**3GPP TSG-SA/WG4 Meeting # *S4***-***210730***

**Electronic meeting, May 18-29, 2021**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *CR-Form-v12.0* | | | | | | | | |
| **DRAFT CHANGE REQUEST** | | | | | | | | |
|  | | | | | | | | |
|  |  | **CR** | **-** | **rev** |  | **Current version:** | **1.1.2** |  |
|  | | | | | | | | |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* | | | | | | | | |
|  | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME |  | Radio Access Network |  | Core Network |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | |
| ***Title:*** | Editorial changes | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Source to WG:*** | Apple | | | | | | | | | |
| ***Source to TSG:*** |  | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Work item code:*** |  | | | | |  | ***Date:*** | | | 12th May 2021 |
|  |  | | | |  | |  | | |  |
| ***Category:*** |  |  | | | | | ***Release:*** | | |  |
|  | *Use one of the following categories:* ***F*** *(correction)* ***A*** *(mirror corresponding to a change in an earlier release)* ***B*** *(addition of feature),* ***C*** *(functional modification of feature)* ***D*** *(editorial modification)*  Detailed explanations of the above categories can be found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | | | | | | | | *Use one of the following releases: Rel-8 (Release 8) Rel-9 (Release 9) Rel-10 (Release 10) Rel-11 (Release 11) Rel-12 (Release 12)* *Rel-13 (Release 13) Rel-14 (Release 14) Rel-15 (Release 15) Rel-16 (Release 16)* | |
|  |  | | | | | | | | | |
| ***Reason for change:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Summary of change:*** | | Editorial changes to align with MPEG naming convention, and some text clarification.  Added reference to HDRTool  Update the text for some metrics definition (still some comment to be answer)  Addition of editor’s note concerning the test condition  Comments about the JM configuration  Update to some copyrights after review with Netflix | | | | | | | | |
|  | |  | | | | | | | | |
| ***Consequences if not approved:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Clauses affected:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
|  | | **Y** | **N** |  | | | |  | | |
| ***Other specs*** | |  |  | Other core specifications | | | | TS/TR ... CR ... | | |
| ***affected:*** | |  |  | Test specifications | | | | TS/TR ... CR ... | | |
| ***(show related CRs)*** | |  |  | O&M Specifications | | | | TS/TR ... CR ... | | |
|  | |  | | | | | | | | |
| ***Other comments:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***This CR's revision history:*** | |  | | | | | | | | |

Change 1

# 2 References

[…]

[64] Recommendation ITU-T P.910 (2008), Subjective video quality assessment methods for multimedia applications.

[65] HDRTools, version 0.21-dev, https://gitlab.com/standards/HDRTools/-/tree/0.21-dev

End of change 1

Change 2

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

MSE\_Y Mean Square Error of the luma component

PSNRy Peak-Signal to Noise Ratio of the luma component

PSNRu Peak-Signal to Noise Ratio of the chroma u component

PSNRv Peak-Signal to Noise Ratio of the chroma v component

PSNRyuv Average PSNR over all colour components

MS\_SSIM Multi-Scale Structural Similarity Index Metric

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AAC Advanced Audio Coding

ABR Adaptive BitRate

AOM Alliance for Open Media

AOV Arena Of Valor

ATSC Advanced Television Systems Committee

AVC Advanced Video Coding

AVCHD AVC High Definition

AVI Audio Video Interleave

BMFF Based Media File Format

CABAC Context-Adaptive Binary Arithmetic Coding

CAE Content Aware Encoding

CBP Constrained Baseline Profile

CBR Constant BitRate

CGI Computed Generated Imaginary

CHP Constrained High Profile

CMAF Common Media Application Format

CRA Clean Random Access

CSV Comma-Separated Values

CTC Common Test Conditions

CTU Coding Tree Unit

DASH Dynamic Adaptive Streaming over HTTP

DVB Digital Video Broadcasting

EBU European Broadcast Union

EPZS Enhanced Predictive Zonal Search

ERP Equi-Rectangular Projection

EVC Essential Video Coding

FPS Frames Per Second

GOP Group-Of-Pictures

HDMI High-Definition Multimedia Interface

HDR High Dynamic Range

HDTV High Definition TeleVision

HEVC High-Efficiency Video Coding

HFR High Frame Rate

HLG Hybrid Log-Gamma

HLS HTTP Live Streaming

HMD Head-Mounted Display

HRD Hypothetical Reference Decoder

HTML HyperText Markup Language

IDR Instantaneous Decoder Refresh

JCT-VC Joint Collaborative Team on Video Coding

JSON JavaScript Object Notation

JVET Joint Video Experts Team

MIME Multipurpose Internet Mail Extensions

MKV MatrosKa Video

MMO Massive Multiplayer Online

MMORPG MMO Role-Playing Game

MMS Multimedia Messaging Service

MOBA Multiplayer Online Battle Arena

MOS Mean Opinion Score

MPD Media Presentation Description

MPEG Moving Pictures Expert Group

MS-SSIM Multi-Scale Structural Similarity Index

MSE Mean Square Error

MTSI Multimedia Telephony Service over IMS

OBS Open Broadcaster Software

PSNR Peak Signal to Noise Ratio

RAP Random Acces Point

RCS Rich Communication Services

RDPCM Residual Differential Pulse Code Modulation

RGB Red Green Blue

RPG Role-Playing Game

RTMP Real-Time Messaging Protocol

RTP Realtime Transport Protocol

RTS RealTime Strategy

SCC Screen Content Coding

SCM Screen Content coding Model

SDK Software Development Kit

SDR Standard Definition Range

SEI Supplemental Enhancement Information

SI Spatial perceptual Information

SVOD Subscription Video On Demand

TI Temporal perceptual Information

TIFF Tagged Image File Format

UDP User Datagram Protocol

UHD Ultra High Definition

URI Uniform Resource Identifier

URL Uniform Resource Locator

UVG Ultra Video Group

VBR Variable BitRate

VBS Visual Basic Script

VCL Video Coding Layer

VMAF Video Multimethod Assessment Fusion

VSEI Versatile Supplemental Enhancement Information

VTM VVC Test Model

VVC Versatile Video Coding

WCG Wide Colour Gamut

End of change 2

Change 3

## 4.1 Introduction

This clause summarizes the video coding capabilities in relevant existing 5G services as of Releease-16. Video codecs introduced in this clause are referred to as 3GPP video codecs.

As of today, two codecs are prominently referenced and available, namely H.264 (AVC) [7] and H.265 (HEVC) [8].

Both codecs are defined as part of the TV Video Profiles in 3GPP TS 26.116 [3] and are also the foundation of the VR Video Profiles in 3GPP TS 26.118 [4]. The highest defined profile/level combinations are:

- H.264 (AVC) Progressive High Profile Level 5.1 [7] with the following additional restrictions and requirements:

o the maximum VCL Bit Rate is constrained to be 120Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively.

o the bitstream does not contain more than 10 slices per picture

- H.265 (HEVC) Main 10 Profile Main Tier Level 5.1 [8] without any restrictions

Both codecs are also defined for other 3GPP-based services. More details on the codec capabilities and the necessary interoperability requirements for different services are collected in the remainder of this clause.

## 4.2 TV Video Profiles

The TV Video Profiles in TS 26.116 [3] address coded representations of TV distribution signals up to UHD-1 phase 2. Table 4.2-1 provides an overview of the TV relevant formats considered in the context of 3GPP TV Video Profiles.

In the context of TV Video Profiles, the following aspect are defined:

- **Bitstream**: A media bitstream that conforms to a video encoding format and certain Operation Point.

- **Operation Point**: A collection of discrete combinations of different content formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and the encoding format.

- **Receiver**: A receiver that can decode and render any bitstream that is conforming to a certain Operation Point.

