**3GPP TSG SA WG4 Meeting #126S4-231705**

**Chicago, US, 13 – 17 November 2023 Revision of S4aR230101**

|  |
| --- |
| *CR-Form-v12.2* |
| **CHANGE REQUEST** |
|  |
|  | **26.522** | **CR** |  | **rev** |  | **Current version:** | **0.1.1** |  |
|  |
| *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:***  | [5G\_RTP] RTP header extension for pose |
|  |  |
| ***Source to WG:*** | Nokia Corporation, Lenovo |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | 5G\_RTP |  | ***Date:*** | 7 November 2023 |
|  |  |  |  |  |
| ***Category:*** |  |  | ***Release:*** | Rel-18 |
|  | *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:*** | This pCR implements the agreed changes based on the comments from the RTC SWG telcos after SA4#125.This pCR also proposes to enable the usage of the header extension for uplink user pose, in addition to rendered pose sent by a server. |
|  |  |
| ***Summary of change:*** | * Header extension was made generic for the user pose and not restricted to rendered pose.
* Units and format for the extension fields were added.
* Timestamp definition was modified to use the XR system clock.

Changes relative to S4-231705 are highlighted.Changes relative S4-231705-r02 |
|  |  |
| ***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:*** |  |

|  |
| --- |
| **First Change** |

# 2 References

[x] 3GPPP TS 26.119: “Media Capabilities for Augmented Reality”.

|  |
| --- |
| **Second Change** |

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

**Data Burst:** A data burst is a set of multiple PDUs generated and sent by the application such that there is an idle period between two data bursts. A Data Burst can be composed of one or multiple PDU Sets.

**PDU Set:** One or more PDUs carrying the payload of one unit of information generated at the application level (e.g. frame(s), video slice(s), metadata, etc.).

**XR Pose:** A position and orientation in space relative to an XR Space.

**XR Space:** A frame of reference in which an application chooses to track the real world. An XR Space provides a relation of the user’s physical environment with other tracked entities.

**Third Change**

## 4.4 RTP Header Extensions

### 4.4.1 General

TBA

### 4.4.2 RTP Header Extension for PDU Set Marking

### 4.4.3 RTP Header Extension for Pose

An RTP sender that uses RTP to deliver pre-rendered video streams to a UE should include an RTP header extension for pose to indicate the XR pose used for rendering the media (rendered pose). The RTP header extension for pose may also be used with audio streams.

The RTP header extension for pose may also be used by a UE to indicate the XR pose to another UE or to a server.

An RTP client that supports the RTP header extension for pose shall negotiate the use the of the extension using SDP. The signaling of the RTP header extension for pose shall follow the SDP signaling design, the syntax, and semantics of the "extmap" attribute as outlined in RFC8285. The header extension shall be registered with IANA.

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

The ABNF syntax for this header extension extends the "extmap" attribute as follows:

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

*extensionattributes* = [”media:” 1\*(SP token)]

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

An RTP client that supports the RTP header extension for 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.

If the RTP header extension for pose is used by a server, the server should use the RTP header extension for pose to associate the selected pose with the rendered frame.

If negotiated successfully, an RTP sender should add the RTP header extension for pose to the RTP stream. The frequency of RTP header extension for 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 (RFC8285) RTP header extension format shall be used for signalling the RTP header extension for pose as follows:

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=36+2n | rx …

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

 | ry …

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

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

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

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

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

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

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

| timestamp continued …

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

| 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 60559:2020).

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

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

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

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

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

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

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

[Editor’s Note: Definition of the header extension for 3DoF pose is FFS.]

**timestamp (64 bits)**: Timestamp for the pose. If the header extension is used for uplink pose delivery, this timestamp indicates the XR runtime sampling time of the XR pose. If the header extension is used for downlink delivery of the pose used for rendering the frame (rendered pose), this timestamp indicates the predicted XR runtime display time. This timestamp uses the XR system clock and is represented in nanoseconds. There is no requirement to synchronize the timestamps of the RTP stream to the XR system clock. The timestamp is passed to the XR runtime together with the rendered swapchain images (e.g. as part of the xrEndFrame call in OpenXR).

NOTE: For downlink delivery, XR timestamp takes precedence over RTP timestamp for the calculation of the actual display time.

**action\_id (32 bits)**: A list of actions corresponding to the pose x, y, z, rx, ry, rz, rw coordinates. An action\_id uniqely identifies an action and it may be an action identifier as defined in the action format of TS 26.119 Clause 6.2.3. The number of action identifiers in one RTP header extension for pose shall be no more than 10. Hence, the size of the header extension is 36+2\*n, where n is the number of action identifiers in the header extension.

If the RTP header extension for pose is sent by a server, it 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 header extension for pose is sent by a UE, it 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: A peer to a UE XR client should be aware of the UE actions configuration in an action space. Signalling aspects for the UE actions configuration are defined in other specifications such as TS 26.119 and TS 26.565.

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

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 header extension data, e.g., the same pose may have been used for generating multiple streams. This may lead to sending the same header extension data multiple times in different streams.

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

NOTE: 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.

## **Second change**

Annex A (informative):
Examples of SDP offers and answers

A.1 SDP example for pose RTP header extension

An example SDP description using the pose RTP header extension is presented below. Using the extension attribute media, the pose RTP header extension with URI urn:3gpp:xr-pose provided in the video stream with MID m1 is also applicable to another video stream with MID m3.

v=0

o=alice 2890844526 2890844526 IN IP4 host.atlanta.example.com

s=SDP Session

c=IN IP4 host.atlanta.example.com

t=0 0

m=application 1001 UDP/DTLS/SCTP webrtc-datachannel

a=sendonly

m=video 23458 RTP/AVP 96

a=mid:m1

a=recvonly

a=rtpmap:96 H264/90000

**a=extmap:1 urn:3gpp:xr-pose media:m3**

m=audio 23468 RTP/AVP 97

a=mid:m2

a=recvonly

a=rtpmap:97 PCMU/8000

m=video 23478 RTP/AVP 97

a=mid:m3

a=recvonly

a=rtpmap:96 H264/90000