|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *CR-Form-v12.0* | | | | | | | | |
| **Pseudo CHANGE REQUEST** | | | | | | | | |
|  | | | | | | | | |
|  | **26.956** | **CR** | **<CR#>** | **rev** | **-** | **Current version:** | **0.0.1** |  |
|  | | | | | | | | |
| *For* ***[HE](http://www.3gpp.org/3G_Specs/CRs.htm" \l "_blank)******[LP](http://www.3gpp.org/3G_Specs/CRs.htm" \l "_blank)*** *on using this form: comprehensive instructions can be found at  <http://www.3gpp.org/Change-Requests>.* | | | | | | | | |
|  | | | | | | | | |

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

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | |
| ***Title:*** | [FS\_Beyond2D] Proposed Scenario: Free Viewpoint Video | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Source to WG:*** | China Mobile Com. Corporation | | | | | | | | | |
| ***Source to TSG:*** | SA4 | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Work item code:*** | FS\_ Beyond2D | | | | |  | ***Date:*** | | | 2024-05-14 |
|  |  | | | |  | |  | | |  |
| ***Category:*** | **B** |  | | | | | ***Release:*** | | | Rel-19 |
|  | *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:*** | | The study item asks for typical scenarios | | | | | | | | |
|  | |  | | | | | | | | |
| ***Summary of change:*** | | Free-viewpoint video is an immersive experience in which a viewer can actively select any viewpoint to view the action of a scene of a 3D world. | | | | | | | | |
|  | |  | | | | | | | | |
| ***Consequences if not approved:*** | | Missing relevant B2D experience | | | | | | | | |
|  | |  | | | | | | | | |
| ***Clauses affected:*** | | Annex A | | | | | | | | |
|  | |  | | | | | | | | |
|  | | **Y** | **N** |  | | | |  | | |
| ***Other specs*** | |  | **X** | Other core specifications | | | | TS/TR ... CR ... | | |
| ***affected:*** | |  | **X** | Test specifications | | | | TS/TR ... CR ... | | |
| ***(show related CRs)*** | |  | **X** | O&M Specifications | | | | TS/TR ... CR ... | | |
|  | |  | | | | | | | | |
| ***Other comments:*** | | This pCR is based on TR26.956v0.0.1 | | | | | | | | |
|  | |  | | | | | | | | |
| ***This CR's revision history:*** | |  | | | | | | | | |

# Introduction

During SA4#127-bis e-meeting and Video SWG Adhoc post 127-bis-e, some scenarios and related proposals have been discussed. According to market-relevant discussions, we believe that providing a B2D experience on market-relevant UEs is critical point. Therefore, for Free Viewpoint Video (FVV) or Free Viewpoint TV (FTV) scenarios,

* For display end, it is limited to traditional 2D UE, including mobile devices [3], PCs or TVs [1][2].
* For collection end, it is further away from the B2D experience. It can support media pre-processing based on B2D experience and performance. Most of the commercial applications use camera matrix to collect [3], and in some current end-to-end solutions, you can also use mobile devices to capture multiple video streams [4].