Table 4.2-1: TV over 3GPP services Video Profile Operation Points (TS 26.116 [3])

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Operation Point name | Resolution format | Picture aspect ratio | Scan | Max. frame rate | Chroma format | Chroma sub-sampling | Bit depth | Colour space format | Transfer  Characteristics |
| H.264/AVC 720p HD | 1280 × 720 | 16:9 | Progressive | 30 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.265/HEVC 720p HD | 1280 × 720 | 16:9 | Progressive | 30 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.264/AVC Full HD | 1920 × 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.265/HEVC Full HD | 1920 × 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 8; 10 | BT.709; BT.2020 | BT.709; BT.2020 |
| H.265/HEVC UHD | 3840 × 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2020 |
| H.265/HEVC Full HD HDR | 1920 x 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 PQ |
| H.265/HEVC UHD HDR | 3840 x 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 PQ |
| H.265/HEVC Full HD HDR HLG | 1920 x 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 HLG |
| H.265/HEVC UHD HDR HLG | 3840 x 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 HLG |

Each one of the Operation Points is associated with a video coding specification and a particular Profile, Level, and Tier (for HEVC). The combination of a specific profile, level, and tier indicate the maximum decoding capabilities, such as chroma format, resolution, frame rate, and bit depth, that can be supported by a decoder conforming to such combination of parameters. An operation point is also associated with additional bitstream constraints defined in TS 26.116 [3], clause 4. Table 4.2-2 summarizes the video coding specification profile, tier, and level associated with each operation point.

Table 4.2-2: Operation point video codec Profile/Tier/Level

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Operation Point name | Video Codec | Profile | Tier | Level |
| H.264/AVC 720p HD | AVC | High | - | 3.1 |
| H.265/HEVC 720p HD | HEVC | Main | Main | 3.1 |
| H.264/AVC Full HD | AVC | High | - | 4.2 |
| H.265/HEVC Full HD | HEVC | Main 10 | Main | 4.1 |
| H.265/HEVC UHD | HEVC | Main 10 | Main | 5.1 |
| H.265/HEVC Full HD HDR | HEVC | Main 10 | Main | 4.1 |
| H.265/HEVC UHD HDR | HEVC | Main 10 | Main | 5.1 |
| H.265/HEVC Full HD HDR HLG | HEVC | Main 10 | Main | 4.1 |
| H.265/HEVC UHD HDR HLG | HEVC | Main 10 | Main | 5.1 |

For TV Video profiles, interoperability with ISO BMFF based systems and the DASH Streaming is of most relevance. Hence, for a codec to be used in the context of TV Video Profiles, the following is defined in terms of interoperability:

1. The receiver requirements on elementary stream level
2. The encapsulation of an elementary stream into an ISO Base Media File Format track
3. The provisioning of the media as part of the DASH Adaptation Set to support seamless switching
4. All MPD-level signalling for the codec to support capability discovery

For details, refer to TS 26.116 [3], clause 4 and clause 5.

## 4.3 VR Video Profiles

The VR profiles for streaming applications defined in TS 26.118 [4] address the coded representation of 360º VR distribution signals. Table 4.3-1 provides an overview of the 360 º VR relevant formats considered in the context of 3GPP VR profiles. The VR profiles follow the same logic as the TV Video profiles; they represent a list of interoperability points that are amended by the 3GPP services such as 5GMS.

Table 4.3-1: High-level Summary of Operation Points

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Operation Point name | Decoder | Bit depth | Typical  Original Spatial Resolution | Frame Rate | Colour space format | Transfer  Characteristics | Projection | RWP | Stereo |
| Basic H.264/AVC | H.264/AVC HP@L5.1 | 8 | Up to 4k | Up to 60 Hz | BT.709 | BT.709 | ERP w/o padding | No | No |
| Main H.265/HEVC | H.265/HEVC MP10@L5.1 | 8, 10 | Up to 6k in mono and 3k in stereo | Up to 60 Hz | BT.709  BT.2020 | BT.709 | ERP w/o padding | Yes | Yes |
| Flexible H.265/HEVC | H.265/HEVC MP10@L5.1 | 8, 10 | Up to 8k in mono and 3k in stereo | Up to 120 Hz | BT.709  BT.2020 | BT.709,  BT.2100 PQ | ERP w/o padding CMP | Yes | Yes |
| Main 8K H.265/HEVC | H.265/HEVC MP10@L6.1 | 10 | Up to 8k in mono and 6k in stereo | Up to 60 Hz for 8K and 120 Hz for 4k | BT.709  BT.2020 | BT.709, BT.2100 PQ, | ERP w/o padding | Yes, but restricted to coverage | Yes |

## 4.4 5G Media Streaming

### 4.4.1 Introduction

5G Media Streaming (5GMS) services are associated with a series of 5GMS profiles, each of which contains a set of capability requirements associated with a service scenario. Default profiles (for downlink and uplink streaming) are also defined in case no other profile is claimed to be supported. The detailed requirements for 5G Media Streaming profiles and their associated codec requirements are defined in 3GPP TS 26.511 [13]. The following clauses focus on video codec requirements and recommendations in 5GMS specification as well as the encapsulation format.

NOTE: In some profiles the HD-HDR capability is used as the ability for a UE to present video signals with all the following features: at least Full-HD resolution, bit depth of at least 10, at least 60 frames per second, Wide Colour Gamut and High Dynamic Range.

### 4.4.2 5GMS Downlink Streaming default profile

#### 4.4.2.1 H.264 (AVC)

TS 26.511 [13] requires the support of H.264 (AVC) Progressive High Profile Level 3.1 [7] decoding, with the maximum VCL Bit Rate constrained to be 14 Mbps.

If the 5GMS client supports the reception of video and HD-HDR capabilities, then TS 26.511 [13] requires the support of H.264 (AVC) Progressive High Profile Level 4.0 [7] decoding.

#### 4.4.2.2 H.265 (HEVC)

TS 26.511 [13] recommends the support of H.265 (HEVC) Main Profile, Main Tier, Level 3.1 [8] decoding with no interlace support.

If the 5GMS client supports the reception of video and HD-HDR capabilities, then TS 26.511 [13] requires the support of H.265 (HEVC) Main 10 Profile, Main Tier, Level 4.1 [8] decoding with no interlace support.

#### 4.4.2.3 Encapsulation format

TS 26.511 [13] requires the support of Common Media Application Format (CMAF) encapsulation as defined in ISO/IEC 23000-19 [30].

End of change 3

Change 4

### 5.4.3 H.265/HEVC

For H.265/HEVC generated anchor bitstreams, the following reference software has been used:

- *HM16.22* H.265/MPEG-H HEVC Reference Software: The reference software for H.265/MPEG-H HEVC is called HM (HEVC Test Model). The HM software is maintained and can be downloaded from in the repository https://vcgit.hhi.fraunhofer.de/jct-vc/HM and development versions are available. For HM-based H.265/MPEG-4 HEVC anchors, version 16.22 is used except specified otherwise. The reference software supports the following 3GPP profiles

- H.265 (HEVC) Main Profile

- H.265 (HEVC) Main 10 Profile

- *SCM8.8*: Extensions for screen content coding are provided with the SCM (Screen Content Coding Model) software available here <https://vcgit.hhi.fraunhofer.de/jct-vc/HM/-/tags/HM-16.21+SCM-8.8>. The reference software supports the following 3GPP profile

- H.265 (HEVC) [8] Screen-Extended Main Profile

## 5.5 Metrics

### 5.5.1 General

Each anchor bitstream gets assigned multiple performance metrics, in particular:

- the bitrate of the sequence, i.e. the sum of the size of all compressed pictures divided by duration of the sequence.

- If Standard Dynamic Range (SDR) is used, then the metrics in clause 5.5.2 apply.

- If High Dynamic Range (HDR) is used, then the metrics in clause 5.5.3 apply.

An overview of the metrics is provided below. These metrics are implemented in software scripts defined in Annex F. This software is used to compute and report all the metrics.

Editor’s Note: At this stage the metrics are defined as the output of the reference software that generates the anchors. In an updated version of the TR, an independent metric computation tool will be provided that allows to generate the identical metrics based on the references sequence and the decoded test sequence.

Subjective evaluation of anchors streams is not excluded per se, but this specification does not define any recommended metric or method until now.

