**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG/N**18089

**October 2018, Macau SAR, CN**

|  |  |
| --- | --- |
| **Source** | Video |
| **Status** | Approved |
| **Title** | Common Test Conditions on 3DoF+ and Windowed 6DoF |
| **Authors** | Joel Jung, Bart Kroon, Renaud Doré, Gauthier Lafruit, Jill Boyce |

# Introduction

Common test conditions are desirable to conduct coding experiments in a well-defined environment and ease the comparison of the outcome of experiments. This document specifies the common test conditions for 3DoF+ (limited amount of parallax) and Windowed-6DoF (considering a virtual window separating the viewer from the content) activities also called 6DoF for convenience in the rest of the document. The common test conditions are defined to evaluate the coding efficiency, subjective quality and user experience of 3DoF+ and 6DoF solutions.

The technical approach, common to both activities is following the steps:

1. Compress test content,
2. Synthesize intermediate views from decoded views and metadata,
3. Render viewports of real/virtual pose traces of limited head movement, typically sitting in a seat for 3DoF+, or wider movement for 6DoF,
4. Evaluate coding efficiency and parallax effect, considering both decoded views and synthesized views.

The bitstream shall be viewer independent, meaning that neither the position nor the orientation of the viewer shall be considered when compressing the test content. The range of supported possible viewer positions is constrained and known.

Both omnidirectional and perspective capture video content is included in these conditions. The 3DoF+ anchors and test conditions use single layer HEVC coding. The 6DoF anchors use MV-HEVC coding.

# Test material

This section describes the test material that is used in the common test conditions and defines the numbering of the views. References to input documents are included for a more detailed description of each sequence. Subsequent subsections provide download links for sequence data and metadata. CTC-specific configuration files are provided as attachments to this document, as reported in Table 23.

## Omnidirectional test material

This class of test material is omnidirectional, and stored as multiple views in equirectangular projection (ERP) format.

The *ClassroomVideo*, *TechnicolorMuseum* and *TechnicolorHijack* sequences have a common format as defined in the *Call for MPEG-I Visual Test Materials on 3DoF+ and 6DoF* [1], determining texture and depth representation, filenames and metadata. For this CTC, views are numbered according to the ordering of the metadata files, counting from zero. For 3DoF+, the view positions in the tables are rounded to millimeter accuracy or hundredth of degree, with exact positions and orientation available in the metadata files. The values in the metadata files should be considered the fully accurate values, and used in all experiments, rather than the values in the Word document.

Note that the view numbering exists sometimes in a matrix version for the rig description and an ordinal version for the json file. Annex 2 “Source View label conversion” gives the correspondence between the two.

The test material is provided as a collection of separate color and depth frames, but RVS and HM expect raw YUV 4:2:0 sequences per view and for texture and depth. Because the format of the omnidirectional sequences differ from YUV 4:2:0 10-bit, HDRTools scripts mentioned in §4.1.1 are used to convert to YUV 4:2:0 10-bit texture and YUV 4:2:0 10-bit depth. The per-frame output of HDRTools can be simply concatenated to form sequences. The 10-bit sequences are little-endian.

### ClassroomVideo

The general characteristics of the ClassroomVideo sequence are summarized in Table 1. Source view positions are according to a hexagonally-packed circular disc with an additional top and bottom view, as shown in Figure 1.

Table 1: Summary of the ClassroomVideo sequence

|  |  |
| --- | --- |
| **Input contributions** | [m42415](http://wg11.sc29.org/doc_end_user/current_document.php?id=61486&id_meeting=), m42756 and [m42944](http://wg11.sc29.org/doc_end_user/current_document.php?id=62216&id_meeting=) |
| **Length & frame rate** | 120 frames (30 fps) |
| **Number of source views** | 15 |
| **Texture format** | YUV 4:2:0 8-bit |
| **Depth format** | YUV 4:2:0 16-bit, normalized disparity in [0.8m, ) range |
| **Source view resolution** | 4096 × 2048 |
| **View FoV & mapping** | 360° × 180° ERP |
| **Global FoV** | 360° × 180° |
| **Download** | [http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent /MPEG-I/Philips/ClassroomVideo](http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Philips/ClassroomVideo)  Texture maps: ClassroomVideo\_v\*.gz  Depth maps: ClassroomVideo\_fulldepth\*.gz  Metadata file: ClassroomVideo.json [A13] |

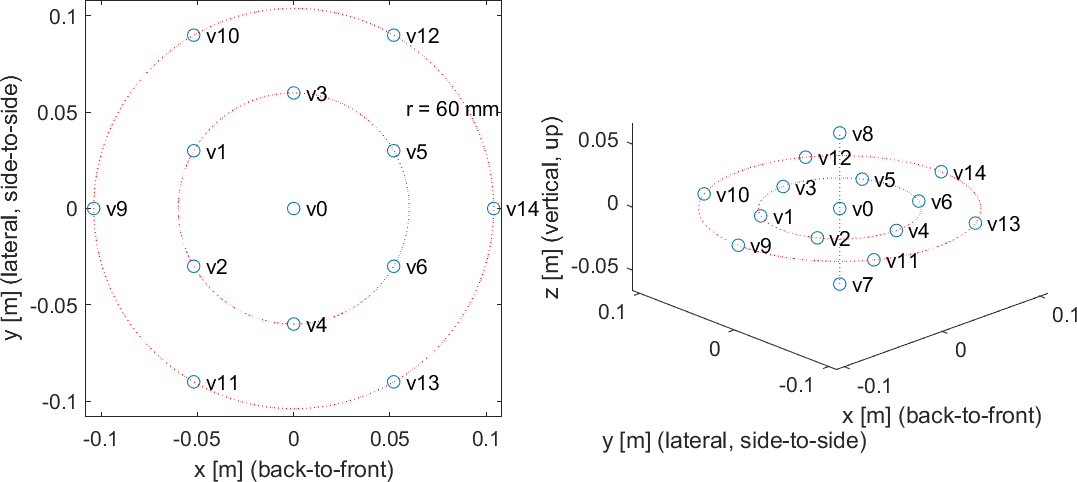


Figure 1: Visualization of the viewpoints of the ClassroomVideo sequence

### TechnicolorMuseum

The general characteristics of the TechnicolorMuseum sequence are summarized in Table 2. The cameras are disposed on a spherical surface of 30 cm radius, and divergent in the direction of the sphere radius. Figure 2 provides the (X, Y, Z) coordinates and the spherical dimension, with an example using the 11th view. The metadata file comprising source and intermediate view positions is attachment A12 to this output document.

The 1st column of Table 24 corresponds to the original view names in the uploaded content. In order to have name consistency between all content, the naming convention v0, v1, … has been used and is given in the 2nd column of Table 24.

