|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  5th Meeting: Vienna, AT, 27 July-2 Aug. 2013 | Document: JCT3V-E1100 |

|  |  |  |
| --- | --- | --- |
| *Title:* | **Common Test Conditions of 3DV Core Experiments** | |
| *Status:* | Output Document | |
| *Purpose:* | Common Test Conditions | |
| *Author(s)):* | Dmytro Rusanovskyy  Karsten Müller Anthony Vetro | [dmytro.rusanovskyy@ieee.org](mailto:dmytro.rusanovskyy@ieee.org)  [karsten.mueller@hhi.fraunhofer.de](mailto:karsten.mueller@hhi.fraunhofer.de) [avetro@merl.com](mailto:avetro@merl.com) | |
| *Source:* | JCT-3V | |

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

This document specifies the common test conditions for 3D video core experiments with AVC-based and HEVC-based test models.

# Test Cases

There are two general test scenarios under consideration:

* multiview/stereo video coding with depth data
* multiview/stereo video coding without depth data

For each test scenario, experiments consider a two-view case (C2) and a three-view case (C3), where C2 is subset of C3 coding configuration. Simulations are run under the C3 configuration and the corresponding results are reported for C3. The C2 results are derived from the C3 results using the recommended Excel template.

Proponents are encouraged to directly simulate the C2 configuration if they find it useful for their specific proposal.

# Test Sequences

The multiview test sequences with associated depth data, and corresponding input and output views, to be used for experiments are specified in the table below. For accessibility of test sequences see section 8.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Seq. ID** | **Test Sequence** | **Frames to be encoded** | **2-view input** | **3-view input** |
| S01 | Poznan\_Hall2 | 200 | 7-6 | 7-6-5 |
| S02 | Poznan\_Street | 250 | 5-4 | 5-4-3 |
| S03 | Undo\_Dancer | 250 | 1-5 | 1-5-9 |
| S04 | GT\_Fly | 250 | 9-5 | 9-5-1 |
| S05 | Kendo | 300 | 1-3 | 1-3-5 |
| S06 | Balloons | 300 | 1-3 | 1-3-5 |
| S08 | Newspaper1 | 300 | 2-4 | 2-4-6 |
| S10 | Shark | 300 | 1-5 | 1-5-9 |

# Basic Encoder Configuration

The encoder configuration settings that are common to both HEVC- and AVC-based experiments are given below:

* Inter-view coding structure
  + 3 view case: center-left-right (in coding order)
  + 2 view case: right-left (in coding order)
  + P-I-P inter-view prediction
* Temporal prediction structure
  + GOP 8
  + intra every 24 frames (random access at roughly each second)

Specific encoder configurations for AVC-based experiments are specified as follows:

* Anchor software: 3DV-ATM v9.0
* <http://mpeg3dv.research.nokia.com/svn/mpeg3dv/tags/>
* Encoder HP/EHP configurations are provided as part of the “Starter-kit”
  + ./configs
* Changes to encoder configurations compare to previous configuration:
  + EHP profile:
    - None
  + HP profile
    - None
* Independent view texture QP values: 26, 31, 36, 41 (Full resolution)
* Independent view depth QP values: 26, 31, 36, 41 (Reduced resolution)
* In the case, the CTC coding configuration is applied for coding of test sequences with full resolution depth map, the following depth map QP settings apply:
  + Base view depth QP values: 30, 35, 40, 45 (Full resolution)
* Full resolution texture coding
  + Check sum of input data to be provided in the “Starter-kit”
* Reduced resolution depth coding (50% in each direction)
  + Downsampling: JSVM utility, method1 (included in the Starter-kit)
  + Up-sampling: built-in 3DV-ATM (bilinear).
  + Depth map post-processing (JCT2-A0038/m25885)
  + Check sum of the input data to be provided in the “Starter-kit”

Specific encoder configurations for HEVC-based experiments are specified as follows:

* **MV-HEVC:** Anchor software for MV-HEVC-based experiments: 3DV-HTM v8.0
  + <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware>
* **3D-HEVC:** Anchor software for 3D-HEVC-based experiments: 3DV-HTM v8.0.
  + <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware>
  + Changes to encoding configuration compare to the previous version:
    - Fast encoder decision (texture) ON (E0173)
    - Fast encoder mode selection (DMM/RBC) ON (E0238)
    - Inter-view MAD prediction OFF (E0227)
* 8-bit input data to be used
* HEVC codecs are configured with 8-bit internal processing
* Encoder configuration files and renderer configuration provided with each software version
* Full resolution of video and depth (3D-HEVC only)
* Texture QP values for independent view: 40, 35, 30, 25
* Depth QP values: fixed relation to video QP as shown below (3D-HEVC only)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| QPV0 | 51 | 50 | 49 | 48 | 47 | 46 | 45 | 44 | 43 | 42 | 41 | 40 | 39 | 38 | 37 | 36 | 35 | 34 | 33 | 32 | 31 | 30 | 29 | 28 | 27 | 26 | 25 |
| QPD0 | 51 | 50 | 50 | 50 | 50 | 49 | 48 | 47 | 47 | 46 | 45 | 45 | 44 | 44 | 43 | 43 | 42 | 42 | 41 | 41 | 40 | 39 | 38 | 37 | 36 | 35 | 34 |

