**3GPP TSG SA WG4 Meeting #128 *-241068***

**Jeju,South Korea 4**

|  |
| --- |
| *CR-Form-v12.2* |
| **pseudo CHANGE REQUEST** |
|  |
|  |  | **CR** |  | **rev** | - | **Current version:** | **0** |  |
|  |
| *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 | **X** | Radio Access Network |  | Core Network | **X** |

|  |
| --- |
|  |
| ***Title:***  | [IBACS]pCR Spatial and Scene descriptions |
|  |  |
| ***Source to WG:*** | Nokia |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | IBACS |  | ***Date:*** | 14 May 2024 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***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 canbe 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-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19)* |
|  |  |
| ***Reason for change:*** | Ambiguity about spatial and scene descriptions in the TS |
|  |  |
| ***Summary of change:*** | 1. Restructuring of clause 6.3.1
2. Introduction of a sub clause 6.3.2
3. Expansion of A.1.4
 |
| ***tr*** |   |
| ***Consequences if not approved:*** | The TS does not address how spatial descriptions and scene descriptions are used in an AR call |
|  |  |
| ***Clauses affected:*** |  |
|  |  |
|  | **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 CR's revision history:*** | N/A |

1. Start of Change 1

## 6.3 Spatial descriptions

### 6.3.1 Spatial description format

#### 6.3.1.1 General

A spatial description format is used for defining the physical space around a UE or trackable in which. This clause includes the supported formats and the method for exchanging the information between AR-MTSI clients.

#### 6.3.1.2 Available visualization space

An AR-MTSI client in terminal may send available visualization space, user position and other trackable poses to AR MF/MRF for scene creation and update.

The available visualization space defines an occlusion-free space around the user for rendering the AR scene as a geometric primitive. The format for available visualization space is defined in clause 6.2.4 of [3]. The type of the message containing visualization space as a payload shall be “**urn:3gpp:ar:v1:visualization-space**”. The availableVisualizationSpace object [3] shall contain a xrSpaceId. The xrSpaceId is used for determining the local coordinate axis of the visualization space. The xrSpaceId shall be a unique identifier for an XR space of one AR-MTSI client in terminal. The viewer pose is assumed to be the center of the visualization space, i.e., orientation [x,y,z,w] = [0,0,0,1] and position [x,y,z] =[0,0,0].

If the visualization space is not anchored around the user or if the viewer is not the centre of the visualization space, an initial pose as defined in clause 6.3.1.3 can be used.

6.3.1.3 Initial Pose

Trackable is a real-world object (e.g., the UE, floor, controllers, table etc.) that the UE can detect, which can be used as a reference to anchor virtual objects to the real world.

The AR-MTSI client in terminal that sends the available visualization space may also send at least one pose for a trackable e.g., to determine the position of the user within the visualization space. The AR-MTSI client in terminal may send additional poses for anchoring virtual objects. The poses shall be sent using the format defined in clause 6.2.2 of [3]. The poseInfo (as defined in Table 6.2.2-1 [3]) shall contain an xrSpaceId that is the same as the one used for visualization space. The poses may additionally contain a label string to identify the type of anchor. The labels are application-dependent, but for example, user, floor, left controller etc., can be used as labels. The type of the message for a pose sent for scene creation shall be set to “**urn:3gpp:ar:v1:initial-pose**”.

1. End of Change 1
2. Start of Change 2

6.3.2. Spatial descriptions processing

An AR-MTSI client may share visual features extracted from its surroundings (using algorithms like ORB, SIFT, etc.) with the MF for spatial computing. Visual features shall be sent using application data channel using AR metadata message format as defined in clause 6.2. The type of the message containing visual features as a payload shall be “**urn:3gpp:ar:v1:visual features**”.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name**  | **Type**  | **Cardinality**  | **Description**  |
| XRSpaceId  |   |   | The XR space identifier used.   |
| cameraIntrinsics  | string | 1..1  | Camera lens intrinsic parameters signalled as a string using the format defined for 3gpp-camera-calibration as defined in TS 26.114 clause Y.6.6. |
| cameraPose  | object  | 1..n  | Camera extrinsics or current pose of the camera. The format shall be based on poseInfo (as defined in Table 6.2.2-1 [3]) |
| featureId  | string  | 0..1  | A unique identifier of the feature.  |
| virtualFrameId  | string  | 0..1  | A unique identifier of the frame used by the device for extracting features. This can be a keyframe identifier for easy mapping in the XR Runtime.   |
| featureType  | string  | 1..1  | A string indicating the type of features. The value shall be “visual” for visual features extracted from frames captured by the UE camera.  Other application-defined feature types may be used if allowed by the bootstrap application.  |
| featureDescriptor  | string  | 1..1  | Identifies the localization technique used to extract the feature. The following values may be used:  “SIFT” “ORB”   |
| featureData  | object  | 0..1  | Vector of numbers (floats or integers) defined by the localization algorithm being used.   |
| featurePositions  | array  |   |   |
|     featurePosition   | object  | 0..1  | Array of x and y coordinates of features in the keyframe with respect to the origin which is the top left corner of the frame.  The x and y coordinates are expressed in pixels.  |
| frameResolution  | string  | 0..1  | The resolution of the frame in pixels expressed as a string of type “1024x786”, where  1024 is the resolution along the horizontal axis and y is the resolution along the vertical axis  |
|   |   |   |   |

An MF can create 3D points based on the visual features for spatial mapping. An MF facilitating AR applications that require real-world interactions shall include the XR spatial description using scene description capabilities of UEs as specified in TS 26.119.

A spatial description shall consist of real-world objects as meshes or point clouds and shall use glTF 2.0 scenes as specified in [X]. The gLTF scene consisting of spatial description may also include virtual XR objects.

1. End of Change 2
2. Start of Change 3

## A.1.4 Scene Description Processing

Figure A.1.4-1 illustrates a detailed call flow for scene description processing procedure.



Figure A.1.4-1: Scene Description Processing

The steps are as follows:

1. The UE1 initiates an AR communication session and establishes audio and video session connections with the UE2. Then the bootstrap and application data channels are established for the UE1 and UE2.

2. The split rendering negotiation procedure has been finished.

3. [Optional] The AR AS provides an initial scene description to the MF. This initial scene description may be specific to the service provided by the AR-AS or AP

4,5. UE1 and UE2 send spatial information to the MF. The spatial information may be spatial descriptions as specified in clause 6.3.1. If a UE is unable to perform spatial mapping and localization, for example, due to low computational or energy resources, it may share only spatial visual features, as specified in clause 6.3.1.3, with the MF for the MF to perform spatial mapping and localization.

6. MF/MRF prepares the scene description based on received spatial information, media descriptions and assets for the call.

7,8. MF/MRF delivers the initial scene description to the UEs.

9. A UE may trigger a scene update e.g., when a new object is added/removed in the scene, or a spatial information update is sent or UE capabilities change. The figure shows the update is triggered by UE1, but this can be either UE.

10. The MF/MRF processes the new information and creates a scene description update. It is also possible for the MF/MRF to initiate an update without an update from the UEs, for example if link conditions change.

11,12. MF/MRF distributes scene description update to all UEs.

NOTE: Spatial data related updates may be required for collaborative AR calls, e.g., when multiple users are physically collocated and also part of the same AR experience. The type of spatial description updates is FFS.

13. Subsequent procedures continue.

1. End of Change 3