**3GPP TSG SA WG4#116e S4-211552**

**E-meeting, 10th – 19th November 2021**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *CR-Form-v12.1* | | | | | | | | |
| **PSEUDO CHANGE REQUEST** | | | | | | | | |
|  | | | | | | | | |
|  | 26.998 | **CR** |  | **rev** |  | **Current version:** | 1.0.3 |  |
|  | | | | | | | | |
| *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:*** | [FS\_5GSTAR] Proposed Updated Conclusions | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Source to WG:*** | Qualcomm Incorporated, Samsung, Orange | | | | | | | | | |
| ***Source to TSG:*** |  | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Work item code:*** | FS\_5GSTAR | | | | |  | ***Date:*** | | | 2021-11-02 |
|  |  | | | |  | |  | | |  |
| ***Category:*** | C |  | | | | | ***Release:*** | | | Rel-17 |
|  | *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 can be 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-15 (Release 15) Rel-16 (Release 16) Rel-17 (Release 17) Rel-18 (Release 18)* | |
|  |  | | | | | | | | | |
| ***Reason for change:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Summary of change:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Consequences if not approved:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Clauses affected:*** | | 4.5.2, 4.5.3 | | | | | | | | |
|  | |  | | | | | | | | |
|  | | **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:*** | |  | | | | | | | | |

**===== CHANGE =====**

8.1 General

This clause documents and clusters potential standardisation areas identified in the context of this Technical Report.

**===== CHANGE =====**

8.2 5G Augmented Reality Experiences Architectures (5G-AREA)

Based on the initial conclusions in TR 26.928 [2], clause 7, and the evaluation of architectures in clause 4 and 6 of this report, it is clear that for AR experiences, additional architectural components need to be added to media workflows. This includes both, network as well as device architectures. The architectures for different scenarios described in clause 6 of this report need to be defined. In particular, the following architectural enhancements need to be considered:

* Extensions to device architectures to add rendering and AR run time
* Network-architectures to support split rendering and spatial computing
* Operator and third-party services need to be supported
* 5G Integration through different methods (OTT-based and IMS-based)

Based on the above, it is considered to be important to specify 5G Augmented Reality Experiences Architectures addressing the following stage-2 work objectives:

* A generic AR/MR architecture to define relevant core building blocks, interfaces as well as rendering-centric end-points
* Provide extensions to existing 5G system architecture including 5G Media Streaming Architecture or MTSI to address AR experiences
* Call flows and procedures for AR/MR experiences based on the context of clause 6
* Specify a split-rendering and spatial computing architecture for AR devices on top of a 5G System
* Provide all relevant reference points and interfaces to support different collaboration models between 5G System operator and third-party AR application provider

**===== CHANGE =====**

8.3 5G-Media Service Enablers

AR applications rely on functionalities provided by devices and networks. On devices, such functionalities are typically bundled in software development kits (SDKs) in order to get access to complex hardware functionalities. SDKs typically expose APIs to simplify the communication with the underlying hardware and network functionalities.

What is clearly needed for AR and provided for example by Khronos with OpenXR, are standardized APIs to access underlying AR hardware functions. However, the standardized APIs and functions in OpenXR are restricted to local device processing. In order to enable and simplify the access to 5G network, system and media functionalities for AR, it is beneficial to provide packages and bundles for application providers. Typical assets for Media Service enablers are:

* Set of functions that may be used to develop applications on top of 5G Systems
* Set of robust features and functionalities which reduce the complexity of developing applications
* Functions to leverage system and radio optimizations as well as features defined in 5G System (5G Core Network and 5G NR)
* Provision and documentation of APIs to enable or at least simplify access to these functionalities
* Provision of network interfaces to connect to the 5G System
* A testable set of functions. Testing and conformance may be addressed outside 3GPP by an appropriate Marketing and Public Relations (MPR) or Industry Forum.
* Guidelines and examples to make use of the functionalities

It is proposed to use the concept of 5G-Media Service enablers to define relevant specifications for AR and possibly other applications. A common set of properties and functionalities for Media Service Enabler specifications is needed and hence it is proposed to provide a 3GPP internal report that:

* Define the principal properties of Media Service Enablers
* Define minimum and typical functionalities of Media Service Enablers
* Define a specification template for Media Service Enablers
* Identify possibly relevant stage-2 and stage-3 work for Media Service Enablers
* Collect a set of initially relevant Media Service Enablers for normative work

**===== CHANGE =====**

8.4 5G Real-time Communication

As documented in clause 4.2.5 and further developed in the context of clause 6, there are several use cases that require a 5G Real-time communication. The use cases include:

1. EDGAR-based UEs relying on rendering on the network. In this case, the downlink requires sending pre-rendered viewports with lowest latency, typically in the range below 50ms.
2. Uplink streaming of camera and sensor information for cognitive/spatial computing experiences, in case the environment tracking data and sensor data is used in creating and rendering the scene.
3. Conversational AR services require real-time communication both in the downlink and the uplink, even independent from MTSI for app integration of the communication.

