**Source: Huawei**

**Title: AR media types and transport discussion for MeCAR**

## Document for: Discussion and Agreement

## Agenda Item: 9.5

# 1 Introduction

At the RTC SWG Telco #4, contribution [S4aR220018](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/3GPP_SA4_AHOC_MTGs/SA4_RTC/Docs/S4aR220018.zip) was discussed and an agreement was reached (cf. [S4-220924](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/TSGS4_120-e/Docs/S4-220924.zip), “Report for SA4 RTC SWG 3 August 2022 Teleconference”). This contribution attempts to implement the agreed way forward and provides an opportunity for further discussion before reaching a final decision.

# 2 AR/MR metadata definition and transport requirement

For the intended AR/MR applications, it is expected to have a number of media metadata (cf. Clause 2 of [S4aR220018](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/3GPP_SA4_AHOC_MTGs/SA4_RTC/Docs/S4aR220018.zip)) that have different characteristics and levels of dependency on real-time transport. The table below shows how and where some of the AR/MR metadata are defined and an assessment on their levels of dependency on real-time transport.

|  |  |  |  |
| --- | --- | --- | --- |
| **AR/MR Metadata**  **Media Type** | **Media Type Description (Examples)** | **Definition Source** | **Real Time Characteristics** |
| **Scene Description** | **Type**: glTF2.0, JSON format  **Organization**: MPEG-I and Khronos | Clause 4.4.2 of 3GPP TR 26.998[1] | Non timed |
| **Spatial Description** | **Type**: For example:  [https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html#XR\_FB\_spatial\_entity](https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html" \l "XR_FB_spatial_entity)  **Organization:** OpenXR | Clause 4.4.7.3 of 3GPP TR 26.998[1] | Sparsely timed (event based) |
| **3D Model** | **Type**: PoLYgon  **Organization:** None | Clause 4.6.3.5.2 of 3GPP TR 26.928[2] | Non timed |
| **AR Tracking** | **Type**: Metadata allowing accurate overlaying/rendering of text, graphics or video contents to support Use Case 8 of TR 26.928.  **Organization:** None | None | Sparsely timed (event based) |
| **User Pose** | **Type**: It consists of a quaternion for orientation and 3 vectors for position. Timestamp is represented by a 64 bit monotonically increasing nano-second-based integer.  **Organization**: Khronos OpenXR | Clause 4.4.3.1 of 3GPP TR 26.998[1] | Continuous |
| **FOV** | **Type**: It consists of vertical fov and horizontal fov.  **Organization:** None | Y.6.2.3 of 3GPP TS 26.114[3] | Continuous |
| **Viewport** | **Type**: It shall contain all of the parameters Viewport\_azimuth, Viewport\_elevation, Viewport\_tilt, Viewport\_azimuth\_range and Viewport\_elevation\_range  **Organization:** None | Y.7.2 of 3GPP TS 26.114[3] | Continuous |
| **Gesture** | **Type**: A array of finger joint position.  For example: [https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html#XR\_EXT\_hand\_tracking](https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html" \l "XR_EXT_hand_tracking)  **Organization:** OpenXR | None | Continuous |
| **Body action** | **Type**: bvh format.  **Frequency**: at least 1kHz  **Organization:** BioVision company | None | Continuous |
| **Facial expression** | **Type**: An array of key point position.  For example: [https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html#XrSystemFacialTrackingPropertiesHTC](https://www.khronos.org/registry/OpenXR/specs/1.0/html/xrspec.html" \l "XrSystemFacialTrackingPropertiesHTC)  **Organization:** None | OpenXR | Continuous |
| **Sensor information** | **Type**: None  **Organization:** None | None | Non timed |
| **Camera information** | **Type**: The camera parameters such as focal length, principal points, calibration parameters and the pose of the camera all contribute in understanding the relevance between points in the volumetric scene and pixels in the captured image.  **Organization:** None | None | Sparsely timed (event based) |
| **Projection information** | **Type**: Projection mapping information (indicating the projection format in use, e.g., Equirectangular projection (ERP) or Cubemap projection (CMP)), for the projection sample location remapping process as specified in clauses 7.5.1.3 and 5.2 of ISO/IEC 23090-2 [179]  **Organization:** None | Clause Y.3 of 3GPP TS 26.114 and H264 protocol | Sparsely timed (event based) |

# 3 Proposal

We propose to include the Clause 2 of this contribution in the MeCAR PD under a sub-clause 3.x entitled “Media Type Categories and Transport Characteristics”.

# References

1. 3GPP TR 26.998: "Support of 5G Glass-type Augmented Reality / Mixed Reality (AR/MR) devices ".
2. 3GPP TR 26.928: "Extended Reality (XR) in 5G"
3. 3GPP TS 26.114: " IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction".