3GPP TSG|WG-4 Meeting #117 S4-220127

Electronic Meeting, 14th – 23rd February 2022 (revision of S4-211568)

**Source: Xiaomi, Qualcomm Incorporated, Samsung Electronics Co., Ltd, AT&T, MediaTek, Dolby Laboratories Inc.**

**Title: Draft WID on Media Capabilities for Augmented Reality**

**Document for: Approval**

**Agenda Item: 10.10**

3GPP™ Work Item Description

Information on Work Items can be found at <http://www.3gpp.org/Work-Items>
See also the [3GPP Working Procedures](http://www.3gpp.org/specifications-groups/working-procedures), article 39 and the TSG Working Methods in [3GPP TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm)

Title: Media Capabilities for Augmented Reality

Acronym: MeCAR

Unique identifier:

Potential target Release: Rel-18

# 1 Impacts

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Affects: | UICC apps | ME | AN | CN | Others (specify) |
| Yes |  | X |  |  |  |
| No | X |  | X | X |  |
| Don't know |  |  |  |  | X |

# 2 Classification of the Work Item and linked work items

## 2.1 Primary classification

### This work item is a …

|  |  |
| --- | --- |
| X | Feature |
|  | Building Block |
|  | Work Task |
|  | Study Item |

## 2.2 Parent Work Item

For a brand-new topic, use “N/A” in the table below. Otherwise indicate the parent Work Item.

|  |
| --- |
| Parent Work / Study Items  |
| Acronym | Working Group | Unique ID | Title (as in 3GPP Work Plan) |
| N/A | N/A | N/A | N/A |

### 2.3 Other related Work Items and dependencies

|  |
| --- |
| Other related Work /Study Items (if any) |
| Unique ID | Title | Nature of relationship |
| 810006 | Extended Reality (XR) in 5G | Initial study on AR/MR and key use cases. |
| 880011 | Study on 5G Glass-type AR/MR Devices | Study on the support of AR/MR with 5G glass-type devices. TR 26.998 concludes 5G Real-time Communication as an area for potential standardisation. |
| [770024](https://www.3gpp.org/DynaReport/WiVsSpec--770024.htm%22%20%5Ct%20%22_blank) | EVS Codec Extension for Immersive Voice and Audio Services | Codec for spatial audio in conversational services |

# 3 Justification

In TR 26.928 and TR 26.998, XR and AR device architectures have been developed and details on relevant media formats are documented, for example in TR 26.998, clause 4.4. In particular, it is identified in TR 26.998 that for AR design glasses (referred to as WLAR and EDGAR in Table 4.2.2.1-1), implementation and operational requirements are significantly more stringent than for smart phones. As an example, consuming media on AR glasses requires functionalities to address very low power consumption, restricted System-on-Chip (SoC) area size, low latency options, new formats, operation of multiple decoders in parallel, etc.

To support basic interoperability for AR applications in context of 5G System based delivery, a set of well-defined media capabilities are essential to create the conditions of a successful ecosystem. These capabilities may be used in different services and applications, therefore it is relevant to define them in a service-independent manner. The media capabilities typically address three main scenarios:

* Support of basic media applications on such AR glasses with simple rendering functionalities, also applicable on other device types
* Support of split-rendering, for which a pre-rendering of the views is carried out in the cloud/edge
* Support of sensor and device data streaming to the network in order to support network-based processing of device sensor information

Media capabilities are relevant for the Media Access function as defined in clause 4.2.5 of TR 26.998. They are importantly driven by realistic deployment options addressing device capabilities as documented in TR 26.998, clause 4.5.2 as well as the relevant KPIs.

In particular, media capabilities cover the aspects of AR Runtime, Lightweight Scene Manager as well as the Basic Codecs defined in the Media Access Function as shown below in Figure 1 for EDGAR type of devices in clause 4.2.2.3 of TR 26.998.

Concerning audio processing, 3GPP SA4 is working on the development of the EVS Codec Extension for Immersive Voice and Audio Services (IVAS) codec. It targets encoding, decoding and rendering of speech, music and generic sound, with low latency operation and support of high error robustness under various transmission conditions. The IVAS codec is expected to provide support for a range of service capabilities, e.g., from mono to stereo to fully immersive audio, implementable on a wide range of UEs.

Other aspects such as the Content Delivery, the Description Delivery and the Media Session Handler of the Media Access Function do not fall under the media capabilities umbrella but work in synergy with. Those delivery and network aspects of the terminal architecture may be addressed by other Work Items (e.g. AREA).



Media capabilities

Figure 1 - Functional structure for Type 2: 5G EDGe-Dependent AR (EDGAR) UE

Lastly, the media capabilities may be referenced and added to 3GPP Media service enablers. In particular, inclusion into 5G Media Streaming is foreseen.

# 4 Objective

This work item defines service-independent media capabilities for AR devices. In particular, the following objectives are considered:

* Define at least one AR device category that addresses the constraints of an EDGAR-type AR glass
	+ Note: Additional device categories may be defined, but with lower priority.
* For each AR device category
	+ Define a reference terminal architecture regarding media capability aspects for this AR device category
	+ Define media types and formats produced and consumed by the AR device, including basic scene descriptions, audio, graphics and video as well as sensor information and metadata about user and environment.
	+ Define the integration of the relevant codecs, including IVAS, into the reference terminal architecture
	+ Define decoding capabilities, including support for multiple parallel decoders
	+ Define encoding capabilities

Define security aspects related to the media capabilities

Define the required, recommended and optional media capabilities for this AR device category

* Define capability exchange mechanisms based on complexity of AR media and capability of device to support EAS KPIs for provisioning of edge/cloud resources
	+ Note: Identify a suitable existing capability framework, or if it does not exist, we need to work with the broader industry (e.g., IETF, KHRONOS, W3C, etc.) to get this done.
* Define relevant KPIs and QoE Metrics for AR media
* Define encapsulations into RTP, ISOBMFF and CMAF
* Specify the relevant parameters for session setup and negotiation of the media delivery and provide instantiations for SDP and DASH MPD
* Enable AR media in 5G Media Streaming by defining suitable 5GMS profiles based on AR media capabilities
* Define typical traffic characteristics for AR media

# 5 Expected Output and Time scale

|  |
| --- |
| New specifications |
| Type  | TS/TR number | Title | For info at TSG#  | For approval at TSG# | Rapporteur |
| TS | 26.XXX (suggest 26.119) | Media Capabilities for Augmented Reality | SA#99 (March 2023) | SA#100 (June 2023) | Gilles Teniou (teniou@tencent.com) |

|  |
| --- |
| Impacted existing TS/TR |
| TS/TR No. | Description of change  | Target completion plenary# | Remarks |
| TS 26.511 | Addition of AR Media Capabilities for 5G Media Streaming | SA#100 (June 2023) |  |
| TS 26.117 | AR Audio Capabilities |  |  |
| TR 26.925  | Typical traffic characteristics for AR media |  |  |

# 6 Work item Rapporteur(s)

Emmanuel Thomas, Xiaomi, thomase@xiaomi.com

# 7 Work item leadership

SA4

# 8 Aspects that involve other WGs

None

# 9 Supporting Individual Members

|  |
| --- |
| Supporting IM name |
| Xiaomi |
| Qualcomm Incorporated |
| Samsung Electronics Co., Ltd |
| AT&T |
| MediaTek |
| Dolby Laboratories Inc. |
| [Facebook] |
|  |
| [Tencent] |