### 5.5.2 SDR Metrics

For standard dynamic range (SDR) sequences, the following metrics are used:

- Peak-Signal to Noise Ratio (PSNRy) of luma component, as specified in [44]:

For an individual frame, the mean square error is calculated between the luma channel of the decoded output image and the luma channel of the original image according to

where and are the luma sample values at position of the decoded and original image. and are the width and height of the luma component, respectively. The luma PSNR value for the frame is then calculated as:

where bitDepth is the bit-depth of the input video. For reporting, bitDepth shall be equal to 10 for all sequences. All 8-bit content shall be converted to 10-bit input in the encoder by shifting 2 bits to the left and using the 10-bit PSNR calculation to report testing results. The max PSNR for a frame is fixed to 999.99 dB.

Then the PSNR for each bitstream is computed as the sum of all individual frame PSNR values divided by the number of frames in the sequence.

* Similarly, the Peak-Signal to Noise Ratio for chroma u and v, PSNRu and PSNRv respectively, are computed using the chroma sample values.

- Average PSNR over all colour components (PSNRyuv):

PSNRyuv = (6\*PSNRy + PSNRu + PSNRv)/ 8

- Structural similarity metric MS-SSIM, as specified in [54] [55] and [56]:

The multi-scale SSIM method is illustrated in Figure x. Taking the reference x and distorted y image signals as the input, the system iteratively applies a low-pass filter and downsamples the filtered image by a factor of 2. We index the original image as Scale 1, and the highest scale as Scale M, which is obtained after M-1 iterations. At the j-th scale, the contrast and structure components are calculated and denoted as Cj (x, y) and Sj (x, y), respectively. The luma component (inappropriately named as the luminance component in the references) is computed only at Scale M and is denoted as lM (x, y). The overall metric for each frame is obtained by combining the measurement at different scales.

Diagram

Description automatically generated

Figure x: Multi-scale structural similarity measurement system (L:low-pass filter, 2↓: downsampling by 2)

The MS-SSIM between the original image, *I,* and the reconstructed image component, *I’*, is calculated as:

where are pixels of original and reconstructed frames, is the number of pixels per frame, , , , , , are the constants:

β1 = γ1= 0.0448, β2 = γ2 = 0.2856, β3 = γ3 = 0.3001, β4 = γ4 = 0.2363, and α5 = β5 = γ5 = 0.1333.

C1= (K1\*maxValue)2

C2= (K2\*maxValue)2

C3=C2/2,

K1 = 0.01, K2 = 0.03, maxValue =(1<<bitDepth)-1, and .

Average values , at pixels and are computed as weighted sum of neighbors with Gaussian weights. , represent the variance at pixels and . is covariance of these two pixels.

The MS\_SSIM value for each bitstream is computed as the sum of all individual frame MS\_SSIM values divided by the number of frames in the sequence.

See section 5.7, on the use of [59] for computation of MS-SSIM.

It is quite common to convert the MS-SSIM numbers to a dB representation since that representation can be more easily interpreted and is somewhat similar to the PSNR representation. Such computation also has an impact in the BD-rate numbers since the resulting points end up having more similar properties to the PSNR ones for the BD-rate computation. Both HDRTools and VMAF support this conversion, using:

  log\_MS\_SSIM = (-10.0 \* log10(1 – MS\_SSIM));

The log\_MS\_SSIM value for each bitstream is computed as the sum of all individual frame log\_MS\_SSIM values divided by the number of frames in the sequence.

Note that if the log\_MS\_SSIM is used, the MS\_SSIM values for each bitstream is no more available.

- Video Multimethod Assessment Fusion (VMAF), as specified in [57]. See section 5.7, on the use of [59] to compute VMAF.

### 5.5.3 HDR Metrics

For high dynamic range (HDR) sequences, the following metrics are used:

- PSNRL100, as specified in [44] and [58],

- Weighted PSNR over colour components (wPSNR), as specified in [44] and [58],

- DE100, as specified in [44] and [58].

Editor’s Note: For now, we use the output of the HM test model encoder – addition to scripts will be provided based on a fix of the HDRTools.

### 5.5.4 Anchor Tuple Metrics Reporting

For each anchor tuple, the Metrics are reported in a single csv file as defined in IETF RFC 4180 [60].

Additional metrics may be reported, for example encoding times, etc.

Table 5.5.4-1 provides the result format for SDR/HLG.

Table 5.5.4-1 Result Format for Full HD Scenario for SDR/HLG

|  |  |  |
| --- | --- | --- |
| Name | Type | Semantics |
| parameter | BIGINT | the associated variation parameter as defined for the anchor, for example the QP |
| bitrate | DOUBLEPRECISION | The size of the file divided by the duration of the reference sequence in bit/s as defined in clause 5.5.1 with 2 decimal digits accuracy. |
| y\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for Y planes in dB as defined in clause 5.5.2 with 2 decimal digits accuracy. |
| u\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for U planes in dB as defined in clause 5.5.2 with 2 decimal digits accuracy. |
| v\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for V planes in dB as defined in clause 5.5.2 with 2 decimal digits accuracy. |
| ms\_ssim | DOUBLEPRECISION | structural similarity between frames as defined in clause 5.5.2 with 4 decimal digits accuracy. |
| vmaf | DOUBLEPRECISION | Video Multimethod Assessment Fusion (VMAF) as defined in clause 5.5.2 with 2 decimal digits accuracy. |
| bitrate\_log | DOUBLEPRECISION | The bitrate as documented by the encoder log. If not known, it is set to 0. |
| encode\_time | DOUBLEPRECISION | Total time spent to encode the sequence with reference encoder in seconds. If not known, it is set to 0. |
| decode\_time | DOUBLEPRECISION | Total time spent to decode the sequence with reference decoder in seconds. If not known, it is set to 0. |

Table 5.5.4-2 provides the result format for HDR.

Table 5.5.4-2 Result Format for HDR

|  |  |  |
| --- | --- | --- |
| Name | Type | Semantics |
| Parameter | BIGINT | the associated variation parameter as defined for the anchor, for example the QP |
| Bitrate | DOUBLEPRECISION | The size of the file divided by the duration of the reference sequence in bit/s as defined in clause 5.5.1. |
| y\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for Y planes in dB as defined in clause 5.5.2. |
| u\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for U planes in dB as defined in clause 5.5.2. |
| v\_psnr | DOUBLEPRECISION | Peak signal to noise ratio for V planes in dB as defined in clause 5.5.2. |
| Wpsnr | DOUBLEPRECISION | Weighted peak signal to noise ratio for Y, U and V planes in dB as defined in clause 5.5.3. |
| psnrl100 | DOUBLEPRECISION | PSNRL100 as defined in clause 5.5.3 |
| de100 | DOUBLEPRECISION | DE100 as defined in clause 5.5.3. |
| bitrate\_log | DOUBLEPRECISION | The bitrate as documented by the encoder log. If not known, it is set to 0. |
| encode\_time | DOUBLEPRECISION | Total time spent to encode the sequence with reference encoder in seconds. If not known, it is set to 0. |

## 5.6 Tests

Tests may be executed to compare codecs not yet in 3GPP specifications against anchors defined in this specification. Tests, equivalently to anchors, are collected in tuples to address different quality and bitrates that can then be used for evaluation over a larger set of operation points.

A test is developed against an anchor and is a combination of:

- The corresponding anchor, which includes

- Scenario

- Reference Sequence

- Test encoder

- Test encoder configuration that provides an equivalent setting to the anchor configuration

Editor’s Note: This needs to be clarified. In particular what do we mean by equivalent setting? Are we trying to match the coding parameters of the anchor or are we trying to match a particular behavior that the anchor is exhibiting? If so, how is that behavior defined? This is very vague and not at all well defined.

- Variable test encoder configuration to create multiple quality/bitrate variants (using for example QP variations or other bitrate/quality evaluation tools).

- Test tuples creating multiple variants, each including

- Test bitstream

- Test Metrics

- Additional recommended test information includes

- MD5 check sum of the complete reconstructed yuv file (reconstructed test sequence)

- Output picture log for reference encoder

- Output picture log for reference decoder

- Tests are an integral part of the Technical Report

The generation of test tuples is shown in Figure 5.6-1.