Table 2: Summary of the TechnicolorMuseum sequence

|  |  |
| --- | --- |
| **Input contribution** | [m42349](http://wg11.sc29.org/doc_end_user/current_document.php?id=61510&id_meeting=) |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 24 |
| **Source view resolution** | 2048 × 2048 |
| **Texture format** | YUV 4:2:0 8-bit |
| **Depth format** | YUV 4:2:0 16-bit, normalized disparity in [0.5m, 25m] range |
| **View FoV & mapping** | 180° × 180° ERP |
| **Global FoV** | 360° × 180° |
| **Download** | [http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/ MPEG-I/Technicolor/TechnicolorMuseum](http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Technicolor/TechnicolorMuseum)  Texture and depth maps: f\*.zip  Metadata file: TechnicolorMuseum.json [A13] |

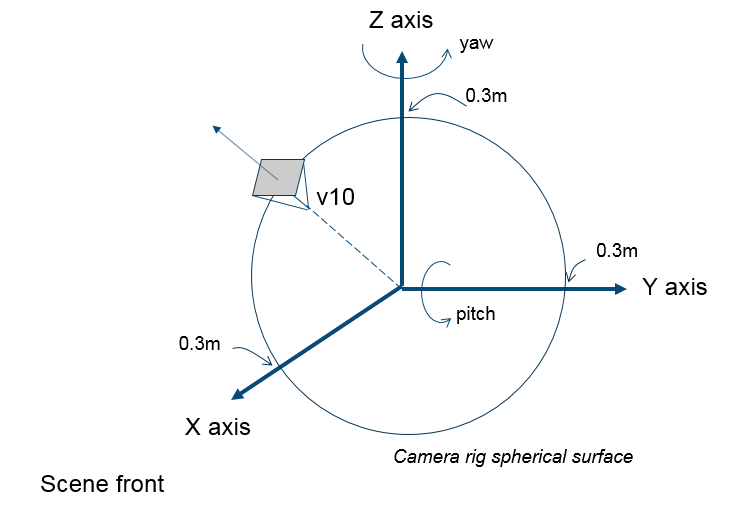


Figure 2: Coordinate system as used by 3D Audio, OMAF and the 3DoF+ investigation, with view 10 of the TechnicolorMuseum sequence superimposed

### TechnicolorHijack

The general characteristics of the TechnicolorHijack sequence are summarized in Table 3, and source view naming in Table 25. Figure 3 provides a visualization of the virtual camera rig in bias, top and front view respectively. The metadata file comprising source and intermediate view positions is in attachment [A13] of this output document.

The 1st column of Table 25 corresponds to the original view names in the uploaded content. The address of location is given below. In order to have name consistency between with all contents, the naming convention v0, v1, … has been used and is given in the 2nd column of Table 25.

Table 3: Summary of the TechnicolorHijack sequence

|  |  |
| --- | --- |
| **Input contribution** | [m42349](http://wg11.sc29.org/doc_end_user/current_document.php?id=61510&id_meeting=) |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 10 |
| **Source view resolution** | 4096 × 4096 |
| **Texture format** | YUV 4:2:0 8-bit |
| **Depth format** | YUV 4:2:0 16-bit, normalized disparity in [0.5m, 25m] range |
| **View FoV & mapping** | 180° × 180° ERP |
| **Global FoV** | 180° × 180° |
| **Download** | [http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent /MPEG-I/Technicolor/TechnicolorHijack](http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Technicolor/TechnicolorHijack)  Texture and depth maps: f\*.zip  Metadata file: TechnicolorHijack.json [A13] |

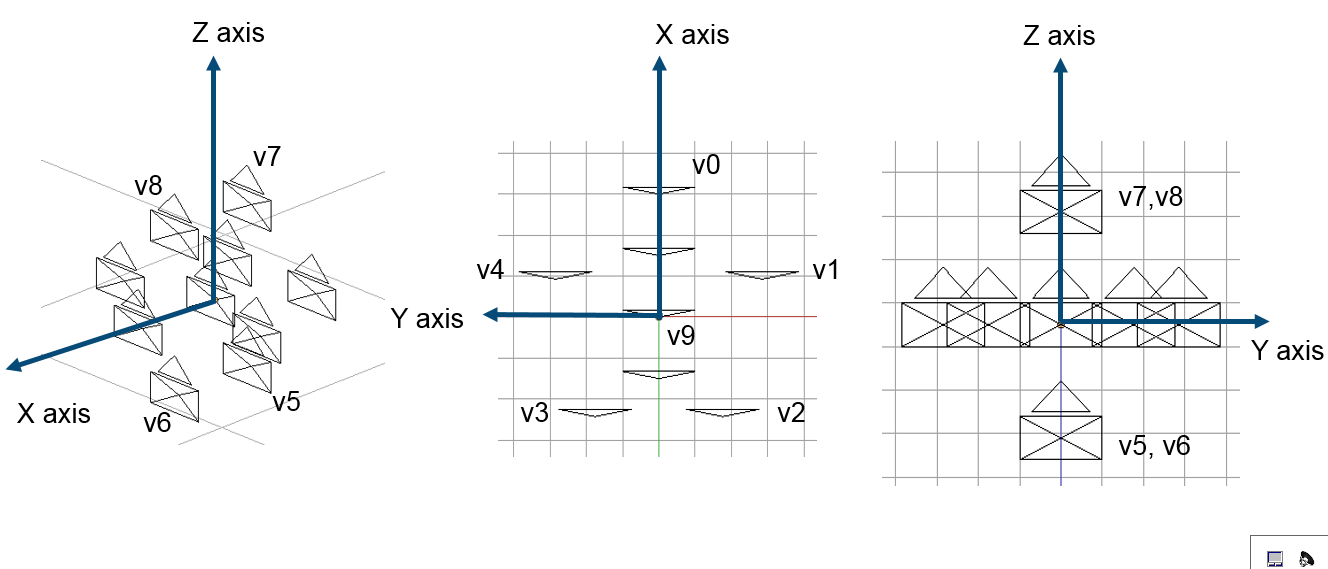


Figure 3: Visualization of the view positions of the TechnicolorHijack sequence

## Perspective test material

The sequences have a common format as defined in the *Call for MPEG-I Visual Test Materials on 6DoF* [2] determining texture and depth representations.

### TechnicolorPainter

The general characteristics of the TechnicolorPainter sequence are summarized in Table 4 and source view naming in Table 26 form a 4x4 planar array and are numbered v0-0 to v3-3 following left to right and top to bottom scan order, as shown in Table 5.

Table 4: Summary of the TechnicolorPainter sequence

|  |  |
| --- | --- |
| **Input contributions** | m40010, m40011 and m43366 |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 16 (4x4) |
| **Source view resolution** | 2048 × 1088 |
| **Download** | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Technicolor/TechnicolorPainter/pseudo-rectified>  Texture maps: source\_TechnicolorPainter\_pr2\_yuv.zip  Depth maps :DERS7\_Technicolorpainter\_depth\_pr2\_TE1\_420.zip  Metadata file: TechnicolorPainter.json [A13] |

Table 5: View numbering of the TechnicolorPainter camera array

|  |  |  |  |
| --- | --- | --- | --- |
| v0-0 | v1-0 | v2-0 | v3-0 |
| v0-1 | v1-1 | v2-1 | v3-1 |
| v0-2 | v1-2 | v2-2 | v3-2 |
| v0-3 | v1-3 | v2-3 | v3-3 |

### ULBUnicorn Plane A'

ULB Unicorn content is made of two planes (A’ and B’).

The general characteristics of the ULB Unicorn Plane A' sequence are summarized in Table 6. The captured views form a 9×9 planar array and are numbered v0-0 to v8-8 following left to right and top to bottom scan order, in a similar way as shown in Table 5 for TechnicolorPainter.

Table 6: Summary of the ULBUnicorn Plane A' sequence

|  |  |
| --- | --- |
| **Input contribution** | [m41083](http://wg11.sc29.org/doc_end_user/current_document.php?id=58741&id_meeting=) |
| **Length & frame rate** | 1 frame |
| **Number of source views** | 81 (9x9) |
| **Source view resolution** | 1920 × 1080 |
| **Download** | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/ULB/ULB_Unicorn/PLANES/Data_2018-08>  Texture maps: Plane\_A'\_Texture\_w\_Midpoints\_2018-08.zip  Depth maps: Plane\_A'\_Depth\_eDERS\_2018-08.zip  Metadata file(s): cams\_unicorn\_vsrs.txt |

### ULBUnicorn Plane B'

The general characteristics of the ULB Unicorn Plane B' sequence are summarized in Table 7. The captured views form a 9×5 planar array and are numbered v0-0 to v9-5 following left to right and top to bottom scan order.

