**3GPP TSG SA WG4 Meeting #127 S4-240427**

**Sophia-Antipolis, FR, 29th January 2024 - 2nd February 2024**

**Title: Work Item Summary of “****Media Capabilities for Augmented Reality (MeCAR)”**

**Type: Work Item Summary**

**Agenda Item: 14.4**

**Source: SA4**

This document provides a summary of the WI Media Capabilities for Augmented Reality.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Unique ID** | **Name** | **Acronym** | **WG** | **WID** | **WI Rapporteur name** |
| **950015** | Media Capabilities for Augmented Reality | MeCAR | S4 | SP-220242 | Emmanuel Thomas, Xiaomi |

# Introduction

The work item MeCAR “Media Capabilities for Augmented Reality” [1] specifies the expected media capabilities for a UE that can be leveraged by AR services and AR applications. The deliverable of this Work Item is the 3GPP TS 26.119 [2] which defines audio, video, scene description and XR system capabilities for four device types representing different form factors of XR devices. The four device types are type 1 “Thin AR glasses”, type 2 “AR glasses”, type 3 “XR phone” and type 4 “XR head-mounted display”. Note that the targeted services and applications belong to the AR domain. However, the device types 3 (XR phone) and 4 (XR head-mounted display) may be capable of running other types of XR applications but those other mode operations are out of scope of TS 26.119.

To relate those media capabilities definition with hardware functions and components, TS 26.119 defines the XR Baseline terminal architecture that is representative of the four device types.

Lastly, TS 26.119 has been developped to serve as a common basis for service and application-oriented AR specifications. In that spirit, TS 26.119 provides data structures and semantics of common metadata (e.g., pose information) and QoE metrics generated by the UE. An example of specifications leveraging TS 26.119 is the TS 26.565 “Split Rendering Media Service Enabler”.

# Description

## Device types

To address a wide range of commercialised devices, TS 26.119 defines four device types targeted by AR services. Those four device types are:

* Device type 1: Thin AR glasses
  + It represents an optical see-trough display device with power-constrained and limited computing power.
* Device type 2: AR glasses
  + It represents s an optical see-trough display device with higher computation power compared to the thin AR glasses device type.
* Device type 3: XR phone
  + It represents a type of device which corresponds to a smartphone with capacities and resources sufficient to offer rich AR experiences.
* Device type 4: XR Head Mounted Display (HMD)
  + It represents a type of device which corresponds to HMDs capable of offering AR experiences but also not precluding other types of XR experiences.

## Media capabilities definition

For each device type, TS 26.119 specifies the mandatory and optional media capabilities to be supported by the UE. Those media capabilities pertain to audio, video, scene processing and XR systems capabilities and are summarised below:

* Video capabilities
  + Single decoder capabilities for AVC and HEVC video codecs from FullHD to 8K resolutions
  + Concurrent decoder capabilities for AVC and HEVC video codecs ranging from 2 to 8 instances
  + Single encoder capabilities for AVC and HEVC video codecs from FullHD to UHD resolutions
* Audio capabilities
  + Single decoder capabilities for EVS, IVAS and AAC-ELDv2 audio codecs
  + Concurrent decoder capabilities for EVS, IVAS and AAC-ELDv2 audio codecs ranging from 2 to 4 instances
  + Single encoder capabilities for EVS, IVAS and AAC-ELDv2 audio codecs
* Scene processing capabilities
  + Process glTF2.0 [4] scene description files and to render the described scenes with the defined restrictions
  + Process glTF2.0 [4] scene description files with additional selected extensions defined in MPEG-I Scene Description [5]

## Metadata

Regarding metadata, TS 26.119 defines common metadata that can be reused for different services and application. Those metadata are:

* A pose information object representing a pose in a 3D space, possibly predicted for a future point in time.
* An action object representing actions performed by a user within an AR application.
* An available visualization space object representing a 3D space within the user’s real-word space that is suitable for rendering virtual objects.

## QoE metrics

To assess the quality of experience of an AR application, TS 26.119 defines a set of QoE metrics which can be reported by the UE. Those metrics are:

* Time duration from pose to render to photon
* Time duration from render to photon
* Time duration for roundtrip interaction delay
* Time duration for user interaction delay
* Time of content age
* Time duration for scene update

# **References**

[1] 3GPP SP-220242, Work Item on “Media Capabilities for Augmented Reality”.

[2] 3GPP TS 26.119: “Media Capabilities for Augmented Reality applications/services/experiences[TBD]”.

[3] 3GPP TS 26.565: “Split Rendering Media Service Enabler”.

[4] ISO/IEC 12113:2022 Information technology Runtime 3D asset delivery format Khronos glTF™2.0.

[5] ISO/IEC 23090-14:2023 Information technology Coded representation of immersive media Part 14: Scene description.