**3GPP TSG-SA4 Meeting # 127-bis-eS4-240952r3**

**Online, 8-12 April 2024**

|  |
| --- |
| *CR-Form-v12.2* |
| **PSEUDO CHANGE REQUEST** |
|  |
|  | **64** | **CR** | pseudo | **rev** | **-** | **Current version:** |  |  |
|  |
| *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 | **X** | Radio Access Network |  | Core Network | **X** |

|  |
| --- |
|  |
| ***Title:***  | [IBACS] pCR for completing missing gaps in TS 26.264 |
|  |  |
| ***Source to WG:*** | KPN N.V., Samsung |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | IBACS |  | ***Date:*** | 14-05-2024 |
|  |  |  |  |  |
| ***Category:*** | **C** |  | ***Release:*** | Rel-18 |
|  | *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-16 (Release 16)Rel-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19)* |
|  |  |
| ***Reason for change:*** | Completion of the IBACS TS (filling placeholders and addressing editor notes). |
|  |  |
| ***Summary of change:*** | 1st and 2nd change: editorial change3rd change: improve text on Immersive AR Media ( including adding speech)4th change: adding reference in AR metadata table to TS 26.5655th change: editorial change6th change: Note to clarify network media rendering7th change: editorial change8th change: Adding RTP reference to TS 26.114 and TS 26.5529th change: Adding QoE reference to TS 26.114 |
|  |  |
| ***Consequences if not approved:*** | TS will be incomplete |
|  |  |
| ***Clauses affected:*** | 1,2,5.2,6.2,6.4,6.5,7,8,9 |
|  |  |
|  | **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:*** |  |

FIRST CHANGE

# 1 Scope

The present document focuses on IMS-based conversational AR (Augmented reality) services. AR services can overlay media (e.g., video, audio, text, etc.) on top of the user’s real perception. Conversational AR services as described by the present document typically include a bidirectional conversational A/V connection in addition to other non-real-time AR media for collaboration or communication between two or more users.

End of FIRST CHANGE

Second CHANGE

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".

[3] 3GPP TS 26.119: "Media Capabilities for Augmented Reality".

[4] 3GPP TS 23.228: "IP Multimedia Subsystem (IMS); Stage 2".

[5] 3GPP TS 24.229: "IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3".

[6] 3GPP TS 26.565: "Split Rendering Media Service Enabler".

[7] Khronos glTF 2.0, [glTF™ 2.0 Specification (khronos.org)](https://registry.khronos.org/glTF/specs/2.0/glTF-2.0.html)

[8] ISO/IEC 23090-14 AMD 2, Information technology — Coded representation of immersive media — Part 14: Scene description — Amendment 2: Support for haptics, augmented reality, avatars, Interactivity, MPEG-I audio, and lighting

[9] 3GPP TS 26.522: "5G Real-time Media Transport Protocol Configurations".

End of Second CHANGE

Third CHANGE

# 5 Immersive AR Media

## 5.1 General

An AR-MTSI client supports simultaneous transfer of multiple media components with real-time characteristics. An AR-MTSI client supports the core media components in [2] for a conversational AR scenario including text, image, video and speech (also referred to as audio).

## 5.2 Speech

AR-MTSI client in terminal offering speech communication shall follow clause 5.2.1 in TS 26.114 [2]. In order to support minimum service interoperability, an AR-MTSI client in terminal shall implement the UE codec and media handling requirements as specified in TS 26.114 [2]

## 5.3 Video

AR-MTSI client in terminal offering video communication shall follow clause 5.2.2 in TS 26.114 [2] and may render it as based on AR metadata (Clause 6) and media configuration (Clause 7). In order to support minimum service interoperability, an AR-MTSI client in terminal shall implement the UE codec and media handling requirements as specified in TS 26.114 [2].

Specifically, the AR-MTSI client in terminal may support Overlays and Scene Description-Based Overlays (as described in TS 26.114 [2] in clause Y.6.4 and Y.6.9) to render video elements in parts of the AR environment. This may result into rendering the video stream (or parts of the video stream) in a sub-area of the display device. Further, the UE may negotiate a stream characteristic most suitable for the sub-area and may renegotiate the stream characteristics in case the sub-area changes.

## 5.4 Real-time text

AR-MTSI client in terminal offering real-time text shall follow clause 5.2.3 in TS 26.114 [2] and may render it as defined in the AR metadata (Clause 6).

## 5.5 Still images

An AR-MTSI client in terminal supporting still images shall follow clause 5.2.4 in TS 26.114 [2] and may render it as defined in the AR metadata (Clause 6).

End of Third CHANGE

Fourth CHANGE

## 6.2 Metadata data channel message format

For the carriage of metadata defined in this clause the AR-MTSI clients shall use the data channel. The data channel sub-protocol shall be identified as “3gpp-ar-metadata”, which shall be included in the dcmap attribute of the SDP.

The transmission order for the data channel shall be set to in-order and the transmission reliability shall be set to reliable.

The metadata message format shall be set to text-based and the messages shall be UTF-8 encoded JSON messages.

A data channel message may carry one or more AR metadata messages as defined in Table 6.2-1.

Table 6.2-1 AR Metadata Messages Format

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| messages | Array(Message) | 1..n | A list of AR metadata messages. Each message shall be formatted according to the Message data type as defined in Table 6.2-2 |

Each metadata message shall follow the format specified in Table 6.2-2.

Table 6.2-2 AR Metadata Message Data Type

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| id | string | 1..1 | A unique identifier of the message in the scope of the data channel session. |
| type | string | 1..1 | A urn that identifies the message type.  |
| payload | object | 1..1 | The message payload depends on the message type. |
| sendingAtTime  | number | 0..1 | The wall clock time when the AR metadata message is transmitted. (clause 9.3.2.1 in TS 26.565 [6]) |

End of fourth CHANGE

fifth CHANGE

## 6.4 Scene descriptions

An AR-MTSI client in terminal shall comply with the capabilities requirements for scene description as described in clause 10 of TS26.119 for their respective device type.