Table 7: Summary of the ULBUnicorn Plane B' sequence

|  |  |
| --- | --- |
| **Input contribution** | [m41083](http://wg11.sc29.org/doc_end_user/current_document.php?id=58741&id_meeting=) |
| **Length & frame rate** | 1 frame |
| **Number of source views** | 45 (9x5) |
| **Source view resolution** | 1920 × 1080 |
| **Download** | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/ULB/ULB_Unicorn/PLANES/Data_2018-08>  Texture maps: Plane\_B'\_Texture\_w\_Midpoints\_2018-08.zip  Depth maps: Plane\_B'\_Depth\_eDERS\_2018-08.zip  Metadata file(s): cams\_unicorn\_vsrs.txt |

### OrangeShaman

The general characteristics of the OrangeShaman sequence are summarized in Table 8 and source view naming in Table 27. The captured views form a 5×5 planar array and are numbered v0-0 to v4-4 following left to right and top to bottom scan order.

Table 8: Characteristics of the OrangeShaman sequence

|  |  |
| --- | --- |
| **Input contribution** | m43318 |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 25 (5x5) |
| **Source view resolution** | 1920x1080 |
| **View FoV & mapping** | 77.3° × 48.5° Rectilinear |
| **Lens** | 20 mm |
| **Camera spacing** | 10cm x 10cm |
| **zNear** | 0.4 |
| **zFar** | 4.2 |
| **Download** | http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Orange/OrangeShaman  Texture maps: OrangeShaman\_x\*\*y\*\*.zip  Depth maps: OrangeShaman\_d\_x\*\*y\*\*.zip |

### OrangeDancing

The general characteristics of the OrangeDancing sequence are summarized in Table 9. The captured views form a 23×3 planar array and are numbered v0-0 to v22-2 following left to right and top to bottom scan order.

Table 9: Characteristics of the OrangeDancing sequence

|  |  |
| --- | --- |
| **Input contribution** | m43318 |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 42 (14 x 3) |
| **Source view resolution** | 1920x1080 |
| **View FoV & mapping** | 90° × 58.7° Rectilinear |
| **Lens** | 16 mm |
| **Camera spacing** | 3.9° along ellipse with rx=5 and ry=4 |
| **zNear** | 1.2 |
| **zFar** | 14.2 |
| **Download** | http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Orange/OrangeDancing  Texture maps: OrangeDancing\_x\*\*y\*\*.zip  Depth maps: OrangeDancing\_d\_x\*\*y\*\*.zip |

### OrangeKitchen

The general characteristics of the OrangeKitchen sequence are summarized in Table 10 and source view positions in Table 27. The captured views form a 5×5 planar array and are numbered v0-0 to v4-4 following left to right and top to bottom scan order.

Table 10: Characteristics of the OrangeKitchen sequence

|  |  |
| --- | --- |
| **Input contribution** | m43318 |
| **Length & frame rate** | 90 frames (30 fps) |
| **Number of source views** | 25 (5x5) |
| **Source view resolution** | 1920x1080 |
| **View FoV & mapping** | 53.1° × 31.4° Rectilinear |
| **Lens** | 32 mm |
| **Camera spacing** | 10cm x 10cm |
| **zNear** | 2.2 |
| **zFar** | 7.2 |
| **Download** | http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Orange/OrangeKitchen  Texture maps: OrangeKitchen\_x\*\*y\*\*.zip  Depth maps: OrangeKitchen\_d\_x\*\*y\*\*.zip |

### IntelKermit

The general characteristics of the IntelKermit sequence are summarized in Table 11 and source view positions in Table 28. The captured views form a 14x1 line and are numbered v0-0 to v13-0 following left to right scan order.

Table 11: Characteristics of the IntelKermit sequence

|  |  |
| --- | --- |
| **Input contribution** | m43748 and m44914 |
| **Length & frame rate** | 300 frames (30fps) |
| **Number of source views** | 15 (15x1) |
| **Source view resolution** | 1920x1080 |
| **View FoV & mapping** | 63.65° × 38.47° Rectilinear |
| **Lens** | 2.16 mm |
| **Camera spacing** | 3.675 cm |
| **zNear** | 0.3 |
| **zFar** | 1.62 |
| **Download** | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/Intel/Kermit>  TextureContent/vAll\_1920\_1080\_8bYUV420.tar.gz  DERS7\_DepthMaps/vAll\_1920\_1080\_0\_3\_1\_62d\_16bYUV400.tar.gz  Metadata file(s): IntelKermit.json [A13] |

### ETRIChef

The general characteristics of the ETRIChef sequence are summarized in Table 12 and source view positions in Table 29. The captured views form a 5×4 planar array and are numbered v0-0 to v4-3 following left to right and top to bottom scan order.

Table 12: Characteristics of the ETRIChef sequence

|  |  |
| --- | --- |
| **Input contribution** | m42542 |
| **Length & frame rate** | 300 frames (30 fps) |
| **Number of source views** | 20 (5x4) |
| **Source view resolution** | 1920x1080 |
| **Download** | <http://mpegfs.int-evry.fr/mpegcontent/ws-mpegcontent/MPEG-I/etri_chef>  Texture: Cam\*\*.yuv420 Depth: Depth/depth\_view\*\*.yuv420y  \*\* in [00; 19] |

### PoznanFencing

The general characteristics of the PoznanFencing sequence are summarized in Table 13 and source view positions in Table 30. The captured views form a 10x1 linear arc and are numbered v0-0 to v9-0 following left to right scan order.

Table 13: Characteristics of the PoznanFencing sequence

|  |  |
| --- | --- |
| Input contribution | m38247 |
| Length & frame rate | 250 frames (25 fps) |
| Number of source views | 10 |
| Source view resolution | 1920x1080 |
| View FoV & mapping | 63° × 48° |
| Lens | 4.5 mm |
| Camera spacing | 5 stereopairs (baseline: 22 cm) placed on arc (radius: 4 m),  angle between neighboring stereopairs: 15 degrees,  total angle of the system: 60 degrees |
| zNear | 3.5 |
| zFar | 7.0 |
| Download | <http://mpegfs.int-evry.fr/>mpegcontent/ws-mpegcontent/MPEG-I/Poznan/Poznan\_Fencing2  Texture maps: Poznan\_Fencing2\_tex.rar  Depth maps: Poznan\_Fencing2\_depth\_8bps\_cf420.rar  Metadata: Poznan\_Fencing2\_camparam.txt |

# Tools

## Introduction

The referenced tools are listed in the following table with related source location, documentation and version number, and detailed in the following subsections.

Table 14: List of used tools

|  |  |  |  |
| --- | --- | --- | --- |
| **Tool name** |  | **Location** | **Tag/branch** |
| RVS | [3] | <http://mpegx.int-evry.fr/software/MPEG/Explorations/3DoFplus/RVS> | v3.1 |
| WS-PSNR | [4] | <http://mpegx.int-evry.fr/software/MPEG/Explorations/3DoFplus/WS-PSNR> | v2.0 |
| HDRTools |  | <https://gitlab.com/standards/HDRTools/tree/v0.18> | v0.18 |
| 360Lib | [5] | <https://jvet.hhi.fraunhofer.de/svn/svn_360Lib/branches/360Lib-5.1-dev> | 5.1-dev |
| HM |  | <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.16> | 16.16 |
| VVS | [9] | http://mpegx.int-evry.fr/software/MPEG/Explorations/6DoF/VVS | v1.0 |
| HTM |  | <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-13.0> | 13.0 |

## RVS

The 3DoF+ Reference View Synthesizer [3], named RVS in the rest of the document is used to synthesize views at all positions, both at intermediate and source positions. RVS works with configuration files that specify a synthesis task from multiple input to multiple output views. Input and output views are described by a content metadata format that includes camera parameters and video formats. Configuration files may reference pose trace files that are CSV tables with per-frame modification of the output view pose. The default parameters of RVS 3.1 are aligned with the CTC and should not be specified.

