**3GPP TSG- Meeting # *xxx***

**, , - revision of S4-241141**

|  |
| --- |
| *CR-Form-v12.3* |
| **CHANGE REQUEST** |
|  |
|  |  | **CR** |  | **rev** | **3** | **Current version:** |  |  |
|  |
| *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 |  |

|  |
| --- |
|  |
| ***Title:***  |  |
|  |  |
| ***Source to WG:*** |  |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** |  |  | ***Date:*** |  |
|  |  |  |  |  |
| ***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 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-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19) Rel-20 (Release 20)* |
|  |  |
| ***Reason for change:*** | • The terms, e.g., UE, server, RTP client, are too particular and not consistent.• The RTP header extension for the XR pose has two meanings: “pose to render” and “rendered pose”. In the case of “pose to render”, there is no association between the XR pose in the RTP header extension and the media in the RTP payload. In the case of “rendered pose”, the media in the RTP paylaod has been rendered using the XR pose in the RTP header extension. Some text on how to determine the meaning is needed.  |
|  |  |
| ***Summary of change:*** | * Removed UE, server, RTP client, and used RTP sender, RTP receiver and RTP endpoint instead.
* Added a note – NOTE 4 - on determming the meaning of the HE.
 |
|  |  |
| ***Consequences if not approved:*** | Unclear and inconsistent specification content relative to the definition and usage of RTP header extension for XR pose. |
|  |  |
| ***Clauses affected:*** | 4.3.1, 4.3.2, 4.3.3, Annex B.1 |
|  |  |
|  | **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:*** | S4-241xxx CR0002r3: Removing of changes over changes editsS4-241141 CR0002r2: Corrections to the cover pageS4-241055 CR0026r1: Updates based on received comments to remove additional paragaph on ambiguity of RTP HE for XR pose.S4-240725 CR0026r0: Upgraded pCR to CR stage since TS 26.522 sent for approval and approved by TSG SA plenary in March meetingS4aR240013: pCR with original content |

\* \* \* \* 1st change \* \* \* \*

## 4.3 RTP Header Extension for XR Pose

### 4.3.1 General

An RTP sender that uses RTP to deliver rendered video streams to an RTP receiver should include an RTP HE for XR pose to indicate the XR pose used for rendering the media (rendered pose). The RTP HE for XR pose may be used for signaling either a 6DoF XR pose or a 3DoF XR pose. The RTP HE for XR pose may also be used with audio streams.

The RTP HE for XR pose may also be used by an RTP sender to indicate the XR pose to be rendered to an RTP receiver.

The IANA registration information for the RTP HE for XR pose is provided in Annex D.3.

### 4.3.2 SDP Signaling

An RTP client that supports the RTP HE for XR pose shall negotiate the use of the extension using SDP. The signaling of the RTP HE for XR pose shall follow the SDP signaling design, the syntax, and semantics of the "extmap" attribute as outlined in RFC8285 [11].

For IANA registration, the "reference" field in the registry is 3GPP TS 26.522.

The ABNF syntax for this RTP HE extends the "extmap" attribute as follows:

*extensionname* = "urn:3gpp:xr-pose"

*extensionattributes* = "3DOF" / "6DOF" ["media:" 1\*(SP token)]

The extension attribute "3DOF" indicates that the sender uses the RTP HE to signal a 3DoF XR pose, i.e., an XR pose that does not include the position fields x, y, z.

An RTP client that supports the RTP HE for XR pose and receives an SDP offer with "a=extmap" attribute with the URN "urn:3gpp:xr-pose" and the extension attribute "3DOF", shall include the extension attribute "3DOF" in the SDP answer, if the SDP offer is accepted.

The extension attribute "6DOF" indicates that the sender uses the RTP HE to signal a 6DoF XR pose, i.e., an XR pose that includes both the position fields x, y, z and the orientation fields rx, ry, rz, rw.

An RTP client that supports the RTP HE for XR pose and receives an SDP offer with "a=extmap" attribute with the URN "urn:3gpp:xr-pose" and the extension attribute "6DOF", shall include the extension attribute "6DOF" in the SDP answer, if the SDP offer is accepted.

The extension attribute "media" is followed by a list of tokens for "mid" (as defined in RFC 5888 [19]) for media streams that can reuse the pose included in the RTP HE. Further details on reuse are provided later in the section.

An RTP endpoint that supports the RTP HE for XR pose and receives an SDP offer with "a=extmap" attribute with the URN "urn:3gpp:xr-pose" shall remove the attribute from the answer for any media that will not use the extension, and retain it for any media that will use it.

### 4.3.3 Header Extension Format

If the RTP HE for XR pose is for rendered pose, the RTP sender should use the RTP HE for XR pose to associate the selected pose with the rendered frame. The RTP sender delivers the rendered frames using one or more video streams, depending on the view and projection configuration that is selected by the UE.

If negotiated successfully, an RTP sender should add the RTP HE for XR pose to the RTP stream. The frequency of RTP HE for XR pose shall be at least once in a frame. It may be sent more often but not necessarily in every RTP packet.