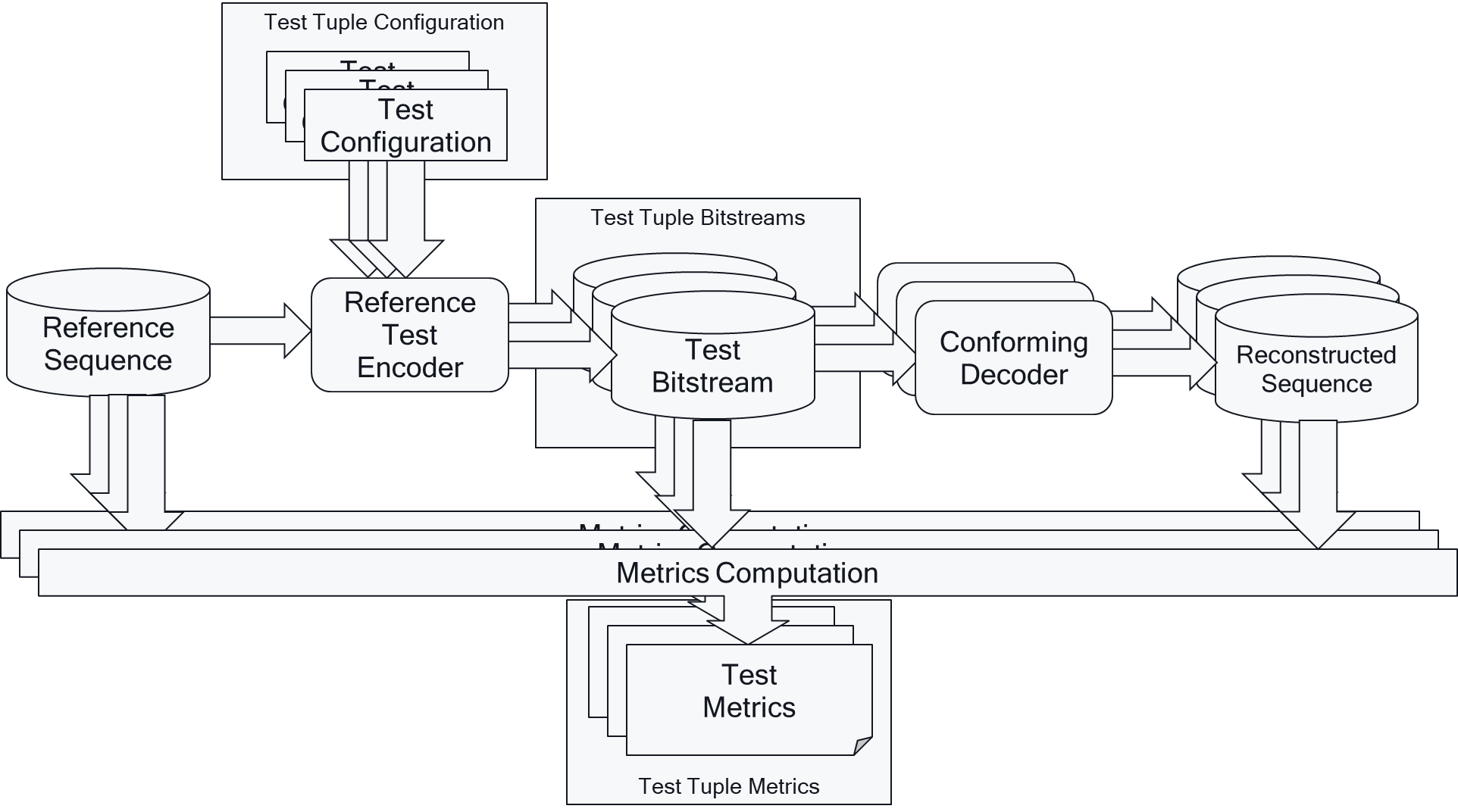


Figure 5.6-1: Test Tuple Generation Framework and Test Tuple Metrics Generation

## 5.7 Characterization

Characterization is the comparison of a codec under test with an anchor based on the framework introduced in this clause. Characterization in this report is based on Bjöntegard-Delta (BD)-rate information according to [44].

A full characterization of a codec for a scenario against a 3GPP codec shall provide at least the following metrics

- The BD-rate gain for each defined anchor tuple and each required metric

- The average BD-rate gain across all anchors of the scenario for each required metric.

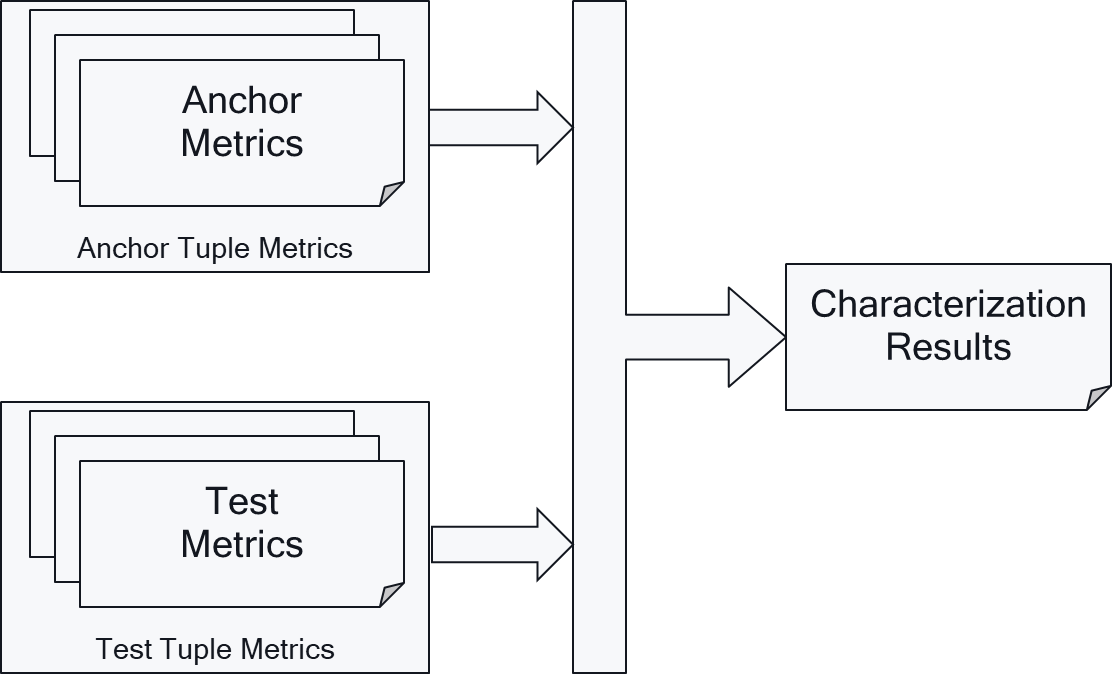


Figure 5.7-1: Characterization Framework

BD-Rate is computed according to the CTC method used in JVET and specified in [44] from the tools publicly available: Reference codec software, Excel file available in [57] for SDR and in [58] for HDR.

The Excel files include the VBS script bdrate( ) to compute the BD-Rate performance between a test codec and a reference from four or five rate-distortion points.

These excel files have been extended in the Random-Access and low delay tabs to contain new columns for the new metrics: VMAF and MS-SSIM, in the SDR case only. The “SA4 extended excel files” for SDR and HDR are attached as S4-template-HDR.xlsx and S4-template-SDR.xlsx.

If a metric does not have a monotonic behavior for particular sequence, an error should be reported, and the metric should not be computed for that sequence.

Editor’s Note: At this stage the BD-Rate is defined as result of the above excel sheets. In an updated version of the TR, an independent metric computation tool may be provided that allows to generate the BD-Rate values based on the metrics results. The reference for an extended version with 5 points is still tbd.

For the computation of VMAF and MS-SSIM, the C++ executable “vmafossexec” [59], open source provided by Netflix could be used (Licence BSD + Patent) (Note: a tag need to be defined for libvmaf and vmafossexec). MS-SSIM is computed in Vmafossexec with the default 11 Gaussian Window and default K1=0.01 and K2=0.03.