## WS-PSNR

WS-PSNR (old name ERP\_WS-PSNR) is a tool in charge of computing WS-PSNR for objective metrics on images and is used to compare coding and synthesis results against uncompressed source views. For the omnidirectional ERP sequences, computation is done according to §4.2 of [5] and is applied to Y, U and V components, but only the Y component is used for evaluation.

The second version of WS-PSNR adds support for perspective images and configuration files. For perspective images the WS-PSNR method reduces to the regular PSNR method.

## HDRTools

HDRTools 0.18 is used to convert 8-bit YUV 4:2:0 texture frames into 10-bit YUV 4:2:0 texture videos, and 16-bit YUV 4:0:0 depth frames into 10-bit YUV 4:2:0 videos.

## 360Lib

360Lib-5.1 with HM-16.16 is used for encoding omnidirectional texture and depth streams.

## HM

HM 16.16 is used for encoding 3DoF+ content. For perspective sequences, it is used directly, while for omnidirectional sequences, it is used with 360Lib.

## VVS

The 6DoF Versatile View Synthesizer [9] – named VVS in the rest of the document is used to synthesize views at all positions, both at intermediate and source positions. VVS configuration files are provided in attachments (see Annex 1).

## HTM

MV-HEVC is used for encoding 6DoF content. The anchor is generated with MV-HEVC version 13.0 (macro HEVC\_EXT set to 1) [7], with the patch described in [8] applied to solve an implementation issue restricting the total number of views to be above 16.

# Anchor definition

The general structure of the anchors generation is represented in Figure 4. It consists of selecting a subset of the source views to be included in the anchor (possibly all), encoding those multiple views using a single layer HEVC bitstream for 3DoF+, or using MV-HEVC for 6DoF. The resulting bitstream is decoded and provides decoded views. From this set of decoded views, non-coded source views and intermediate views are synthesized using a reference tool, RVS and VVS for 3DoF+ and 6DoF respectively.

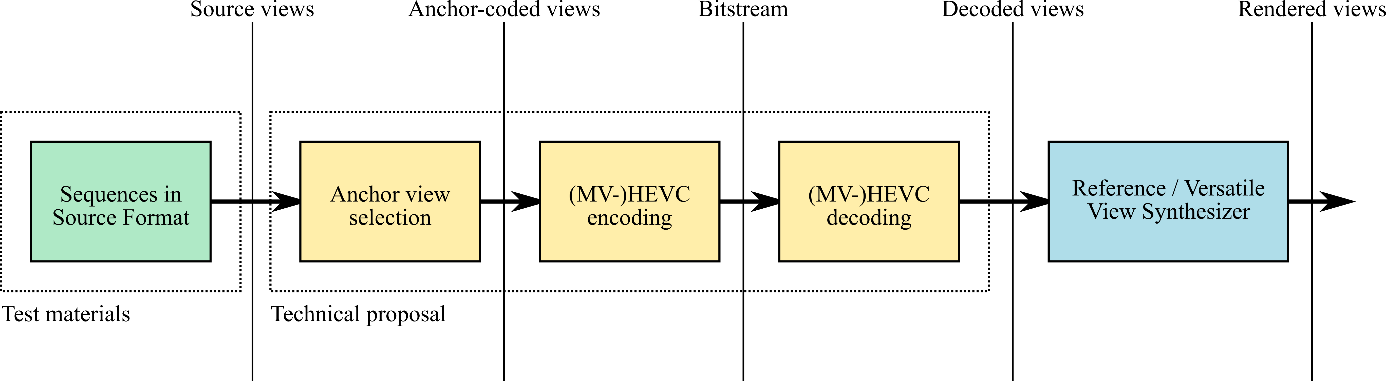


Figure 4: Definition of the anchor

## Coding of the anchor views

### Coding of 3DoF+ anchors

The 3DoF+ test conditions are aligned with the JVET common test conditions for the HM anchor using the random access 10-bit case. Compared with the 3DoF+ CfTM, the omnidirectional sequences are converted using a script [A8]. For all sequences (including perspective ones) there are HDRTools 0.18 configuration files to convert texture and depth to the format used for 3DoF+ [A9]:

* There is one texture and one depth stream per view,
* Camera ID’s have been renamed v0, v1, … in the order of the metadata file,
* Texture: From YUV 4:2:0 8-bit raw images to YUV 4:2:0 10-bit raw streams (BT.709),
* Depth: From YUV 4:0:0 16-bit raw images to YUV 4:2:0 10-bit raw streams (linear).

All views defined in Table 16 are coded using 360Lib-5.1-HM-16.16 software. The 360Lib add-on is only used for omnidirectional test material. The coding configurations to generate the anchors for all sequences (omnidirectional and perspective) are provided in attachment [A10]. The following command is used to encode the ClassroomVideo content in padded ERP format with HEVC RA Main10 based on related config files:

../bin/TAppEncoderStatic -c ../cfg/encoder\_randomaccess\_main10.cfg -c ../cfg/HM\_360Lib\_ClassroomVideo.cfg -i ./test\_seq/v1\_4096\_2048\_420\_10b.yuv -b v0\_4112\_2048\_420\_10b\_22.265 --SEIDecodedPictureHash=1 -o '' -q 22

RVS supports padded ERP through the Crop\_region parameter:

“Resolution”: [4112, 2048],  
 “Crop\_region”: [8, 0, 4096, 2048],

If the 3DoF+ scene is 180°×180°, then the content is coded in non-padded ERP format. For example, the following command is used to encode the TechnicolorMuseum content in ERP format with HEVC RA Main10:

../bin/TAppEncoderStatic -c ../cfg/encoder\_randomaccess\_main10.cfg -c ../cfg/HM\_360Lib\_TechnicolorMuseum -i ./test\_seq/v0\_2048\_2048\_420\_10b.yuv -b v0\_2048\_2048\_420\_10b\_22.265 --SEIDecodedPictureHash=1 -o '' -q 22

If the content is perspective, then HM 16.16 is used without 360Lib.

For each video sequence, two sets of QP points are considered, medium and low, corresponding respectively to QP1, QP2, QP3, QP4, and QP2, QP3, QP4, QP5, as defined in Table 15.

Table 15: QPs used for depth and texture

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **QP1** | **QP2** | **QP3** | **QP4** | **QP5** |
| Texture QP | 22 | 27 | 32 | 37 | 42 |
| Depth QP ClassroomVideo | 7 | 12 | 17 | 22 | 27 |
| Depth QP TechnicolorMuseum | 7 | 12 | 17 | 22 | 27 |
| Depth QP TechnicolorHijack | 7 | 12 | 17 | 22 | 27 |
| Depth QP TechnicolorPainter | 12 | 17 | 22 | 27 | 32 |
| Depth QP IntelKermit | 12 | 17 | 22 | 27 | 32 |

In alignment with JVET and JCT-VC CTC [6], the motion search range is set to 384, and the GOP size is set to 16. The last GOP is truncated to match the test sequence length.