In order to provide adequate QoS as well as possible optimizations when using a 5G System for media delivery, an integration of real-time communication into the 5G System framework is essential.

As identified in clause 4.2.5 and clause 6.5, there is a need for supporting third-party applications in 5G real-time communication as well as server-based real-time streaming. From an app developer perspective, an enabler is preferable, especially to support real-time streaming, for example split-rendering.

Different options may be considered, for example re-use of parts of MTSI such as the IMS data channel and 5G Media Streaming for managed services, or re-use of WebRTC for OTT services. A 5G Real-time communication is expected to be aligned with either IMS or WebRTC but provides additional functions to integrate with the 5G System.

It is proposed to define a general 5G Real-time Communication Media Service Enabler that includes, among others, the following functionalities:

* A protocol stack and content delivery protocol for real-time communication based on RTP
* A set of codecs for different media types
* A common session and connection establishment framework, with instantiations based on SIP and SDP for IMS or SDP and ICE for WebRTC, including further possible investigation of control plane
* A capability exchange mechanism
* A security framework, for example based on SRTP and DTLS for WebRTC
* Uplink and downlink communication
* Suitable control protocols for end-to-end adaptation
* QoS and 5G System integration framework
* Reporting and QoE framework

**===== CHANGE =====**

8.5 Media Capabilities for Augmented Reality Glasses (MeCAR)

In TR 26.928 [2] and this report, XR and AR device architectures have been developed and details on relevant media formats are documented, for example in, clause 4.4. In particular, it is identified that for design AR glasses, implementation and operational requirements are significantly more stringent than for smart phones (see clause 4.5.2 and clause 7). As an example, consuming media on AR glasses requires functionalities to address very low power consumption, low 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. These capabilities may be used in different services and applications and hence service-independent capabilities are relevant. The media capabilities typically address three main scenarios:

* Support of basic media services on such glasses with simple rendering functionalities
* Support of split-rendering, e.g. a pre-rendering of eye buffers is carried out in the cloud/edge
* Support of sensor and device data streaming to the network in order to support network-based processing or device sensor information

Media functions are relevant for the Media Access function as defined in clause 4.2.5. The media capabilities are importantly driven by realistic deployment options addressing device capabilities, as documented in clause 4.5.2, as well as the relevant KPIs.

In particular, the following objectives need to be considered:

* Define a reference terminal architecture for AR devices
* 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 media types and formats, including scene description, audio, 3D/2D graphics and video, as well as sensor data.

- Define decoding capabilities, including support for multiple parallel decoders

- Define encoding capabilities

- Define security aspects related to media capabilities

- Define relevant KPIs and QoE Metrics for AR media

- Encapsulation into RTP and ISO BMFF/CMAF

The media capabilities may be referenced and added to 3GPP Media service enablers and/or 3GPP service specifications such as 5G Media Streaming or MTSI.

**===== CHANGE =====**

8.6 Split Rendering Media Service Enabler with AR profile

In the context of this report, it was clearly identified that AR glasses depend on cloud or edge-based pre-rendering. However, not only AR glasses benefit from such a functionality, also for VR, XR and gaming, as identified in TR 26.928 and TR 26.926, would benefit from split rendering approaches. Hence, a basic Media Service Enabler for split rendering is paramount, in particular in combination with 5G new radio and 5G System capabilities.