Here is the command line:

vmafossexec $VMAF\_FMT $WIDTH $HEIGHT ref.yuv test.yuv $VMAFMODEL --thread 1 --psnr --ssim --ms-ssim --log metrics.vmaf

$VMAF\_FMT: describe yuv subsampling (yuv420p10le or yuv420p8In10leOut)

$VMAFMODEL: vmaf\_4k\_v0.6.1.pkl (4K and more) or vmaf\_v0.6.1.pkl (HD and lower res)

thread: 0 to use all threads available

Note: the VMAF executable allows to extract the psnr which could also be used to check if it matches reference software output.

Another optional method is provided, as described in Annex F, to compute the metrics automatically. BD-Rate computation is supported by a script that uses anchor tuple and test tuple metrics to provide the characterization results as shown in Figure 5.7-1.

For details on BD-Rate computation, refer to [44].

End of change 3

Change 4

### 6.2.4 Encoding and Decoding Constraints

Table 6.2.4-1 provides an overview of encoding and decoding constraints for H.264/AVC Full HD and H.265/HEVC Full HD Profiles. This will support the definition of detailed test conditions.

Table 6.2.4-1 Encoding and Decoding Configurations

|  |  |  |
| --- | --- | --- |
| Encoding and Decoding Constraints | H.264/AVC Full HD | H.265/HEVC Full HD |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | H.264/AVC Progressive High Profile Level 4.2 [7] | HEVC/H.265 Main 10 Profile  Main Tier Level 4.1 [8] |
| Random access frequency | 1 second, 3.84 seconds | 1 second, 3.84 seconds |
| Error resiliency requirements | None | None |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | Fixed QP  CBR 8–12 Mbit/s  VBR capped at 12 Mbit/s  others | Fixed QP  CBR 5–8 Mbit/s  VBR capped at 12 Mbit/s  others |
| ABR encoding requirements (switching frequency, etc.) | 1 second  ABR through multiple QPs | 1 second  ABR through multiple QPs |
| Latency requirements and specific encoding settings | No latency requirements beyond RAP so picture reordering is allowed | No latency requirements beyond RAP so picture reordering is allowed |
| Encoding complexity context | real-time encoding, cloud-based encoding, offline encoding, etc. | real-time encoding, cloud-based encoding, offline encoding, etc. |
| Required decoding capabilities | H.264/AVC Progressive High Profile Level 4.2 [7] | HEVC/H.265 Main 10 Profile  Main Tier Level 4.1 [8] |

### 

End of change 4



### 

Change 5

### 6.3.4 Encoding and Decoding Constraints

Table 6.3.4-1 provides an overview of encoding and decoding constraints for 4K-TV category using legacy codec HEVC. This will support the definition of detailed test conditions. It is noted that no relevant profiles exist in TS26.116 and TS26.511 for HFR 4K-TV content.

Table 6.3.4-1 Encoding and Decoding Configurations for 4K-TV with legacy HEVC codec

|  |  |  |
| --- | --- | --- |
| Encoding and Decoding Constraints | H.265/HEVC 4K-TV | H.265/HEVC 4K-TV HFR |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | H.265/HEVC Main 10 Profile  Level 5.1 [8] | No relevant 3GPP profiles, should be aligned with H.265/HEVC Main 10 Profile Level 5.2 [8] |
| RAP period | 3.84sec, 1sec | 3.84sec, 1sec |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | QP = [tbd]  B = {10,20,30,40} Mbps [49]  CBR and capped-VBR | QP = [tbd]  B = {10,20,30,40} Mbps [49]  CBR and capped-VBR |
| Latency requirements and specific encoding settings | No latency requirements beyond RAP so picture reordering allowed | No latency requirements beyond RAP so picture reordering allowed |
| Encoding complexity context | real-time encoding, cloud-based encoding, offline encoding, etc. | real-time encoding, cloud-based encoding, offline encoding, etc. |
| Required decoding capabilities | H.265/HEVC Main 10 Profile  Level 5.1 [8] | H.265/HEVC Main 10 Profile  Level 5.2 [8] |

End of change 5

Change 6

### 6.4.4 Encoding and Decoding Constraints

Table 6.4.4-1 provides an overview of the different codec tools per profile that may be suitable for coding screen content sequences.

Table 6.4.4-1 Screen Content Tools per Profile

|  |  |  |
| --- | --- | --- |
| Screen content tools | AVC | HEVC |
| main profile | Not applicable | Transform skip |
| range extension profile | not applicable | Residual Differential Pulse Code Modulation (RDPCM) (implicit intra/explicit inter), |
| screen content profile | Not applicable | Intra Block Copy (full frame or less), Palette, Adaptive Colour Transform |

Table 6.4.4-2 provides an overview of encoding and decoding constraints for Screen Content scenario

- general constraints

- using H.264/AVC codecs operating points,

- using H.265/HEVC codecs operating points.

These configurations support the definition of detailed test conditions.

Table 6.4.4-2 Encoding and Decoding Configurations for Screen Content

|  |  |  |  |
| --- | --- | --- | --- |
| Encoding and Decoding Constraints | General | H.264/AVC | H.265/HEVC |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | Profiles suitable for screen content  Levels to meet the above frame rates | H.264/AVC Progressive High Profile  Level 4.2, 5.2 | H.265/HEVC Main 10 Profile  H.265/HEVC Screen-Extended Main 10 profile  Level 4.1, 5.1, 6.1 |
| Random access frequency | 1 second, infinite | 1 second, infinite | 1 second, infinite |
| Error resiliency tools | none | none | none |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | Constant quality | Fixed QP for I and P slices | Fixed QP for I and P slices |
| Bit rates and quality configuration | Covering a range of relevant bitrates and qualities | QP variations between 22 and 42 | QP variations between 22 and 42 |
| Latency requirements and specific encoding settings | Encoding adds no latency | Low-delay P | Low-delay P |
| Encoding complexity context | Real-time encoding is possible | Settings for search ranges, set to 256 | Settings for search ranges, set to 64 |
| Required decoding capabilities | Profiles suitable for screen content  Levels to meet the above frame rates | H.264/AVC Progressive High Profile  Level 4.2, 5.2 | H.265/HEVC Main 10 Profile  H.265/HEVC Screen-extended Main 10 profile  Level 4.1, 5.1, 6.1 |

End of change 6

Change 7

### 6.5.1 Motivation

According to the 2020 Mobile Internet Phenomena Report from Sandvine [9] Video traffic continues to grow worldwide, and the increasing popularity of mobile consumers sharing video has not only caused growth in downstream traffic, but also in upstream traffic as well. Instagram ™ grew in the upstream as more consumers share images and videos. TikTok ™, Snapchat ™ (video), FaceTime ™, and even Facebook ™ Live were all in the top 50 applications worldwide on the upstream that are video-sharing-centric. Messaging applications, especially on the upstream, continue to become a critical part of the mobile experience, replacing old style text messaging, and increasingly are video-based. Four of the top 20 applications on the upstream are messaging apps.

Some typical examples and restrictions in April 2020 are provided in the following:

1. WhatsApp™ [25]

a. The maximum size of the video that you can share is 16 MB.

b. Various container formats that are supported by include MP4, MKV, AVI, 3GP, and MOV. H.264/AVC video codec and AAC audio codec are needed today.

2. YouTube™ [26]

a. H.264/AVC is the recommended codec with the following settings

i. Progressive scan

ii. High Profile

iii. 2 consecutive B frames

iv. Closed GOP. GOP of half the frame rate.

v. CABAC

vi. Variable bitrate. No bitrate limit required, though we offer recommended bit rates below for reference

vii. Chroma subsampling: 4:2:0

b. Resolution Formats: 360p, 480p, 720p, 1080p, 1440p, 2160p

c. Both SDR and HDR are possible

d. The standard aspect ratio is 16:9

3. Facebook Live ™ [27]. To live stream on Facebook™, these video format guidelines are provided:

a. Recommended max bit rate is 4000 Kbps (4 mbps).

b. Max: 1080p (1920x1080) resolution, at 60 frames per second.

c. An I-frame (keyframe) must be sent at least every 2 seconds throughout the stream.

d. H264 encoded video.