The anchor encodes a subset of the source views, and for each sequence two classes are defined (Table 16). For each condition, WS-PSNR bitrate distortion curves are provided for each of the anchor-coded views, and for all anchor-coded views combined. The *X*1 classes provide a high-quality point, while the *X*2 classes *on average* comply with the maximum pixel rate.

WS-PSNR is computed with the WS-PSNR tool referenced in Table 14 both for coding and synthesis anchor.

Table 16: Anchor-coded views per class

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Test class | Sequence Name | Format | # of source views | # of anchor-coded views | Anchor-coded views |
| A1 | ClassroomVideo | ERP | 15 | 15 | All: v0…v14 |
| A2 | ClassroomVideo | ERP | 15 | 9 | v0, v7…v14 |
| B1 | TechnicolorMuseum | ERP | 24 | 24 | All: v0…v23 |
| B2 | TechnicolorMuseum | ERP | 24 | 8 | v0, v1, v4, v8, v11, v12, v13, v17 |
| C1 | TechnicolorHijack | semi-ERP | 10 | 10 | All: v0…v9 |
| C2 | TechnicolorHijack | semi-ERP | 10 | 5 | v1, v4, v5, v8, v9 |
| D1 | TechnicolorPainter | perspective | 16 | 16 | All: v0…v15 |
| D2 | TechnicolorPainter | perspective | 16 | 8 | v0, v3, v5, v6, v9, v10, v12, v15 |
| E1 | IntelKermit | perspective | 13 | 13 | All: v1…v13 |
| E2 | IntelKermit | perspective | 13 | 7 | v1, v3, v5, v7, v9, v11, v13 |

### Coding of Windowed-6DoF anchors

The anchor is generated with MV-HEVC. The coding configurations to generate the anchor are provided in attachment [A2] for textures and [A3] for depths. In particular, the following coding configuration is used:

* The inter-view prediction is set as described in Figure 5.
* The quantization parameters are given in Table 17.

Anchor results for the coded views are reported in the attached template file [A1].



Figure 5: Serpentine scan

Two sets of bitrates are considered, medium and low, corresponding respectively to QP1, QP2, QP3, QP4, and QP2, QP3, QP4, QP5, as defined in Table 17.

Table 17: QP used for depth and texture

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **QP1** | **QP2** | **QP3** | **QP4** | **QP5** |
| Depth QP | 34 | 39 | 42 | 45 | 48 |
| Texture QP | 25 | 30 | 35 | 40 | 45 |

The anchor encodes a subset of the source views or all source views, as defined in Table 18.

Table 18: Anchor-coded views for 6DoF sequences

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test class | Sequence Name | # of source views | # of anchor-coded views | Anchor-coded views |
| A1 | TechnicolorPainter | 16 (4x4) | 16 (4x4) | All |
| A2 | ULBUnicornA’ | 81 (9x9) | 25 (5x5) | v0-0, v2-0, …, v8-0  v0,2, v2-2, …, v8-2  … |
| A3 | ULBUnicornB’ | 45 (9x5) | 15 (5x3) | v0-0, v2-0, …, v8-0  v0,2, v2-2, …, v8-2  … |
| A4 | OrangeShaman | 25 (5x5) | 25 (5x5) | All |
| A5 | OrangeDancing | 69 (23 x 3) | 69 (23 x 3) | All |
| A6 | OrangeKitchen | 25 (5x5) | 25 (5x5) | All |
| A7 | IntelKermit | 13 (13x1) | 13 (13x1) | All: v1…v13 |
| A8 | ETRIChef | 25 (5x5) | 25 (5x5) | All |
| A9 | PoznanFencing | 10 (10x1) | 10 (10x1) | All |

## Synthesis of the intermediate views

### 3DoF+ view synthesis

Both for objective and subjective testing, a range of frames of each sequence are synthesized at source positions. All decoded views except for the target view are used as input views for RVS.

The 10-bit decoded textures and depth maps are used for the synthesis anchor. The format of each synthesized view is an omnidirectional image with equirectangular projection with the same angular resolution (pixels / degree) for ERP or semi-ERP test materials, and linear perspective projection for linear perspective input content. The synthesis result is 8-bit YUV 4:2:0 format for subjective evaluation and 10-bit YUV 4:2:0 for objective evaluation.

Inpainting of invalid pixels is used for both subjective and objective testing.

For 3DoF+ the reference view synthesizer is RVS. For objective evaluation only the first 32 and last 32 frames are used for view synthesis (Table 19). For subjective evaluation all frames are used. Synthesis configuration files for objective evaluation are available as attachment [A12]. Note that there are “head” and “tail” configurations to implement the non-continuous frame selection. The output of both runs should be concatenated to form 64-frame videos.

Table 19: Frame ranges for 3DoF+ view synthesis

|  |  |  |  |
| --- | --- | --- | --- |
| Test class | Sequence Name | Frames Objective Eval. | All Frames / Subjective Eval. |
| A1 | ClassroomVideo | 1-32, 89-120 | 1-120 |
| A2 | ClassroomVideo | 1-32, 89-120 | 1-120 |
| B1 | TechnicolorMuseum | 1-32, 269-300 | 1-300 |
| B2 | TechnicolorMuseum | 1-32, 269-300 | 1-300 |
| C1 | TechnicolorHijack | 1-32, 269-300 | 1-300 |
| C2 | TechnicolorHijack | 1-32, 269-300 | 1-300 |
| D1 | TechnicolorPainter | 1-32, 269-300 | 1-300 |
| D2 | TechnicolorPainter | 1-32, 269-300 | 1-300 |
| E1 | IntelKermit | 1-32, 269-300 | 1-300 |
| E2 | IntelKermit | 1-32, 269-300 | 1-300 |

For expert viewing, each sequence is also synthesized according to a set of pose traces. An HMD pose trace specifies for each frame the position and orientation of the viewport to synthesize. Each pose trace is stored as a comma-separated table with position (X, Y, Z) and orientation (Yaw, Pitch, Roll) columns and exactly one row per frame of the sequence. The format of each synthesized view is an image with perspective projection with at most 2048 × 2048 pixels resolution, at most 90-degree field of view and 8-bit YUV 4:2:0 color format. The purpose is to mimic natural viewing on a head-mounted display (HMD) while using offline tools and a 2D monitor.

Because of the large difference in visual comfort between a viewer that voluntarily initiates head motion versus a viewer watching the same viewport on a 2D monitor, pose traces will have a small amount of motion. For each sequence there are up to four pose traces – named *X*PT1 to *X*PT4 - which are meant to represent a diversity of natural head movement compliant with the overall dimension of the capture rig, as indicated in Table 21. Attachment [A14] contains all pose traces and example RVS configuration files.

Proposals are not required to code views corresponding to all anchor-coded views but are required to be able to generate viewport and omnidirectional or perspective video sequences for any intermediate view position in the designated range for each test sequence. The field of view for ERP (e.g. 180° or 360° degrees) will be the same as the source content.

It is meaningful to define the pose traces according to the conditions of capture, and typically to define the the related path within the volume of the camera rig. It is convenient to formulate this range as a volume in 3D space, as in Table 20.