Based on this discussion it is proposed to specify a generic raster-based Split Rendering Media Service Enabler that includes, among others, the following functionalities:

* A content delivery protocol defined as a profile of 5G-RTC for downlink communications with possible extension
* A relevant subset of codecs for different media types
* A scene description functionality to support a scene manager end point
* Relevant edge compute capabilities, for example Edge procedures, EAS profiles and KPIs for rendering, and rendering context relocation
* Relevant APIs and network communication
* Integration into 5GS and RAN, possibly with support of cross-layer optimizations
* Operational requirements and recommendations for low-latency communications
* Guidelines and examples

In addition to the generic enabler for split rendering a specific profile for AR is recommended to be defined that includes special considerations for:

* The formats to be supported on AR glasses
* The post-processing for pose correction and the integration with XR runtimes
* The power consumption challenge for AR glasses
* The metrics and KPIs for AR glasses
* The required QoS and QoE for AR type of applications as defined in clause 4.5
* Other AR specific considerations

**===== CHANGE =====**

8.7 Smartly Tethering AR Glasses (SmarTAR)

In clause 4.2.2.4, the important aspect of wireless tethering of AR glasses was introduced. The tethering technology between a UE and an AR glass may use different connectivity. Wireless tethered connectivity is provided through WiFi or 5G sidelink. BLE (Bluetooth Low Energy) connectivity may be used for audio. Two main types are identified:

- Functional structure for Type 3a: 5G Split Rendering WireLess Tethered AR UE

- Functional structure for Type 3b: 5G Relay WireLess Tethered AR UE

In the first case, the motion-to-render-to-photon loop runs from the glass to the phone, whereas in the second case the 5G Phone acts as a relay to forward IP packets. The architectures result in different QoS requirements, session handling properties, and also media handling aspects. For enhanced end-to-end QoS and/or QoE, AR glasses may need to provide functions beyond the basic tethering connectivity function, and the resulting AR glasses may be referred to as Smartly Tethering AR Glasses (SmarTAR). Generally, smartly tethering AR glasses is an important aspect. Based on these observations, it is proposed to further study this subject including specific topics such as:

- Defining different tethering architectures for AR Glasses including 5G sidelink and non-5G access

- Documenting end-to-end call flows for session setup and handling

- Identify media handling aspects of different tethering architectures

- Identify end-to-end QoS-handling for different tethering architectures and define supporting mechanisms to compensate for the non-5G link between the UE and the AR glasses

- Provide recommendations for suitable architectures to meet typical AR requirements such as low power consumption, low latency, high bitrates, security and reliability.

- Collaborate with relevant other 3GPP groups on this matter

- Identify potential normative work for stage-2 and stage-3

**===== CHANGE =====**

## [

## 8.8 MTSI-based AR conversational services

As identified in Table 6.1-1, AR conversational services is one of the service scenarios and has a number of related use cases. 3GPP TR 22.873 [14] also addresses use cases relevant to AR conversational services, namely conference calls with AR holography and AR calls, which have similarities with UC#19 and UC#4 in this study, respectively.

As documented in clause 6.5, AR conversational services can be realized using various building blocks, including call setup and control, formats, delivery and 5G system integration, and these building blocks may have different instantiations and/or options. In addition, AR conversational services may support both asymmetrical and symmetrical experiences on various device types, including STAR, EDGAR and WLAR UEs.

In this study, the MTSI architecture is identified as one of the options to map AR conversational services to the 5G system. Furthermore, SA1’s Rel-18 eMMTEL work item introduced new service requirements for 5G IMS Multimedia Telephony Service, including the support of AR media processing in TS 22.261[13] and it is expected that enhancements on the IMS architecture and/or IMS procedures to fulfil new requirements will be handled by SA2 in Rel-18.

It is proposed to define an MTSI-based instantiation for a complete AR communication service, including:

- Terminal architecture(s) considering STAR, EDGAR and WLAR UEs

- Session setup and control procedures for AR media

- Capability negotiation and AR media stream setup procedures

- Transport of AR media and AR metadata via IMS media path including Data Channel

## ]

**===== CHANGE =====**

## [

1. **Other aspects in the FS\_5GSTAR TR (from S4-211539)**

In clause 4.6.1 (related 3GPP work), the ongoing audio work (on IVAS and ATIAS) may be referenced.

In clause 4.6.8 (WebRTC), the description may be amended to split two subclauses: 1) WebRTC as an OTT application, 2) a tentative subset of WebRTC for Media Service Enablers (defined in clause 8). In the latter part, one should clarify that the media handling part of WebRTC (e.g. voice/video engine) is not in scope, and whether signalling servers and STUN/TURN servers for WebRTC are in the operator domain (or not).

## ]

### 4.6.1 3GPP

This clause documents the 3GPP activity related to services using AR/MR device.

- 3GPP TR 26.928 [2] provides an introduction to XR including AR and a mapping to 5G media centric architectures. It also specified the core use cases for XR and device types.