# Proposal

|  |
| --- |
| **1st Change** |

6.X Scenario X: Free Viewpoint Video

**6.X.1 Motivation**

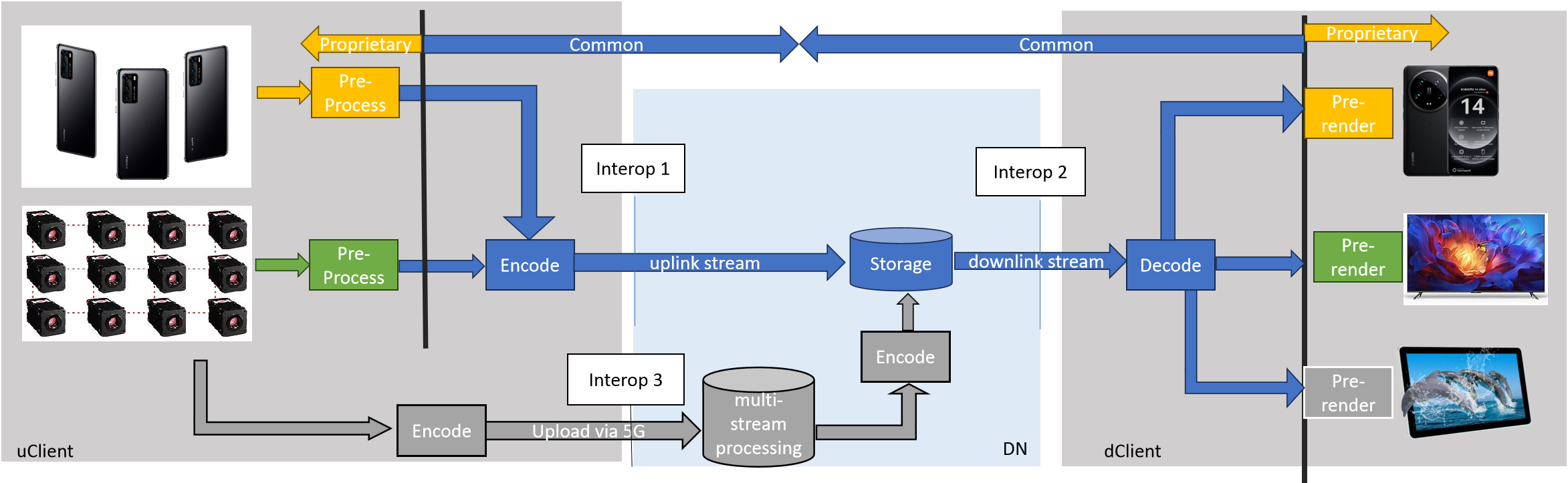
Free-viewpoint video is an immersive experience in which a viewer can actively select any viewpoint to view the action of a scene of a 3D world. It uses multiple high-precision cameras to shoot multi-angle images synchronously at the millisecond level to create a free viewpoint scene so that the user can freely move and watch from the selected view within the scene, including looking up and down, back and forth, left and right. FVV has commercial value in sports events, TV programs, online education, advertising and marketing.

Multi-view video plus depth (MVD) is a source format of FVV which has limited viewpoints of video and associated depth maps. The depth maps and disparity maps associated with every view of the video that can be used to render arbitrary numbers of additional views via view synthesis to approximate the experience of free viewpoint. From the user’s perspective, they can choose the position and viewpoint arbitrarily and watch content of this view on mobile devices or traditional TVs with view control.

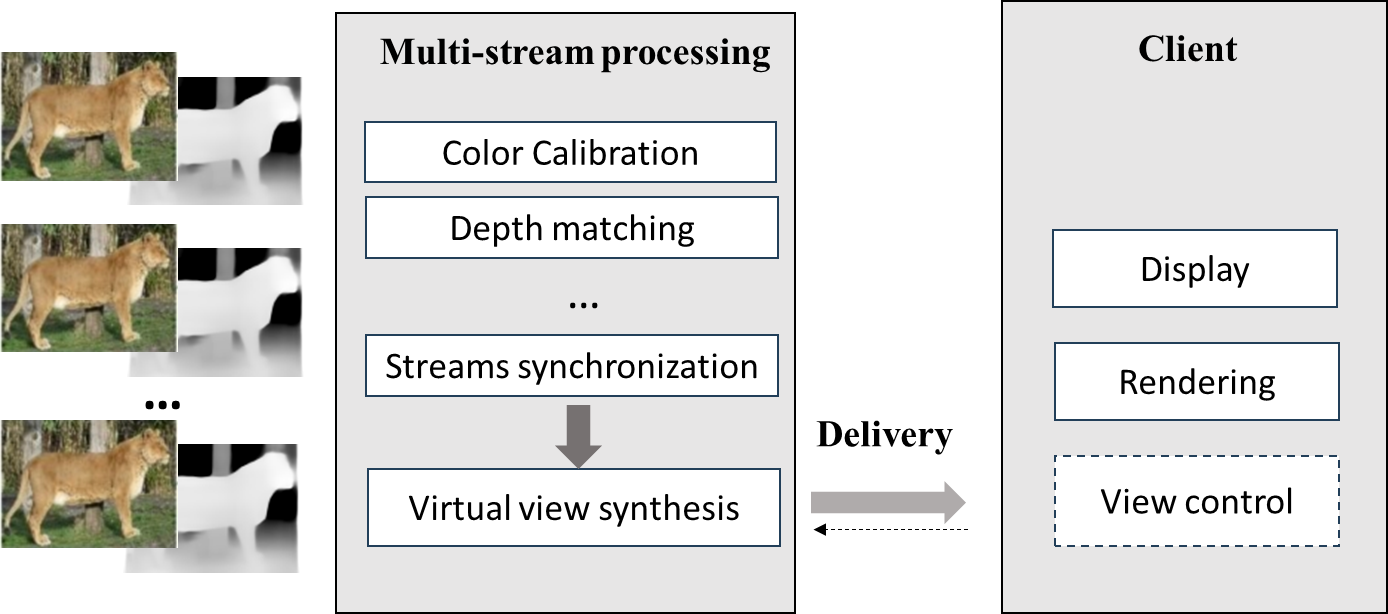
The scenario is shown in the below figure.