Table 20: Intermediate view position ranges

|  |  |  |
| --- | --- | --- |
| Test classes | Name | Range volume description |
| A1, A2 | ClassRoomVideo | Spheroid centered at source view v0 eg (0, 0, 0) meter position, with equatorial radius 104 mm and polar distance 60 mm: |
| B1, B2 | TechnicolorMuseum | Sphere centered at position [0, 0, 1.65] meter with a 300 mm radius: |
| C1, C2 | TechnicolorHijack |
| D1, D2 | TechnicolorPainter | Spheroid centered at position [0, 0, 0] meter, covering a vertical square of side equal to 20cm and developed in the forward axis by 25cm max. |
| E1, E2 | IntelKermit | Rectangle centered at position [0.075, 0, 0] meter with a 15cm width, 44.1cm length, and no z component. |

### 6DoF view synthesis

In 6DoF, VVS is used for synthesis, with the configuration files as provided as attachments in [A4]. The camera parameters representing the positions of the views are provided as attachments in [A5]. Intermediate views are generated using four reference views: if available, the four closest left-right-top-bottom neighbors are used, otherwise depending on their availability, only three or two closest neighbors are used.

Objective evaluation only relies on comparisons made on source view positions. Subjective evaluation considers a navigation path made of coded views and intermediate views. The exact positions of each source view and intermediate views that follow the navigation path, and the reference views used to generate them are reported in attachment [A7]. Those anchor results for the synthesized views from decoded views are reported in the attached template file [A1], columns C-K.

# Rules for proponents

## 3DoF+

For 3DoF+, proponents shall use HEVC Main 10 profile for the compression of video data. The total pixel rate shall be reported, and should not be greater than in HEVC Level 6.2 (i.e. 8K@120Hz), and everything else is up to proponents (e.g. frame packing or coding views in separate bitstreams).

Proponents are allowed to apply transformations to the source format before encoding and after decoding. Examples are view selection, packing and projection (Figure 6).

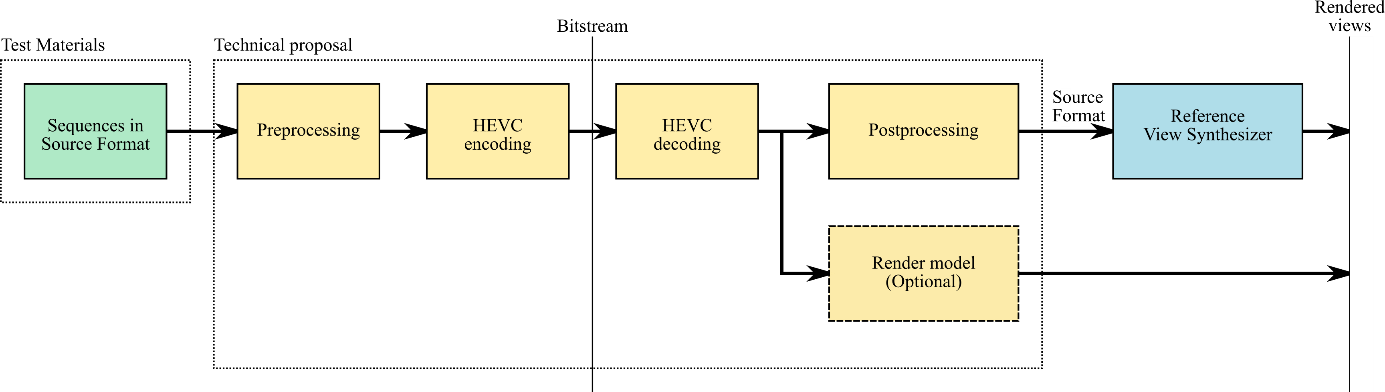


Figure 6: A technical proposal include pre- and post-processing, encoding and decoding and a render model

The default preferred solution consists for the proponent to regenerate all views at the source positions. Applying the RVS on those regenerated views will allow comparing with the output of RVS applied on the source views as illustrated in Figure 4.

Proponents may provide an RVS-modified or alternative render model that is capable of providing rectangular viewports and ERP frames. The render model shall accept the same pose parameters as provided to generate the reference views used for the results. Instead of the source format, this render model takes as input the content in the format as output by a proposed decoder.

When a proponent chooses to provide a render model as part of the proposed technical solution, then the proponent should evaluate the performance difference between RVS and the proposed render model by converting the output back to the source format. Valid reasons for a proponent to ignore this rule are that the conversion from the proposed format to the source format would take too much effort or that it would degrade quality. The proponent shall provide a motivation, because by using the RVS before and after encoding, evaluation concentrates on the format and excludes rendering aspects.

## 6DoF

For 6DoF, no restriction on the coding algorithm is to be met. Normative changes are expected. All source views might be encoded, or not, with or without the depth. The way the synthesis is performed is not restricted. The only constraint is the ability to generate the requested views.

# Evaluation of proposals

## Objective evaluation

### Definitions

With respect to a Test Class, the following definitions apply:

* “**Coded source view**” corresponds to a source view that is coded by the anchor,
* “**Synthesized view**” corresponds to a source view that is synthesized (interpolated) by the anchor using multiple decoded source views, always excluding the target view.

### 3DoF+ objective evaluation

For all test classes, WS-PSNR based BD-rate values will be provided for synthesized source views. For perspective views WS-PSNR reduces to regular PSNR. BD-rate values for coded source views will only be provided for the subset of views coded both by the anchor and proposal.

Coded views (texture or depth) are evaluated on all frames of each of the sequences (Table 19). The original view is the uncompressed source view.

Synthesized views are only evaluated on specified frames (Table 19). Those frames are synthesized and the result concatenated to form sequences. The same frame selection and concatenation is performed on the uncompressed source views to form the reference for WS-PSNR calculation.

The comparison of proponents with the anchors will be expressed in terms of BD rate computed on low- and medium bitrate rate-distortion RD curves, and separately for each Test Class:

* **Coded source view BD rate**, only when anchor and proposal code the same set of views, obtained from:
  + For the anchor RD curve:
    - The average over each view and all frames (Table 19) of the WS-PSNR between the coded view (coded texture) and the corresponding source view,
    - The total bitrate required to encode the views including texture and depths for all frames.
  + For the proponent’s RD curve:
    - The average over each view and all frames (Table 19) of the WS-PSNR between the proponent’s view (coded texture) corresponding to anchor coded view positions, and the source view,
    - The total bitrate of the proponent’s bitstream for all frames including texture, depth and metadata.
* **Synthesized view BD rate,** obtained from:
  + For the anchor RD curve:
    - The average over each source view and specified frames (Table 19) of the WS-PSNR between the intermediate view synthesized from decoded views and the original/non-compressed source views,
    - The total bitrate required to encode the views (including depths) for all frames.
  + For the proponent’s RD curve:
    - The average over each source view and specified frames (Table 19) of the WS-PSNR between the proponent’s synthesized intermediate view and the original/non-compressed source view,
    - The total bitrate of the proponent’s bitstream for all frames.
* Encoding, decoding and rendering runtimes ratios, compared to the anchor.

All the metrics related to a proponent solution are gathered in Table 21.

Table 21: Presentation of the results for new 3DoF+ proposals. Coded source view BD rates are only reported when anchor and proposal code the same set of views.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Configuration | Test Class | Coded source view BD rate | Synthesized view BD rate | Encoding time ratio |  | Synthesis time ratio |
| Random Access | X1 | [ x% ] | x% | x% |  | x% |
| Low bitrate | Y2 | [ x% ] | x% | x% |  | x% |
| **Average** | **[ x% ]** | **x%** | **x%** |  | **x%** |
| Random Access  Medium bitrate | X1 | [ x% ] | x% | x% |  | x% |
| Y2 | [ x% ] | x% | x% |  | x% |
| **Average** | **[ x% ]** | **x%** | **x%** |  | **x%** |

A reporting template [A11] is given as attached document.