A scene description of an AR session is sent from the AR MF/MRF to the AR-MTSI clients in terminal. An MF that supports exchange of scene description shall support the following:

- glTF 2.0 scenes as specified in [7].

- MPEG\_media extension as defined in [8] to refer to RTP media streams. The external media shall be RTP media streams supported by an AR-MTSI client and signalled in the SDP.

- MPEG\_anchor\_extension as defined in [8] to anchor objects in the real-world.

The scene description shall be sent by the AR MF/MRF to the AR-MTSI client in terminal over the data channel. The type of the message shall be set to “**urn:3gpp:ar:v1:sd**”.

An AR MF/MRF that supports scene descriptions should create and distribute the scene for an AR call with audio and video streams based on the visualization space, viewer position and AR media properties. The AR MF/MRF should create the scene description for each participant (AR-MTSI client in terminal) such that the shared experience is symmetrical for the different users in the call, e.g., to maintain relative position of users and objects.

Editor’s Note: To be aligned with Mecar.

End of Fifth CHANGE

sixth CHANGE

6.5 Network media rendering

### 6.5.1 General

The AR-MTSI client in terminal supporting network media rendering (of AR Media objects or 3D scenes) shall support metadata formats for split rendering as specified in clause 8.2.2 of TS 26.565 [6].

Note: In case network media rendering is used, AR Media objects or 3D scenes are rendered in a split-rendering server (AR AS) based on the XR pose information and actions of the AR-MTSI client in terminal. The split-rendering server renders a view of the scene (or object) based on the viewing angles and XR pose, provided by the AR-MTSI client in terminal. The resulting images are send as video streams from the split-rendering server to the AR-MTSI client in terminal, which renders the videos at their respective position in AR.

### 6.5.2 Pose Format

When the network media rendering is activated, the AR-MTSI client in terminal periodically transmits a set of pose predictions to the AR AS. The pose prediction format shall conform to the payload of the message whose type is "**urn:3gpp:split-rendering:v1:pose**" as specified in clause 8.3.2.2 of TS 26.565 [6].

### 6.5.3 Action Format

The action sets and actions are negotiated during the AR media rendering negotiation. The AR-MTSI client in terminal reports any changes to action state as it occurs by sending updated actions to the AR AS after the network media rendering is activated. When the AR-MTSI client in terminal sends updated actions to the AR AS, the action format shall conform to the payload of the message whose type is "**urn:3gpp:split-rendering:v1:action**" as specified in clause 8.3.2.3 of TS 26.565 [6].

End of sixth CHANGE

Seventh CHANGE

# 7 Media configurations

7.1 General

The media configuration requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 6, also apply for AR-MTSI client in terminal.

An SDP framework for AR data exchange for AR communication is presented to negotiate codec support for AR media, AR metadata, as well as RTP/RTCP signalling necessary for AR media rendering processing.

AR-MTSI client in terminal shall use RTP for the real-time transport of AR media for AR communication. Any AR media as an overlay may refer to the overlay configuration described in clause Y 6.4.3 of TS 26.114 [2]*.*

AR-MTSI client in terminal shall use the data channels for exchange of AR metadata and rendering negotiation. The SDP attribute *3gpp\_armetadata\_types* should be used to indicate the types of AR metadata which defined in clause 6 (e.g. pose, action and scene description) within the data channel.

The syntax for the SDP attribute is:

a=3gpp\_armetadata\_types: <metadata-1> / … / <metadata-N>

The poses as part of AR metadata may be transmitted via RTP session as a RTP header extension which specified in clause 4.3 of TS 26.522 [9].

7.2 Network media rendering configuration

As specified in Annex AC.9 of TS 23.228 [4], the AR application server can provides network assisted rendering. An AR-MTSI client in terminal can decide to request network media rendering based on user selection and its status such as power, signal, computing power, internal storage, etc. The AR-MTSI client in terminal shall complete an AR media rendering negotiation with the AR AS before it initiates subsequent procedures to activate the network media rendering. The data channel should be established for rendering negotiation with SDP offer/answer between AR-MTSI client in terminal and MF/MRF with the sub-protocol “3gpp-sr-conf”, and continue to be used for rendering re-negotiation until the end of the AR communication.

The AR media rendering negotiation between the AR-MTSI client in terminal and the AR AS shall determine the split-rendering configuration which exchanged via the data channel of rendering negotiation. The split-rendering configuration shall be in JSON format as specified in clause 8.4.2 of TS 26.565 [6].

End of Seventh CHANGE

Eighth CHANGE

# 8 AR Data Transport

## 8.1 General

The data transport requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 7, also apply for AR-MTSI clients in terminals.

## 8.2 RTP transport

Additionally to the requirements specified in TS 26.114 [2], clause 7, the RTP Header Extension for PDU Set Marking (clause 4.2) and XR Pose (clause 4.3) specified in TS 26.522 [9] also apply for AR-MTSI clients in terminals.

## 8.3 RTCP usage

Additionally to the requirements specified in TS 26.114 [2], clause 7, the Transmission of timing information data for QoE measurements specified in TS 26.522 [9], clause 5.2, also applies for AR-MTSI clients in terminals.End of Eighth CHANGE

Ninth CHANGE

# 9 Quality of Experience

## 9.1 General

Quality of Experience (QoE) requirements for MTSI clients in terminals specified in TS 26.114 also apply for terminals to be specified by this specification. Further, extensions to those QoE requirements are for future studies (and expected once extensions are made to the AR media formats).

End of Ninth CHANGE