 3 view case: center-left-right (in coding order)
  + P-I-P inter-view prediction
* Temporal prediction structure: GOP 8, intra every 24 frames (random access at ~1sec)
* Anchor software: HTM-15.1 (<https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware>), modified to enable VSO for MV-HEVC.
* View Synthesis Optimization (VSO) enabled.

For the subjective comparison, 4 rate points were produced with the following Texture QPs for the independent view: 45, 40, 35, and 30 (according to the most recent CTC document N12352 that recommends QPs for subjective viewing). For each rate point, a bit stream was produced for both coding methods, such that PSNR values of enhancement texture views and synthesized views are equal for 3D-HEVC and MV-HEVC with depth map coding. Accordingly, 2 different bit rates were obtained, as given in the table below:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test Sequence** | **MV-HEVC bit rates (kbps)** | | | | | **3D-HEVC bit rates (kbps)** | | | | |
| R1 | R2 | R3 | R4 | R1 | | R2 | R3 | R4 |
| S01: Poznan\_Hall2 | 126 | 215 | 388 | 711 | 98 | | 171 | 308 | 603 |
| S02: Poznan\_Street | 199 | 379 | 770 | 1669 | 171 | | 328 | 651 | 1444 |
| S03: Undo\_Dancer | 378 | 758 | 1553 | 3240 | 285 | | 591 | 1261 | 2778 |
| S04: GT\_Fly | 296 | 601 | 1237 | 2632 | 228 | | 469 | 974 | 2138 |
| S05: Kendo | 193 | 332 | 579 | 1080 | 146 | | 253 | 443 | 827 |
| S06: Balloons | 184 | 333 | 579 | 1069 | 149 | | 264 | 458 | 840 |
| S08: Newspaper1 | 184 | 329 | 618 | 1156 | 151 | | 266 | 477 | 923 |
| S10: Shark | 414 | 864 | 1783 | 3521 | 307 | | 646 | 1342 | 2786 |

For each rate point and test sequence, the bit rate reduction of 3D-HEVC over MV-HEVC is given in the table below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Test Sequence** | **Bit rate difference ( 3D − MV ) / MV [ % ]** | | | | | |
| R1 | R2 | R3 | R4 | Average |
| S01: Poznan\_Hall2 | −22,2 | −20,5 | −20,6 | −15,2 | −19,6 |
| S02: Poznan\_Street | −14,1 | −13,5 | −15,5 | −13,5 | −14,2 |
| S03: Undo\_Dancer | −24,6 | −22,0 | −18,8 | −14,3 | −19,9 |
| S04: GT\_Fly | −23,0 | −22,0 | −21,3 | −18,8 | −21,3 |
| S05: Kendo | −24,4 | −23,8 | −23,5 | −23,4 | −23,8 |
| S06: Balloons | −19,0 | −20,7 | −20,9 | −21,4 | −20,5 |
| S08: Newspaper1 | −17,9 | −19,1 | −22,8 | −20,2 | −20,0 |
| S10: Shark | −25,8 | −25,2 | −24,7 | −20,9 | −24,2 |
| Average | −21,4 | −20,9 | −21,0 | −18,5 | −20,4 |

1. **Evaluation Procedure**

The different codec configurations were evaluated through formal subjective testing on stereoscopic displays. The evaluation was performed based on the output stereo views as specified in the table of section 2.3.

* 1. ***Evaluation Methodology***

The Absolute Category Rating test method (ACR) was used with 11 quality levels, where 10 indicates the highest quality and 0 indicates the lowest quality. The tests were carried out with expert viewers. The ACR method contains a Basic Test Cell (BTC) timing as shown in figure below.

|  |  |  |
| --- | --- | --- |
| Video 1-A | Video 1-B | Vote  1A + 1B |

***ACR test method, Basic Test Cell***

In a BTC, video A and video B are the same test sequence coded at approximately same PSNR, and using MV-HEVC for video A and 3D-HEVC for video B.

The message showing the sentence “Vote” is shown for 6 seconds. The subjects are asked to judge the overall quality of the videos A and B for evaluation. The test subjects were allowed to express their vote during that time period only.

The quality votes were expressed by the test subjects using paper scoring sheets.

The scoring sheets show one numbered box for each BTC (Basic Test Cell); the viewing subjects were instructed to write a number score (from 0 to 10) in the boxes for the observed videos on the screen (see example below).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***Test 1*** | |  | ***Test 2*** | |  | ***Test 3*** | |
| ***A*** | ***B*** | ***A*** | ***B*** | ***A*** | ***B*** |
|  |  |  |  |  |  |

***ACR Scoring Sheet***

The numbering of the boxes allow for easy synchronization with the message present on the screen to avoid uses of wrong scoring boxes.

The whole test was organized in separate test sessions. At the beginning of each test session, a sequence with very low quality and a sequence with very high quality were presented. This allowed subjects to familiarize with the full quality range to be used during the test.

To reduce contextual effects, the presentation order of the BTCs was randomized, in particular by applying different permutation for each (group of) subject(s). Moreover, the mapping of MV-HEVC and 3D-HEVC coded sequences to videos A and B was permuted.

Considering the particular care that a 3D visual test requires, the length of the test sessions did not exceed 15 minutes. The length and the design of the test sessions were determined by the Test Coordinator in agreement with other experts.

* 1. ***Raw Data Processing***

Mean Opinion Scores (MOS), Standard Deviation (SD) and Confidence Interval (CI) with 95% precision were computed on the raw data. Overlapping of CI values, centred on the MOS values, indicate that no Statistical Significant Difference exists between two test points.

* 1. ***Test subjects and equipment***

The tests were conducted using expert subjects only. A 46” JVC stereo display with passive glasses was used for the subjective evaluation and the tests were conducted with 20 subjects.

1. **Results**

The results of the subjective tests are provided in this section.

* 1. ***Graphical Results***









1. **Conclusions**

The subjective test results show that the 3D-HEVC codec achieves comparable quality relative to the MV-HEVC codec with approximately 20% less bit rate, on average (see table on bit rate difference in Section 2). This also confirms the objective coding performance in PSNR increase for 3D-HEVC vs. MV-HEVC.