### 6DoF objective evaluation

For 6DoF, new methods should be compared with the anchor coding results, by reporting the metrics as in Table 22, for medium-bitrates by considering the QP set {25, 30, 35, 40}, and for low-bitrates by considering the QP set {30, 35, 40, 45}.

* **Coded source view BD rate**, obtained from:
  + For the anchor RD curve:
    - the average over each view and all frames of the PSNR between the coded view and the corresponding source view,
    - the total bitrate required to encode the views (including depths) for all frames.
  + For the proponent’s RD curve:
    - the average over each view and all frames of the PSNR between the proponent’s view (coded or synthesized) corresponding to anchor coded view positions, and the source view,
    - the total bitrate of the proponent’s bitstream for all frames.
* **Synthesized view BD-Rate,** obtained from:
  + For the anchor RD curve:
    - the average over each source view position of the PSNR between the intermediate view synthesized from decoded views and the original/non-compressed source views,
    - the total bitrate required to encode the views (including depths) for all frames.
  + For the proponent’s RD curve:
    - the average over each source view position of the PSNR between the proponent’s synthesized intermediate view and the original/non-compressed source view,
    - the total bitrate of the proponent’s bitstream for all frames.
* Encoding, decoding and rendering runtimes ratios, compared to the anchor.

The synthesis algorithm used as an anchor considers four reference views. When more than four reference views are used in a proposal, it is recommended to additionally report the results with the four references considered for the anchor. Results can be captured in the attached template file [A1].

Table 22: Presentation of the results for 6DoF new proposals

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Configuration | Sequence | Coded source view BD rate | Synthesized view BD rate | Encoding time | Decoding time |  | Synthesis time |
| Random Access | Sequence 1 | x% | x% | x% | x% |  | x% |
| Low bitrate | Sequence 2 | x% | x% | x% | x% |  | x% |
| **Average** | **x%** | **x%** | **x%** | **x%** |  | **x%** |
| Random Access  Medium bitrate | Sequence 1 | x% | x% | x% | x% |  | x% |
| Sequence 2 | x% | x% | x% | x% |  | x% |
| **Average** | **x%** | **x%** | **x%** | **x%** |  | **x%** |

## Subjective evaluation

### 3DoF+ subjective evaluation

3DoF+ expert viewing will make use of a subset of synthesized source views as well as viewport sequences that are rendered out according to HMD pose files. All videos (ERP and perspective) will be viewed on a 2D monitor at 8-bit color depth.

### 6DoF subjective evaluation

Navigation paths are considered for subjective evaluation. The exact position of each frame from the navigation paths are provided in [A7]. The anchor for subjective evaluation is following the predefined path, and corresponds to synthesized views obtained from uncompressed source views with 8bits depth maps, as well as compressed views (anchor and proponent’s codec).

As an example for 2D arrays, the navigation path for the TechnicolorPainter sequence is shown in Figure 7. When the total length of the path is higher than the number of frames available. The sequence is consequently displayed back and forth according to the navigation path. The navigation path includes coded view positions and intermediate view position, obtained from diagonal, horizontal or vertical references, intermediate views located in-between two anchor coded views, or on arbitrary positions. In addition, the navigation path includes three stops, for which the same view is displayed for three seconds. The other 2D linear array sequences are adopting a similar kind of path, yet adapted to the number of source views.

As an example for 1D array, the navigation path for the IntelKermit sequence is show in Figure 8. It sweeps from one view to another, and includes two intermediate stops. The other 1D linear sequences are adopting a similar kind of path, yet adapted to the number of source views.



Figure 7: Navigation path for TechnicolorPainter sequence (and similar 2D array sequences)

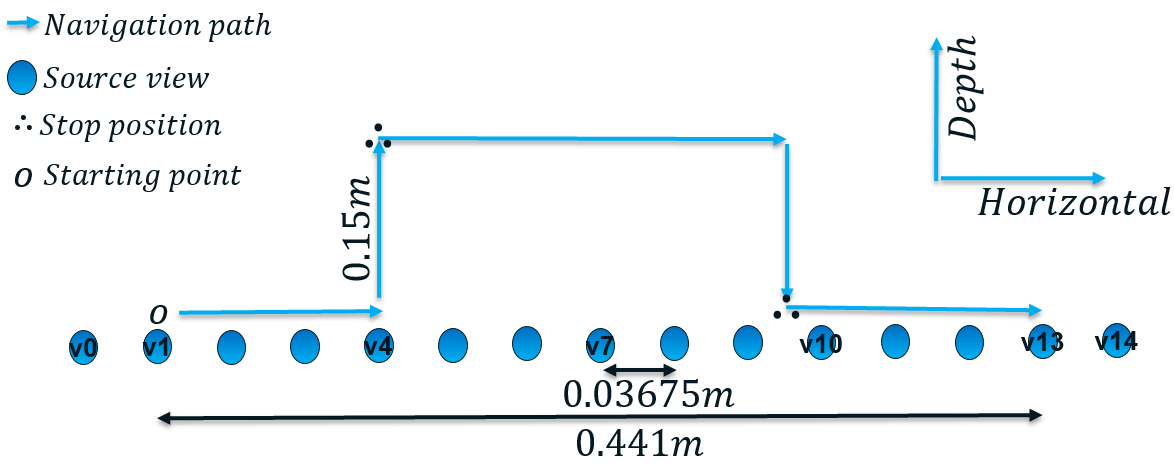


Figure 8: Navigation path for IntelKermit sequence (and similar 1D linear sequences)

# Annex 1: List of attachments

Table 23 provides an overview of files that are provided together with this document.

Table 23: List of attachments to this document

|  |  |  |
| --- | --- | --- |
| **ID** | **Filename** | **Short description** |
| [A1] | 6DoFTemplateV2.12.xlsm | Excel template v2.0 for reporting 6DoF anchor results |
| [A2] | OL\_mvhevc\_texture\_cfg.zip | MV-HEVC encoder configuration for 6DoF content (texture) |
| [A3] | OL\_mvhevc\_texture\_cfg.zip | MV-HEVC encoder configuration for 6DoF content (depth) |
| [A4] | OL\_vvs\_cfg.zip | VSRS configuration for 6DoF content |
| [A5] | OL\_vvs\_camparam\_json.zip | Camera parameters representing view positions for 6DoF content |
| [A6] | OL\_vsrs\_depth\_range.zip | ZNear and ZFar values for each view, for 6DoF content |
| [A7] | OL\_navigation\_paths.zip | Navigation paths for subjective evaluation of 6DoF content |
| [A8] | convert\_CfTM\_to\_CTC | Script to convert 3DoF+ sequences from the CfTM format to YUV420 10b streams. |
| [A9] | 3DoF+ HDRConvert\_cfg.zip | HDRTools configurations for all sequences (omnidirectional and perspective) for converting texture and depth from original source format to YUV 4:2:0 10-bit. |
| [A10] | 3DoF+ HM\_360Lib\_cfg.zip | HM and HM-360Lib configurations for all sequences (omnidirectional and perspective) including the Random Access Main10 configuration.t |
| [A11] | 3DoF+ template.xlsm | Excel template for reporting 3DoF+ anchor results |
| [A12] | 3DoF+ synth\_cfg.zip | View synthesis configurations for all QP’s |
| [A13] | 3DOF+ camparam.zip | Metadata files according to the updated format [3] with source and intermediate view positions for all sequences (omnidirectional and perspective). |
| [A14] | 3Dof+ posetraces\_cfg.zip | Pose traces APTP1…EPT3 and example configruations for RVS |