- 3GPP TS 22.261 [13] identified use cases and requirements for 5G systems including AR and 3GPP TR 22.873 [14] is currently developing new scenarios of AR communication for IMS Multimedia Telephony service.

- 3GPP SA4 is working on the documentation of 360-degree video support to MTSI in 3GPP TS 26.114 [15]. It will provide the recommendations of codec configuration and signalling mechanisms for viewport-dependent media delivery.

- In the context of Release-17, 3GPP RAN work [16] is ongoing in order to identify a traffic model for XR application and an evaluation methodology to access XR performance.

- 3GPP SA4 is working on the documentation of 3GPP TS 26.250 [56] to extend the EVS codec for Immersive Voice and Audio Service (IVAS). It targets new features and performances including improved audio quality, low delay, spatial audio coding support, appropriate range of bit rates, high-quality error resiliency, and practical implementation complexity.

- 3GPP SA4 is working on the documentation of a set of test specifications in 3GPP TS 26.131 [57] and TS 26.132 [58] for subjective and objective test methodologies for immersive audio. **===== CHANGE =====**

[56] 3GPP TS 26.250: “Codec for immersive voice and audio services - General overview”

[57] 3GPP TS 26.131: “Terminal acoustic characteristics for telephony; Requirements”

[58] 3GPP TS 26.132: “Speech and video telephony terminal acoustic test specification”

**===== CHANGE =====**

9 Conclusions and Proposed Next Steps

AR/MR experiences involve augmenting visual/auditory contents into the real world to improve the user’s experience with better immersiveness, unlike VR, which provides an entirely virtual world. To realize these experiences, glass-type AR/MR devices may be a good candidate device, easily combining the lights from the real world and those from the display without a need of holding a device in one’s hand.

In this study, the generic finding for eXtended Reality (XR) in TR 26.928 [2] have been further analysed with specific focus on Augmented Reality (AR) experiences and in particular also with a new device type, AR glasses. Different device centric functions of AR glasses are defined, and different device types are defined. Of particular relevance are 5G STandalone AR (STAR) UEs, i.e. devices that have sufficient capabilities to render rich AR experiences on the device as well as 5G EDGe-Dependent AR (EDGAR) UEs for which edge-based rendering support is a must to provide rich AR experiences. Three basic functions are introduced, the AR Runtime, the Scene Manager and the 5G Media Access Function. Basic AR processes are defined, and a comprehensive summary of AR related media formats is provided. The relevant work in external organizations is summarized.

Based on core use cases, different scenarios are mapped to the 5G System architecture, namely (i) Immersive media downlink streaming (ii) Interactive immersive services (iii) 5G cognitive immersive services as well (iv) AR conversational services. Potential normative work is identified and summarized in clause 8.

Based on the details in the report, the following next steps are proposed.

In the short-term:

* Document the relevant 5G Augmented Reality Experiences Architectures (5G-AREA) according to the considerations in clause 8.2. It may leverage the existing 5G System such as 5G media streaming or MTSI.
* Establish the concept of 5G Media Service Enablers as introduced in clause 8.3 and make use of the concept to define relevant AR media service enablers. It also includes identifying the relevant stage-2 and stage-3 works and providing a set of initially relevant functions of Media Service Enablers for normative works.
* Define a 5G Real-Time Communication (5G-RTC) Media Service Enabler to support different low-latency streaming and conversational AR related services based on the considerations in clause 8.4.
* Define Media Capabilities for Augmented Reality Glasses (MeCAR) in a service-independent manner based on the considerations in clause 8.5.
* Based on the work on 5G-AREA, 5G-MSE, 5G-RTC and MeCAR, define a Split Rendering Media Service Enabler for AR.
* Develop the extension of IMS-based AR conversational services, including an extended MTSI terminal architecture in consideration of the device types defined in clause 4.2, as well as session setup and control procedures for AR media and the transport of AR media/metadata via the IMS media path (e.g., Data Channel)

In the mid-term:

* Add issues around semantical perception and spatial mapping to an AI/ML study, taking into account also the findings in TR 22.874.
* Study options for Smartly Tethering AR Glasses (SmarTAR) based on the discussion in clause 8.7.

All work should be carried out in close coordination with other groups in 3GPP on 5G System and radio related matters, edge computing and rendering as well in communication with experts in MPEG on the MPEG-I project as well as with Khronos on their work on OpenXR, glTF and Vulkan/OpenGL. A follow-up workshop based on the information in clause 4.6.9 should be conducted in order to explore additional synergies and complementary work in different organizations in the XR/AR domain.