**3GPP TSG SA WG4#117e S4-22xxxx**

**E-meeting, 14th – 23rd February 2022**

|  |
| --- |
| *CR-Form-v12.1* |
| **PSEUDO CHANGE REQUEST** |
|  |
|  | **26.998** | **CR** |  | **rev** | **1** | **Current version:** | **1.2.0** |  |
|  |
| *For* [***HELP***](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:***  | Audio aspects of TR 26.998 |
|  |  |
| ***Source to WG:*** | EVS SWG |
| ***Source to TSG:*** |  |
|  |  |
| ***Work item code:*** | FS\_5GSTAR |  | ***Date:*** | 2022-02-21 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***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 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-15 (Release 15)Rel-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)* |
|  |  |
| ***Reason for change:*** | Audio aspects were so far not fully considered in the draft TR 26.998 and this needs to be clarified. |
|  |  |
| ***Summary of change:*** | Clarification of the main focus of the TR and the need to confirm audio aspects. Conclusion of further study for audio aspects with definition of some particular items that would require further considerations for audio. |
|  |  |
| ***Consequences if not approved:*** | Potentially misleading presentation of audio aspects, in particular lacking consideration of:- the type of audio capture and playback and the related system integration- suitable differentiation of functional structures for immersive media types- suitable system architecture for split between codec and rendering |
|  |  |
| ***Clauses affected:*** | Introduction, 8.9 (new), 9 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  | **X** |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  | **X** |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  | **X** |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

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

# Introduction

Augmented Reality (AR) and Mixed Reality (MR) promise to provide as new experiences for immersive media services. The form factors of the devices for these services should typically not deviate significantly from those of typical glasses, resulting in less physical space for the various required components such as sensors, circuit boards, antennas, cameras, and batteries, when comparing with typical smartphones. Such physical limitations also reduce the media processing and communication capabilities that may be supported by AR/MR devices, in some cases requiring the devices to offload certain processing functions to a tethered device and/or a server.

This report addresses the integration of such new devices into 5G system networks and identifies potential needs for specifications to support AR glasses and AR/MR experiences in 5G.

The focus of this document is on general system aspects, especially targeting visual rendering on glasses, and may not be equally balanced or equally precise on all media types (e.g. on haptics, GPUs, audio). For example, extrapolations on architectural aspects derived for primarily visual media pipelines to audio pipelines may require confirmation based on further study.

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

## 8.9 Audio Media Pipelines for AR Experiences

The current focus of this document is on general system aspects, especially targeting visual rendering on glasses. As such it may lack accuracy on audio media type. For example, extrapolations on architectural aspects derived for primarily visual media pipelines to audio pipelines need confirmation and further considerations. In particular, the following aspects may require further study:

- In device functional architecture, the type of audio capture and playback and the related system integration need to be defined.

- In 5G AR device types, the functional structures identified in this TR may be differentiated for immersive media types, e.g. operating immersive audio standalone while immersive video functions are split, involving tethered and/or cloud/edge entities.

- In the 5G system architecture mapping, the split of codecs and rendering assumed for video may not be appropriate for audio.

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

# 9 Conclusions

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/spatial computing 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 generic architecture for real-time media delivery based on the 5GMS architecture as addressed in clause 8.2.

- 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.

- Define a 5G real-time communication 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 AR glasses in a service-independent manner based on the considerations in clause 8.5. The outcomes may affect the other items, especially the 5G real-time communication media service enabler and the IMS-based conversational services.

- Based on the work on above, define a split rendering media service enabler to support EDGAR devices, as addressed in clause 8.6.

- Study options for smartly tethering AR glasses based on the discussion in clause 8.7.

- Develop the extension of IMS-based AR conversational services and shared AR experiences, including an extended MTSI terminal architecture, as addressed in clause 8.8.

- Complement this TR with the relevant audio aspects in a follow-up study based on the considerations in clause 8.9.

In the mid-term:

- Add issues around semantical perception and spatial mapping to an AI/ML study, taking into account the findings in clause 4.2.3 and 4.2.5 as well as TR 22.874.

All work topics will benefit to 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 may be conducted in order to explore additional synergies and complementary work in different organizations in the XR/AR domain.