# Objective evaluation of video quality

* A new reporting template (NEW\_TEMPLATE\_VIENNA.xls) is provided in the zip file of the CTC document, including the Shark sequence. Note that the shark sequence is not taken for the average values.
* View synthesis distortion (PSNR) computed against reference view synthesis results (R0)
* Reference view synthesis (R0) is produced with original texture and depth map data, all at full resolution, original camera parameters information
* Check sums for R0 yuv files to be provided with “Starter-Kit”.
* Use 3 intermediate views between two cameras (evenly distributed), configuration for VS to be provided in the “Starter-kit”.
* View synthesis uses coded camera parameters information.
* Bjontegard deltas calculated from average PSNR and bitrate for each MVD component of the coded views independently
* Bjontegard deltas calculated from average PSNR for the synthesized views and the overall bit rate (texture, depth and supplementary data)
* Bjontegard deltas calculated from average PSNR for coded texture and the overall bit rate (texture, depth and supplementary data)

# Subjective evaluation of visual quality

Subjective evaluation of proposed techniques relative to anchors generated according to the CTC may be performed with views specified in the table below, unless otherwise specified in the Core Experiment description.

|  |  |  |  |
| --- | --- | --- | --- |
| **Seq. ID** | **Test Sequence** | **Input views** | **Synthesized**  **Stereo pair** |
| S01 | Poznan\_Hall2 | 7-6-5 | (6.250-5.750) |
| S02 | Poznan\_Street | 5-4-3 | (4.250-3.750) |
| S03 | Undo\_Dancer | 1-5-9 | (4.0-6.0) |
| S04 | GT\_Fly | 9-5-1 | (6.0-4.0) |
| 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) |

# Data to be provided

The following data must be provided as part of the CE report, as also contained in the updated report template:

* rate-distortion performance relative to the anchor
  + multiview coding without depth:
    - PSNR for each view
    - rate for each view
  + multiview coding with depth:
    - rate for each coded view of each MVD component
    - rate for any additional data, e.g. camera params (if applied)
    - PSNR for coded views
    - PSNR for intermediate views (see objective measurement)
    - BD rates as defined above(see objective measurement)
* Encoder and decoder run times in comparison to anchor
* Decoder run time is collected with disabled writing YUV data to the disk
* If new rendering is tested: renderer run time in comparison to anchor

# View Synthesis Algorithm

The view synthesis algorithm to be used for both AVC and HEVC experiments is the “VSRS-1D-Fast” software based on the version used in core experiments for HTM development. (<https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware>)

The view synthesis used to generate reference results (R0) shall be based on the HTM rendering software using version 3.0 or later.

# Availability of Test Sequences

## S01: Poznan\_Hall2

## S02: Poznan\_Street

**Poznan University of Technology**

<ftp://multimedia.edu.pl/3DV/>  
username: 3DV / password: ftvftv

directory: CFP

*Copyright: Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing and promulgating technology standards and for academic usage. Acknowledgement and reference to the following source document: Marek Domañski, Tomasz Grajek, Krzysztof Klimaszewski, Maciej Kurc, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner "Poznañ Multiview Video Test Sequences and Camera Parameters", ISO/IEC JTC1/SC29/WG11 MPEG 2009/M17050, Xian, China, October 2009, are required in all documents that report any usage of the materials. The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.*

Owners: Poznan University of Technology, Poznañ, Poland.

## S03: Undo\_Dancer

## S04: GT\_Fly

**Nokia**

<ftp://mpeg3dv.research.nokia.com>

username: mpegmember / password: S9"12#sHD)3

*Copyright: The supplied data and content (“Supplied Data”) is provided free of charge and made available for use by the Licensee, who shall be a member of the MPEG standardization committee or a respondent to a the Call for Proposals for the MPEG standard for 3D video coding that is to be issued March 2011, under the following conditions:*

*The Licensee agrees that the Supplied Data and all intellectual property rights therein remain the property of Nokia Corporation and its licensors (owners). The Supplied Data may only be used for the purpose of responding to the Call for Proposals that is to be issued March 2011, developing, testing and promulgating the MPEG technology standard for 3D video coding resulting from the Call for Proposals that is to be issued March 2011. The respective owners make no warranties with respect to the Supplied Data and expressly disclaim any warranties regarding its fitness for any purpose.*

*As a way of promulgating the MPEG technology standard for 3D video coding, the Licensee may present parts or modifications of the Supplied Data at academic conferences and publications.*

*The Licensee agrees not to provide the data to any third parties without permission from the owners and that the data shall not be sold, let for hire, or by way of trade, offered or exposed for sale or hire. This restriction shall apply to the original material or to any reproduction of it in whole or in part and to any modifications of the Supplied Data.*

*The texture views of the Supplied Data are copyright © UNDO. (See* [*http://www.undo.fi*](http://www.undo.fi)*)*

Owners: Nokia, Finnland.

## S05: Kendo

## S06: Balloons

**Nagoya University**

<http://www.tanimoto.nuee.nagoya-u.ac.jp/~mpegftv/mpeg3dv/CfP/>

username: mpegftv / password: fngOyfTv

*Copyright: Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing and promulgating technology standards. The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.*

*Owners: Nagoya University, Japan.*

## S08: Newspaper1

**GIST**

<ftp://203.253.128.142>

username: 3DV / password: 3dvkr

directory: /GIST\_Test\_Sequence/Newspaper

*Copyright: Individuals and organizations extracting sequences from this archive agree that the sequences and all intellectual property rights therein remain the property of the respective owners listed below. These materials may only be used for the purpose of developing, testing and promulgating technology standards. The respective owners make no warranties with respect to the materials and expressly disclaim any warranties regarding their fitness for any purpose.*

*Owners: Gwangju Institute of Science and Technology (GIST), Republic of Korea.*

## S10: Shark

**NICT**

ftp://ftp.merl.com

username: anonymous password: <email>

directory: /pub/tian/NICT-3D/Shark

*Copyright: The supplied data and content (“Supplied Data”) is provided free of charge and made available for use by the Licensee, who shall be a member of the JCT-3V standardization committee, under the following conditions:*

*The Licensee agrees that the Supplied Data and all intellectual property rights therein remain the property of the National Institute of Information and Communications Technology (NICT) and its licensors (owners). The Supplied Data may only be used for the purpose of developing, testing and promulgating the JCT-3V technology standards and for academic usage. The respective owners make no warranties with respect to the Supplied Data and expressly disclaim any warranties regarding its fitness for any purpose.*

*As a way of promulgating the technology standards, the Licensee may present parts or modifications of the Supplied Data at academic conferences and publications.*

*In the use of the parts or modifications of the Supplied Data, the Licensee shall make the source clear in appropriate manners.*

*The Licensee agrees not to provide the data to any third parties without permission from the owners and that the data shall not be sold, let for hire, or by way of trade, offered or exposed for sale or hire. This restriction shall apply to the original material or to any reproduction of it in whole or in part and to any modifications of the Supplied Data.*

*Owners: National Institute of Information and Communication Technology (NICT), Japan*

# Appendix A: Configuration Settings for View Extrapolation

This section describe configuration settings intended for experiments on view extrapolation, which are provided for reference and are not part of the CTC for all experiments.

In the extrapolation experiments, the input views and output view positions that should be synthesized are specified in the table below.

* Interpolated and Extrapolated View

|  |  |  |  |
| --- | --- | --- | --- |
| **Seq. ID** | **Test Sequence** | **Input views** | **Stereo pairs** |
| S01 | Poznan\_Hall2 | 7-6 | (6.25-5.75) |
| S02 | Poznan\_Street | 5-4 | (4.25-3.75) |
| S03 | Undo\_Dancer | 1-5 | (4.0-6.0) |
| S04 | GT\_Fly | 9-5 | (6.0- 4.0) |
| S05 | Kendo | 1-3 | (2.50-3.50) |
| S06 | Balloons | 1-3 | (2.50-3.50) |
| S08 | Newspaper1 | 2-4 | (3.50-4.50) |
| S10 | Shark | 1-5 | (4.0-6.0) |

* Extrapolated View

|  |  |  |  |
| --- | --- | --- | --- |
| **Seq. ID** | **Test Sequence** | **Input views** | **Stereo pairs** |
| S01 | Poznan\_Hall2 | 7-6 | (6.0-5.5), (5.875-5.375) |
| S02 | Poznan\_Street | 5-4 | (4.0-3.5), (3.875-3.375) |
| S03 | Undo\_Dancer | 1-5 | (5.0-7.0), (5.5-7.5) |
| S04 | GT\_Fly | 9-5 | (5.0- 3.0), (4.5-2.5) |
| S05 | Kendo | 1-3 | (3.0-4.0), (3.25-4.25) |
| S06 | Balloons | 1-3 | (3.0-4.0), (3.25-4.25) |
| S08 | Newspaper1 | 2-4 | (4.0-5.0), (4.25-5.25) |
| S10 | Shark | 1-5 | (5.0-7.0), (5.5-7.5) |

The VSRS-1D-Fast software shall be used to generate the anchor views for comparison.

View synthesis should be carried out from coded and uncompressed data and the configuration files for generating the anchors, including the view synthesis software (VSRS-1D-Fast) as well as the 3D-ATM or 3D-HTM software, are provided as part of the “Starter-kit” package.