4. TikTok™ [28], some video restrictions

a. Upload from Android, up to 72 MB at most. upload videos from iOS up to 287.6 MB.

b. Video dimensions should be 1080 x 1920

c. MP4 or MOV file format. Video should be H.264/AVC encoded

5. Snapchat™ [25][2)], The latest information from 2018

a. Recommended size: 1080 by 1920 pixels (9:16 aspect ratio)

b. Recommended specs: .MP4 or MOV, H.264 encoded, maximum file size 1GB

According to Sandvine's report [9], sharing and uploading content as part of social sharing is predominantly pictures and videos that uploaded directly into a cloud and uploaded to one or many social networks, and then discussed (or shared again) over messaging networks. The relevant quality-of-experience factors include the quality of shared content, the time it takes to upload, the costs associated with the upload and also the processing and battery consumption requirements on the device to prepare the content for upload.

In another activity, three large operators released a 5G messaging white paper [31] to promote enhanced messaging services based on Rich Communication Services (RCS). This also shows the relevance of operator-based messaging services. Generally, uplink resources are even more precious and costly in 5G network operation and hence efficient technologies are vital for mass-scale services. This aspect is also considered by GSMA RCS Universal Profile specification, promoted as the industry standard for RCS Business Messaging, ensuring the telecoms industry remains at the centre of digital communications [34].

### 6.5.2 Description of the Anticipated Application

In the context of 3GPP services, 5G Media Streaming [13] provides the following encoding benchmark capabilities:

- HEVC-FullHD-Enc: the capability to encode a video signal with

- up to 133,693,440 luma samples per second, and

- up to a luma picture size of 2,228,224 samples, and

- up to 240 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being either 8 or 10 bit,

to a bitstream that is decodable by a decoder that is HEVC-FullHD-Dec capable as defined in clause 4.2.2.1 of TS26.511 and defined as the capability to decode H.265 (HEVC) Main 10 Profile, Main Tier, Level 4.1 [3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

Based on the considerations in clause 6.5.1, it is also recommended to take into account the AVC-FullHD-Enc capabilities as defined in TS26.511 [13]:

- AVC-FullHD-Enc: the capability to encode a video signal with

- up to 245,760 macroblocks per second, and

- up to a frame size of 8,192 macroblocks, and

- up to 240 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being 8 bit,

to a bitstream that is decodable by a decoder that is AVC-HD-Dec capable as defined in clause 4.2.1.1 of TS26.511 and defined as the capability to decode H.264 (AVC) Progressive High Profile Level 4.0 [2] bitstreams.

Based on future expectations of higher quality uploads, it is also recommened to take into account the HEVC-UHD-Enc capabilities as defined in TS26.511 [13]:

- HEVC-UHD-Enc: the capability to encode a video signal with

- up to 534,773,760 luma samples per second, and

- up to a luma picture size of 8,912,896 samples, and

- up to 480 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being either 8 or 10 bit,

to a bitstream that is decodable by a decoder that is HEVC-UHD-Dec capable as defined in clause 4.2.2.1 of TS26.511 and defined as the capability the capability to decode H.265 (HEVC) Main 10 Profile, Main Tier, Level 5.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

The considered scenario is the uploading and uplink streaming into the ISO/BMFF and CMAF container formats. Important aspects that are expected to be considered when evaluating a codec in the context of this:

- Quality and Coding Efficiency:

- The ability to compress a video sequence targeting the maximum file size and maintaining high quality.

- The ability to compress a video stream in real time to the available uplink streaming resources.

- Considered settings for encoding:

- Regular random access at least every 2 seconds, preferably more often

- No specific encoding latency constraints are applicable

- Encoding in this scenario is typically done as

- Real-time encoding for social sharing

- Offline encoding for messaging

- UE-based Encoding

### 6.5.3 Source Format Properties

Table 6.5.3-1 provides an overview of the different source signal properties for Social Sharing and Messaging. This information is used to select proper test sequences.

Table 6.5.3-1 Source Format Properties for Social sharing scenario

|  |  |
| --- | --- |
| Source format properties | Social Sharing |
| Spatial resolution | 3840x2160, 1920 x 1080, 1080x1920 |
| Chroma format | Y’CbCr |
| Chroma subsampling | 4:2:0 |
| Picture aspect ratio | 16:9, 9:16 |
| Frame rates | 24, 25, 30 Hz  50, 60 Hz (Full HD only) |
| Bit depth | 8, 10 |
| Colour space formats | BT.709, BT.2020 |
| Transfer characteristics | BT.709, BT.2100 (HDR) |

### 6.5.4 Encoding and Decoding Constraints

Table 6.5.4-1 provides an overview of encoding and decoding constraints for H.264/AVC Full HD and H.265/HEVC for Social Sharing and Messaging scenario. This information supports the definition of detailed anchor conditions.

Table 6.5.4-1 Encoding and Decoding Configurations

|  |  |  |  |
| --- | --- | --- | --- |
| Encoding and Decoding Constraints | General | H.264/AVC | H.265/HEVC |
| Relevant Codec and Codec Profile/Levels | Profile suitable for messaging content, no specific requirements.  Levels to meet the above formats | H.264/AVC Progressive High Profile  Level 4.2, 5.2 | H.265/HEVC Main 10 Profile  Level 4.1, 5.1 |
| Random access frequency | 1 second and 10 seconds | 1 and 10 seconds | 1 and 10 seconds |
| Bit rates and quality configuration | Capped-VBR (social sharing) and VBR (messaging)  Fixed QP | B = {5, 10,15, 20} Mbps  Capped-VBR (social sharing) and VBR (messaging)  Fixed QP | B = {2.5, 5, 7.5,10} Mbps  Capped-VBR (social sharing) and VBR (messaging)  Fixed QP |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | Covering a range of relevant bitrates and qualities | No latency requirements beyond RAP so picture reordering allowed | No latency requirements beyond RAP so picture reordering allowed |
| Latency requirements and specific encoding settings | No latency requirements | No specific requirements | No specific requirements |
| Encoding complexity context | real-time encoding (social sharing), offline encoding (messaging) on mobile device, single path | tbd | tbd |
| Required decoding capabilities | Profile suitable for messaging content, no specific requirements.  Levels to meet the above formats | H.264/AVC Progressive High Profile  Level 4.2, 5.2 | H.265/HEVC Main 10 Profile  Level 4.1, 5.1 |











##### 

End of change 7

### 

Change 8

#### C.3.2.1.2 AOM Sequences

The Alliance for Open Media (AOM) sequences are used for the purpose of AOM codec development. The sequences are available for download at <https://media.xiph.org/video/av2/> and an overview is provided in the Table C.3.2.1.2-1. The following copyrights are applicable:

- Sparks : <http://creativecommons.org/licenses/by-nc-nd/4.0/>

- Meridian : <https://creativecommons.org/licenses/by-nc-nd/4.0/>

- Sol Levante : <https://creativecommons.org/licenses/by-nc-nd/4.0/>

- Cosmos Laundromat : <https://creativecommons.org/licenses/by/3.0/>

- Nocturne : <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Table C.3.2.1.2-1: AOM test material description