# Annex 2: Source view label conversion

Table 24: View naming of the TechnicolorMuseum sequence

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| camE0 | v0 | camE6 | v6 | camE12 | v12 | camE18 | v18 |
| camE1 | v1 | camE7 | v7 | camE13 | v13 | camE19 | v19 |
| camE2 | v2 | camE8 | v8 | camE14 | v14 | camE20 | v20 |
| camE3 | v3 | camE9 | v9 | camE15 | v15 | camE21 | v21 |
| camE4 | v4 | camE10 | v10 | camE16 | v16 | camE22 | v22 |
| camE5 | v5 | camE11 | v11 | camE17 | v17 | camE23 | v23 |

Table 25: View naming of the TechnicolorHijack sequence

|  |  |  |  |
| --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| camEA01 | v0 | camEA11 | v5 |
| camEA02 | v1 | camEA12 | v6 |
| camEA03 | v2 | camEA31 | v7 |
| camEA04 | v3 | camEA32 | v8 |
| camEA05 | v4 | camEA00 | v9 |

Table 26: View naming of the TechnicolorPainter sequence

|  |  |  |  |
| --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Texture file name** | **Depth file name** |
| v0-0 | v0 | TechnicolorPainter\_pr2\_view00 | DERS7\_TechnicolorPainter\_depth00\_pr2\_420\_TE1 |
| v1-0 | v1 | TechnicolorPainter\_pr2\_view01 | DERS7\_TechnicolorPainter\_depth01\_pr2\_420\_TE1 |
| v2-0 | v2 | TechnicolorPainter\_pr2\_view02 | DERS7\_TechnicolorPainter\_depth02\_pr2\_420\_TE1 |
| v3-0 | v3 | TechnicolorPainter\_pr2\_view03 | DERS7\_TechnicolorPainter\_depth03\_pr2\_420\_TE1 |
| v0-1 | v4 | TechnicolorPainter\_pr2\_view04 | DERS7\_TechnicolorPainter\_depth04\_pr2\_420\_TE1 |
| v1-1 | v5 | TechnicolorPainter\_pr2\_view05 | DERS7\_TechnicolorPainter\_depth05\_pr2\_420\_TE1 |
| v2-1 | v6 | TechnicolorPainter\_pr2\_view06 | DERS7\_TechnicolorPainter\_depth06\_pr2\_420\_TE1 |
| v3-1 | v7 | TechnicolorPainter\_pr2\_view07 | DERS7\_TechnicolorPainter\_depth07\_pr2\_420\_TE1 |
| v0-2 | v8 | TechnicolorPainter\_pr2\_view08 | DERS7\_TechnicolorPainter\_depth08\_pr2\_420\_TE1 |
| v1-2 | v9 | TechnicolorPainter\_pr2\_view09 | DERS7\_TechnicolorPainter\_depth09\_pr2\_420\_TE1 |
| v2-2 | v10 | TechnicolorPainter\_pr2\_view10 | DERS7\_TechnicolorPainter\_depth10\_pr2\_420\_TE1 |
| v3-2 | v11 | TechnicolorPainter\_pr2\_view11 | DERS7\_TechnicolorPainter\_depth11\_pr2\_420\_TE1 |
| v0-3 | v12 | TechnicolorPainter\_pr2\_view12 | DERS7\_TechnicolorPainter\_depth12\_pr2\_420\_TE1 |
| v1-3 | v13 | TechnicolorPainter\_pr2\_view13 | DERS7\_TechnicolorPainter\_depth13\_pr2\_420\_TE1 |
| v2-3 | v14 | TechnicolorPainter\_pr2\_view14 | DERS7\_TechnicolorPainter\_depth14\_pr2\_420\_TE1 |
| v3-3 | v15 | TechnicolorPainter\_pr2\_view15 | DERS7\_TechnicolorPainter\_depth15\_pr2\_420\_TE1 |

Table 27: View naming of the OrangeShaman & OrangeKitchen sequences

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| v0-0 | v00 | v0-1 | v05 | v0-2 | v10 | v0-3 | v15 | v0-4 | v20 |
| v1-0 | v01 | v1-1 | v06 | v1-2 | v11 | v1-3 | v16 | v1-4 | v21 |
| v2-0 | v02 | v2-1 | v07 | v2-2 | v12 | v2-3 | v17 | v2-4 | v22 |
| v3-0 | v03 | v3-1 | v08 | v3-2 | v13 | v3-3 | v18 | v3-4 | v23 |
| v4-0 | v04 | v4-1 | v09 | v4-2 | v14 | v4-3 | v19 | v4-4 | v24 |

Table 28: View naming of the IntelKermit sequence

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| cam00 | v14 | cam04 | v10 | cam08 | v06 | cam12 | v02 |
| cam01 | v13 | cam05 | v09 | cam09 | v05 | cam13 | v01 |
| cam02 | v12 | cam06 | v08 | cam10 | v04 | cam14 | v00 |
| cam03 | v11 | cam07 | v07 | cam11 | v03 |  |  |

Table 29: View naming of the ETRIChef sequence

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** |  | **Json view name** | **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| v0-0 | v00 | v0-1 |  | v05 | v0-2 | v10 | v0-3 | v15 |
| V1-0 | v01 | V1-1 |  | v06 | V1-2 | v11 | V1-3 | v16 |
| V2-0 | v02 | V2-1 |  | v07 | V2-2 | v12 | V2-3 | v17 |
| V3-0 | v03 | V3-1 |  | v08 | V3-2 | v13 | V3-3 | v18 |
| V4-0 | v04 | V4-1 |  | v09 | V4-2 | v14 | V4-3 | v19 |

Table 30: View naming of the PoznanFencing sequence

|  |  |  |  |
| --- | --- | --- | --- |
| **Original view name** | **Json view name** | **Original view name** | **Json view name** |
| v0-0 | v00 | v5-0 | v05 |
| v1-0 | v01 | v6-0 | v06 |
| v2-0 | v02 | v7-0 | v07 |
| v3-0 | v03 | v8-0 | v08 |
| v4-0 | v04 | v9-0 | v09 |

# References

[1] *Call for MPEG-I Visual Test Materials on 3DoF+ and 6DoF*, ISO/IEC JTC1/SC29/WG11 MPEG/N17720, July 2018, Ljubljana, Slovenia.

[2] *Call for MPEG-I Visual Test Materials on 6DoF*, ISO/IEC JTC1/SC29/WG11 MPEG/N17462, January 2018, Gwangju, Korea.

[3] *Reference View Synthesizer (RVS) manual*, ISO/IEC JTC1/SC29/WG11 MPEG/N18068, October 2018, Macau SAR, China.

[4] *ERP WS-PSNR manual*, ISO/IEC JTC1/SC29/WG11 MPEG/N18069, October 2018, Macau SAR, China.

[5] *Algorithm descriptions of projection format conversion and video quality metrics in 360Lib Version 5*, JVET-H1004, ISO/IEC JTC1/SC29/WG11 MPEG/N17197, October 2017, Macau SAR, China.

[6] *JVET common test conditions and software reference configurations*, JVET-B1010, ISO/IEC JTC1/SC29/WG11 MPEG/M38202, February 2016, San Diego, USA.

[7] *Limitations of Multiview extensions of HEVC and fixes for MPEG-I Phase 2*, ISO/IEC JTC1/SC29/WG11 MPEG/N17133, October 2017, Macau SAR, China.

[8] *FTV Software Summary*, ISO/IEC JTC1/SC29/WG11 MPEG/N16522, October 2016, Chengdu, China.

[9] *Versatile View Synthesizer (VVS) 1.0 manual*, ISO/IEC JTC1/SC29/WG11 MPEG/N18076, October 2018, Macau SAR, China.