The 2-byte (RFC 8285) RTP HE format shall be used for signalling the RTP HE. Format of the HE for 6DoF XR pose is shown below.

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0x100 |appbits| length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ID | L | rx …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

 | ry …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | rz …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | rw …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | x …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | y …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | z …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | XR timestamp …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| XR timestamp continued …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| XR timestamp continued | action\_id #1 |

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| action\_id #2 | ... |

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

If the RTP HE for XR pose is used for signaling a 3DoF XR pose, the fields x, y, z shall be omitted. Format of the HE for 3DoF XR pose is shown below.

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0x100 |appbits| length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ID | L | rx …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

 | ry …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | rz …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | rw …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | | XR timestamp …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| XR timestamp continued …

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| XR timestamp continued | action\_id #1 |

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| action\_id #2 | ... |

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

The fields rx, ry, rz, rw, x, y, z are defined in single-precision floating-point format (binary32 as per ISO/IEC 60559 [20]).

**rx (32 bits):** x coordinate of the orientation quaternion of the XR pose.

**ry (32 bits):** y coordinate of the orientation quaternion of the XR pose.

**rz (32 bits):** z coordinate of the orientation quaternion of the XR pose.

**rw (32 bits):** w coordinate of the orientation quaternion of the XR pose.

**x (32 bits):** x coordinate of the position of the XR pose in meters.

**Y (32 bits):** y coordinate of the position of the XR pose in meters.

**z (32 bits):** z coordinate of the position of the XR pose in meters.

**XR timestamp (64 bits)**: Timestamp for the XR pose. If the RTP HE is used for rendered pose, this timestamp indicates the display time predicted by the XR runtime for the rendered image. Otherwise, this timestamp indicates the associated XR runtime display time for the predicted XR pose. XR timestamp uses the XR system clock and is represented in nanoseconds. The timestamp is passed to the XR runtime together with the rendered swapchain images (e.g. as part of the xrEndFrame call in OpenXR). A receiver may use the XR timestamp together with the RTP timestamp to determine the playout time of the media. XR timestamp shall not be used for media synchronization.

NOTE 1: It is left to the discretion of the receiver application how to use the XR timestamp. It is not specified how the receiver application determines the playout time using the XR timestamp together with the RTP timestamp. The receiver application may take both the media transport aspects and XR application aspects (e.g., reducing motion judder) into account while determining the playout time.

**action\_id (32 bits)**: A list of actions corresponding to the pose x, y, z, rx, ry, rz, rw coordinates. An action\_id uniquely identifies an action and it may be an action identifier as defined in the action format of TS 26.119 [17] clause 6.2.3. The number of action identifiers in one RTP HE for XR pose shall be no more than 10. Hence, the size of the RTP HE is 36+2\*n, if a 6DoF XR pose is used, or 24+2\*n, if a 3DoF XR pose is used, where n is the number of action identifiers in the HE.

If the RTP HE for XR pose is for rendered pose, the RTP sender should contain an action\_id field as defined above, with the list of action identifiers identifying the processed actions for the rendering of the frame.

If the RTP HE for XR pose is for pose to be rendered, the RTP sender should contain an action\_id field as defined above, with the list of action identifiers identifying the action for which the pose coordinates apply.

NOTE 2: An XR server should be aware of an XR client’s actions configuration in an action space. Signalling aspects for the actions configuration are defined in other specifications such as TS 26.119 and TS 26.565.

NOTE 3: An XR server should be aware of the XR space used by the XR client for the XR pose fields defined above. Signalling aspects for this XR space are defined in other specifications such as TS 26.119 and TS 26.565.

NOTE 4: When a receiver receives an RTP HE for XR pose, it is up to the application to determine whether the HE is for rendered pose or for pose to be rendered.

When both video and audio are delivered to an RTP receiver, or when either audio or video is delivered using multiple real-time streams (e.g., left eye + right eye), multiple RTP streams may be associated with the same RTP HE data, e.g., the same pose may have been used for generating multiple streams. This may lead to sending the same RTP HE data multiple times in different streams.

A sender may reuse the XR pose RTP HE of one stream for multiple RTP streams. For example, only the video stream carries the pose RTP HE, but the pose is applicable also for the audio bitstream. In this case, the sender shall include the extension attribute media followed by a space separated list of media ID (MID) values in the "a=extmap" attribute. The MID values indicate all media streams for which the pose RTP HE is applicable to. If the extension attribute media is present, then the media description of all bitstreams that reuse the RTP HE shall include the attribute "mid", as defined in RFC 5888 [19].

NOTE 5: In case there is a mismatch between the frame rates of the streams, the receiver may use the few most recent samples from the source RTP stream to obtain a synchronized sample in the dependent stream via interpolation. Alternatively, the receiver may choose to not perform any interpolation and simply use the last available sample from the source RTP stream for the dependent stream. It is left to the discretion of the receiver application to select an appropriate synchronization method.

\* \* \* \* End of 1st change \* \* \* \*