• For display end, it is limited to traditional 2D UE, including mobile devices, PCs or TVs

• For collection end, it is further away from the B2D experience. It can support media pre-processing based on B2D experience and performance. Most of the commercial applications use camera matrix to collect, and in some current end-to-end solutions, you can also use mobile devices to capture multiple video streams.



**6.X.2 Description of the scenario**



a. Capturing and processing: RGBD camera, color calibration, multi-streams synchronization, depth matching.

b. Encoding: texture map, depth map and camera parameters.

c. Packaging and delivery: synthesis views or a view based on view control can be delivered to client.

d. Decoding: rendering on client

e. General constraints on bandwidth: the number of views should trade off prediction error and bandwidth.

**6.X.3 Source Formats**

Table 6.X.3-1 provides an overview of typical beyond 2D source signal properties for FVV.

**Table 6.x.3-1 FVV Source Format Properties**

|  |  |
| --- | --- |
| Source format properties | Texture |
| Number of views | 3,6 |
| Spatial resolution for each view | 1920 x 1080 |
| Chroma format | Y’CbCr/RGB |
| Chroma subsampling | 4:2:0 |
| Picture aspect ratio | / |
| Frame rates | 30, 50, 60 Hz |
| Bit depth | 10 |
| Colour space formats | BT.709, BT.2020 |
| Transfer characteristics | BT.709, BT.2100 (HDR) |
| Camera parameters | camera id, intrinsic parameters\*,extrinsic parameters |
| **Depth/Disparity** | |
| Spatial Resolution | Same resolution as view |
| Chroma format | Y’CbCr |
| Chroma subsampling | 4:0:0 |
| Picture aspect ratio | / |
| Frame rates | 30, 50, 60 Hz |
| Bit depth | 10 |
| Camera parameters | camera id, intrinsic parameters,extrinsic parameters |

**6.x.4 Encoding and Decoding Constraints**

Beyond 2D encoding based on HEVC main 10 in this scenario is typically done as cloud encoding and UE Encoding for three interoperability points.

**Table 6.5.4-1 Encoding and Decoding Configurations**

|  |  |  |  |
| --- | --- | --- | --- |
| Encoding and Decoding Constraints | Interop1 | Interop2 | **Interop3** |
| Relevant Codec and Codec Profile/Levels | Multiview HEVC profile | Multiview HEVC profile | Multiple H.265/HEVC Main 10 Level 4.1;depth |
| Random access frequency | 1 | 1 | 1 |
| Bit rates and quality configuration | Fixed QP | Fixed QP | Fixed QP |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | Covering a range of  relevant bitrates and  qualities | Covering a range of  relevant bitrates and  qualities | Covering a range of  relevant bitrates and  qualities |
| Latency requirements and specific encoding settings | No latency requirements | No latency requirements | No latency requirements |
| Encoding complexity context | encoding on UE client | encoding on the server | encoding on UE client |
| Required decoding capabilities | Multiview HEVC profile | Multiview HEVC profile | Multiple H.265/HEVC Main 10 Level 4.1;depth |

**6.5.5 Performance Metrics**

The following metrics are expected to be reported: TBD

The key to synthesize viewpoint is how to accurately infer the information of new perspective from the existing data and maintain the coherence and realism of the scene. This often involves complex image processing and computer graphics techniques. In practice, the quality of the synthetic viewpoint is affected by the quality of the raw data, the performance of the acquisition device, and the algorithms and models used.

**6.X.6 Interoperability Considerations**

**6.X.7 Reference Sequences**

<tbd> A Free-Viewpoint-RGB-D-Video-Dataset is provided here <https://medialab.sjtu.edu.cn/post/free-viewpoint-rgb-d-video-dataset/>

**6.X.8 Anchor Definition**

<tbd>

|  |
| --- |
| **3rd Changes** |

3 reference

[1] Nguyen-Ha et al. Free-Viewpoint RGB-D Human Performance Capture and Rendering. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13676. Springer, Cham. <https://doi.org/10.1007/978-3-031-19787-1_27>

[2] P. Pérez et al., "Live Free-Viewpoint Video in Immersive Media Production Over 5G Networks," in IEEE Transactions on Broadcasting, vol. 68, no. 2, pp. 439-450, June 2022, doi: 10.1109/TBC.2022.3154612.

[3] Tencent Media Lab, Free-Viewpoint Video <https://multimedia.tencent.com/products/free-viewpoint-video>

[4] Huawei Cloud, <https://www.huaweicloud.com/product/live/fvv.html>