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name | Resolution | Frame Rate | Colour Gamut | TF | duration |
| sparks\_aom\_30-511 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 481 |
| sparks\_aom\_2024-2455 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 431 |
| sparks\_aom\_5363-5763 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 400 |
| sparks\_aom\_5764-6024 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 260 |
| sparks\_aom\_6026-6502 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 476 |
| sparks\_aom\_8396-8941 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 545 |
| sparks\_aom\_9774-10071 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 297 |
| sparks\_aom\_11198-11570 | 4096 x 2160 | 59.94 | P3-D65 | PQ | 372 |
| meridian\_aom\_1782-2163 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 381 |
| meridian\_aom\_11872-12263 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 391 |
| meridian\_aom\_12264-12745 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 481 |
| meridian\_aom\_15932-16309 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 377 |
| meridian\_aom\_20988-21412 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 424 |
| meridian\_aom\_22412-22738 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 324 |
| meridian\_aom\_24058-24550 | 3840 x 2160 | 59.94 | P3-D65 | PQ | 492 |
| sol\_levante\_aom\_289-453 | 3840 x 2160 | 24 | P3-D65 | PQ | 164 |
| sol\_levante\_aom\_519-649 | 3840 x 2160 | 24 | P3-D65 | PQ | 130 |
| sol\_levante\_aom\_2268-2412 | 3840 x 2160 | 24 | P3-D65 | PQ | 144 |
| sol\_levante\_aom\_3282-3874 | 3840 x 2160 | 24 | P3-D65 | PQ | 592 |
| sol\_levante\_aom\_4123-4545 | 3840 x 2160 | 24 | P3-D65 | PQ | 422 |
| cosmos\_aom\_1573-1749 | 3840 x 2160 | 24 | P3-D65 | PQ | 176 |
| cosmos\_aom\_8686-8826 | 3840 x 2160 | 24 | P3-D65 | PQ | 176 |
| cosmos\_aom\_9561-9789 | 3840 x 2160 | 24 | P3-D65 | PQ | 228 |
| cosmos\_aom\_11589-11752 | 3840 x 2160 | 24 | P3-D65 | PQ | 163 |
| cosmos\_aom\_12025-12075 | 3840 x 2160 | 24 | P3-D65 | PQ | 50 |
| cosmos\_aom\_12149-12330 | 3840 x 2160 | 24 | P3-D65 | PQ | 181 |
| cosmos\_aom\_12916-13078 | 3840 x 2160 | 24 | P3-D65 | PQ | 162 |
| cosmos\_aom\_13446-13649 | 3840 x 2160 | 24 | P3-D65 | PQ | 203 |
| nocturne\_aom\_2370-2539 | 3840 x 2160 | 60 | P3-D65 | PQ | 169 |
| nocturne\_aom\_8540-9009 | 3840 x 2160 | 60 | P3-D65 | PQ | 469 |
| nocturne\_aom\_9010-9349 | 3840 x 2160 | 60 | P3-D65 | PQ | 339 |
| nocturne\_aom\_17140-17709 | 3840 x 2160 | 60 | P3-D65 | PQ | 569 |
| nocturne\_aom\_18013-18315 | 3840 x 2160 | 60 | P3-D65 | PQ | 302 |
| nocturne\_aom\_23820-24322 | 3840 x 2160 | 60 | P3-D65 | PQ | 502 |
| nocturne\_aom\_27740-28109 | 3840 x 2160 | 60 | P3-D65 | PQ | 369 |
| nocturne\_aom\_32660-32799 | 3840 x 2160 | 60 | P3-D65 | PQ | 139 |

The clip cosmos\_aom\_12025-12075 being fifty frames duration, it is removed from the selection process and is not considered. In order to use these sequences for the selection process, a conversion into the appropriate format is conducted using FFmpeg and HDRTools [65] software v0.21. The conversion according to Figure C.3.2.1.2-1 conducted:

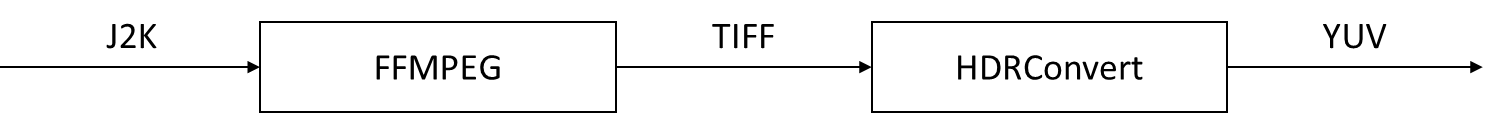


Figure C.3.2.1.2-1: Processing pipeline for AOM sequence conversion

In the first step FFMPEG was used to decode JPEG 2000 sequences and store in uncompressed TIFF format. The following command line was used:

FFMPEG -i $INPUT.j2k -compression\_algo raw -pix\_fmt rgb48le $OUTPUT.tiff

HDRConvert was used to convert the sequences from P3-D65 RGB 12 bit to Rec. BT.2100 10 bit YCbCr 4:2:0. HDRConvert config file used for this process is attached to this contribution.

#### C.3.2.1.3 CableLabs Sequences

The CableLabs dataset is composed by one HDR sequence described in Table C.3.2.1.3-1. The sequence can be used under the CC BY-NC-ND 3.0 license restriction.

Table C.3.2.1.3-1: Cable Labs HDR test material description

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name | Resolution | Frame rate | Colour Gamut | TF | duration |
| Life Untouched | 3840 x 2160 | 59.94 | P3-D65 | PQ | 5:20 |

For this sequence, a selection of some critical time intervals have been conducted based on experience on compression. Especially intervals with fast or complex motion, representative of TV contents have been selected. Static pictures, or pictures with artistic intent with few spatial details have been discarded. To remain in the scope of the scenario, the following section have been selected:

* Life Untouched 1: 0.5s-8s

The extraction of subpart from original ProRes sequences is realized with FFmpeg release 4.3.1 using the following command line:

./ffmpeg.exe -i $IN.mov -vcodec rawvideo -ss $START\_TIME -t $DURATION -pix\_fmt yuv444p12le $OUT.yuv

Where $IN and $OUT are respectively the input and output sequences, $START\_TIME and $DURATION respectively the start time and duration in hh:mm:ss format.

In order to map the colour into a BT.2020 colour container, HDRTools v0.21 has been used. In the first step conversion from YCbCr 4:4:4 to RGB color space was performed. In the second step conversion from P3-D65 RGB to 10-bit YCbCr 4:2:0 was performed. Both configuration files for HDRTools are attached to this report.

End of change 8

Change 9

### C.3.2.3 Selected Sequences

#### C.3.2.3.1 Life Untouched

The Life Untouched test sequence shows nature scenes. The selected part starts from fade from black and includes camera tilt.

Figure C.3.2.3.1-1 shows a screenshot of the Life Untouched test sequence.

A picture containing outdoor, tree, sky, plant

Description automatically generated

Figure C.3.2.3.1-1: Screenshot of Life Untouched test sequence

The Life Untouched test sequence properties are provided in Table C.3.2.3.1-1.

Table C.3.2.3.1-1 Life Untouched sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 3840 x 2160 |
| Scan | progressive |
| Frame Rate | 60/1.001 |
| Bit Depth | 10 |
| Length | 450 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Life-Untouched/Life-Untouched.json

The Life Untouched test sequence is made available under the following copyright disclaimer.

*Copyright © CableLabs LifeUntouched sequence is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. https://creativecommons.org/licenses/by-nc-nd/3.0/deed.en\_US.*

#### C.3.2.3.2 Meridian

The Meridian test sequence is tbd(meridian\_aom\_22412-22738).

Figure C.3.2.3.2-1 shows a screenshot of the Meridian test sequence.



Figure C.3.2.3.2-1: Screenshot of Meridian test sequence

The Meridian test sequence properties are provided in Table C.3.2.3.2-1.

Table C.3.2.3.2-1 Meridian sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 3840 x 2160 |
| Scan | progressive |
| Frame Rate | 60/1.001 |
| Bit Depth | 10 |
| Length | 327 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Meridian/Meridian.json

The Meridian test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence provided above and all intellectual property rights therein remain the property of Netflix Inc. This video sequence is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>*

#### C.3.2.3.3 Sol Levante

The Sol Levante test sequence is tbd(sol\_levante\_aom\_2268-2412).

Figure C.3.2.3.3-1 shows a screenshot of the Sol Levante test sequence.



Figure C.3.2.3.3-1: Screenshot of Sol Levante test sequence

The Sol Levante test sequence properties are provided in Table C.3.2.3.3-1.

Table C.3.2.3.3-1 Sol Levante sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 3840 x 2160 |
| Scan | progressive |
| Frame Rate | 24 |
| Bit Depth | 10 |
| Length | 145 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Sol-Levante/Sol-Levante.json

The Sol Levante test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence provided above and all intellectual property rights therein remain the property of Netflix Inc. This video sequence is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>*

#### C.3.2.3.4 Cosmos

The Cosmos test sequence is tbd(cosmos\_12149-12330).

Figure C.3.2.3.4-1 shows a screenshot of the Cosmos test sequence.



Figure C.3.2.3.4-1: Screenshot of Cosmos test sequence

The Cosmos test sequence properties are provided in Table C.3.2.3.4-1.

Table C.3.2.3.4-1 Cosmos sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 3840 x 2160 |
| Scan | progressive |
| Frame Rate | 24 |
| Bit Depth | 10 |
| Length | 182 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Cosmos/Cosmos.json

The Cosmos test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence "Cosmos Laundromat" has been created by Blender Foundation. This version of the sequence*

*has been remastered as the high dynamic range and wide color gamut video by Netflix Inc. This video sequence*

*is licensed under the Creative Commons Attribution 3.0 International License. To view a copy of this license,*

*visit http://creativecommons.org/licenses/by/3.0*

#### C.3.2.3.5 Elevator

The Elevator test sequence is tbd(sparks\_aom\_2024-2455).

Figure C.3.2.3.5-1 shows a screenshot of the Elevator test sequence.



Figure C.3.2.3.5-1: Screenshot of Elevator test sequence

The Elevator test sequence properties are provided in Table C.3.2.3.5-1.

Table C.3.2.3.5-1 Elevator sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 4096 x 2160 |
| Scan | progressive |
| Frame Rate | 60/1.001 |
| Bit Depth | 10 |
| Length | 432 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Elevator/Elevator.json

The Elevator test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence provided above and all intellectual property rights therein remain the property of Netflix Inc. This video sequence is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>*

#### C.3.2.3.6 Sparks

The Sparks test sequence is tbd(sparks\_aom\_5764-6024).

Figure C.3.2.3.6-1 shows a screenshot of the Sparks test sequence.



Figure C.3.2.3.6-1: Screenshot of Sparks test sequence

The Sparks test sequence properties are provided in Table C.3.2.3.6-1.

Table C.3.2.3.6-1 Sparks sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 4096 x 2160 |
| Scan | progressive |
| Frame Rate | 60/1.001 |
| Bit Depth | 10 |
| Length | 261 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Sparks/Sparks.json

The Sparks test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence provided above and all intellectual property rights therein remain the property of Netflix Inc. This video sequence is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>*

#### C.3.2.3.7 Nocturne

The Nocturne test sequence is tbd(nocturne\_aom\_27740-28109).

Figure C.3.2.3.7-1 shows a screenshot of the Nocturne test sequence.



Figure C.3.2.3.7-1: Screenshot of Nocturne test sequence

The Nocturne test sequence properties are provided in Table C.3.2.3.7-1.

Table C.3.2.3.7-1 Nocturne sequence properties

|  |  |
| --- | --- |
| Parameter | Value |
| Resolution | 4096 x 2160 |
| Scan | progressive |
| Frame Rate | 60/1.001 |
| Bit Depth | 10 |
| Length | 370 |
| YUV format | 4:2:0 |
| Colour components | ITU-T BT.2020 |
| Colour space | HDR PQ |

The sequence can be accessed

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Nocturne/Nocturne.json

The Nocturne test sequence is made available under the following copyright disclaimer.

*Copyright ©*

*The video sequence provided above and all intellectual property rights therein remain the property of Netflix Inc. This video sequence is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>*

End of change 19

Change 10

## C.6.6 Mission Control

### C.6.6.1 Introduction

The sequence Mission-Control combines natural video with synthetic content. Figure C.6.6.1-1 is a screenshot of each of the 3 scenes contained in the *Mission-Control* test sequence.

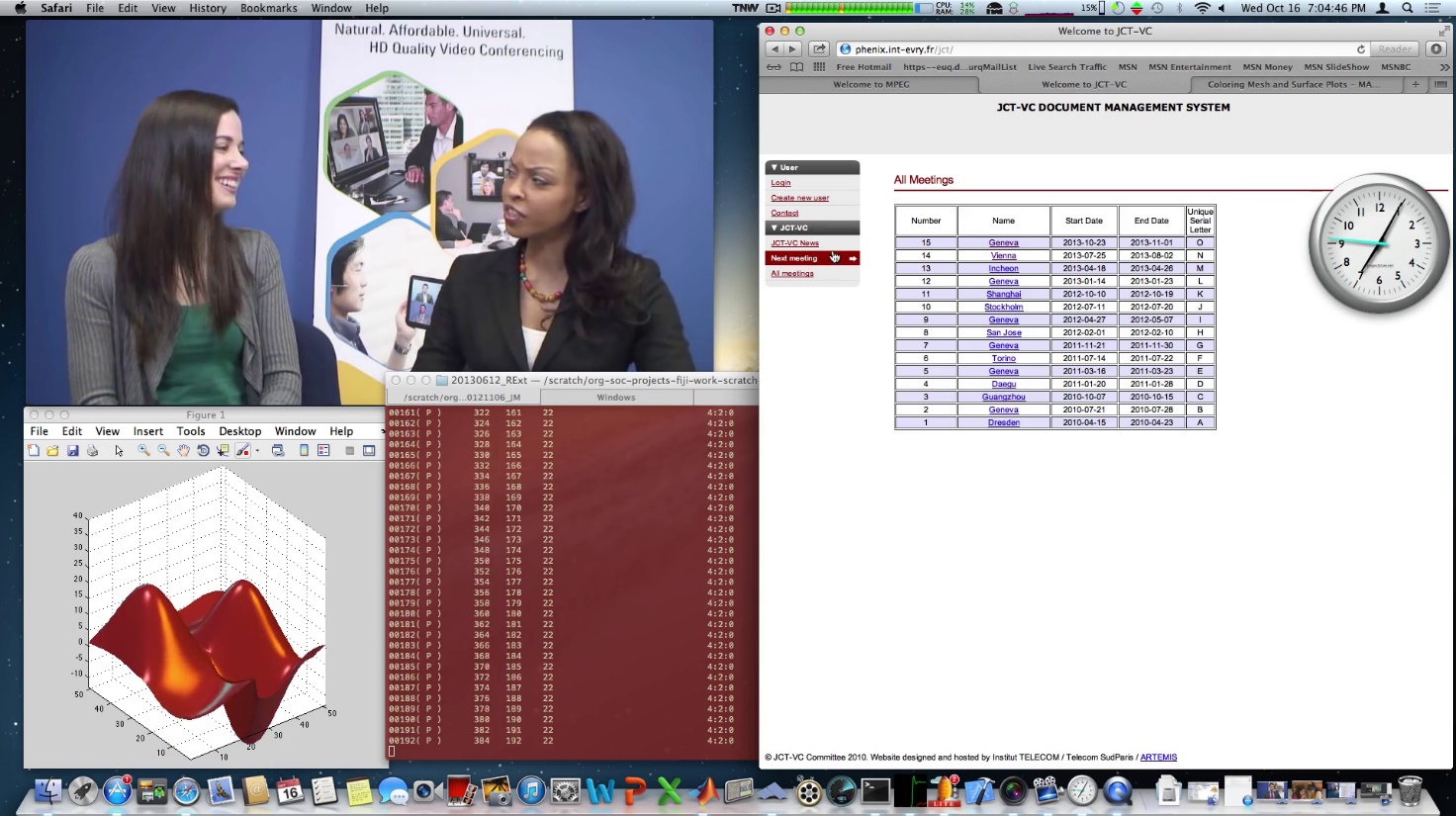
**

Figure C.6.6.1-1: Screenshot of Mission-Control test sequence

### C.6.6.2 Source sequence properties

The *Mission-Control* test sequence properties are provided in Table C.6.6.2-1.

Table C.6.6.2-1 Mission-Control test sequence

|  |  |
| --- | --- |
| Parameter | GMT-03 |
| Resolution | 1920 x1080 |
| Scan | Progressive |
| Frame Rate | 60 fps |
| Bit Depth | 10 |
| Length | 600 frames (10s) |
| YUV format | YUV 4:2:0 |
| Color components | Y’CbCr |
| Colour space | ITU-R BT.709 |

The sequence can be accessed here:

- https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/Mission-Control/Mission-Control.json

### C.6.6.3 Copyright information

*The test sequence and all intellectual property rights therein remain the property of their respective copyright holders. This material can only be used for the purpose of academic research and development of 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.*

*Owners of this sequence: Apple Inc.*

*Production: Apple Inc.*

*The sequence Kristen and Sara included in this video is property of Vidyo, Inc.*

End of change 10