3GPP TSG SA WG4#110-e meeting S4-201139

19th August – 28th August 2020

**Agenda item:** 10.8

**Source:** China Mobile

**Title:** [FS\_XRTraffic] New media Channels of MTSI-based XR conversational services

**Document for** Agreement

6 System Design Assumptions

6.x New media type of MTSI-based XR conversational services

### **6.x.1 Overview**



Figure 1 new media channels between MTSI UE and XR server (SBC / MRF)

Scene: when we use a MTSI-based conversational service, the computing capability of UE cannot meet the requirements if we need to show the other participant a fully detailed AR model. In this proposal, we put the rendering task of high-precision AR model on the side of XR server (SBC/MRF) to solve the problem above.

To realize the rendering on the side of XR server (SBC/MRF), the following steps are needed:

Firstly, MTSI UE collects the data of the current position and pose (such as the position, rotation and FOV of the current device) and uploads them to the XR server (SBC/MRF), which is used to assist the edge cloud on the XR server (SBC/MRF) to render the corresponding AR model in a specific spatial position, and generate the corresponding AR image after rendering.

Then, the AR image is sent back to MTSI UE by XR server (SBC/MRF). Since the AR image received by MTSI UE is not transparent after encoding and decoding, it is necessary to segment the AR image to obtain the AR foreground image.

 Finally, the AR foreground image is superimposed with the real image collected by MTSI UE camera to present the AR scene.

During the process, several new media types are involved between MTSI UE and XR server (SBC/MRF), including the uplink AR action data media type and the downlink Mask+ media type.

The uplink AR action media channel is used to upload the position and pose data of the MTSI UE to the SBC/MRF, ensuring that the XR server can accurately generate AR images corresponding to the MTSI UE. Table 1 lists the transmission parameters. The corresponding transmission parameters are shown in Table 1.

On the other hand, in order to segment the AR image quickly and keep the edge contour well, the Mask+ data of the image need to be generated at the same time when the AR image is generated on the side of XR server (SBC / MRF). Mask+ refers to the lossless data information which is used to describe the edge of the AR foreground image after specific coding. The data is transmitted to MTSI UE through the new Mask+ media channel. It ensures that the MTSI UE side can segment the AR foreground image quickly and well according to the data, so as to complete the later virtual and real superposition. The specific transmission parameters are shown in Table 1.

Table 1 transmission parameters

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Type** | **Synchronization** | **Delay** | **Packet loss** | **Type of GBR** | **Bit rate** |
| Data of the current position and pose | NA | 50 ms | 10-3 | GBR | 50 kbps |
| Mask+ | Synchronize with video | 150 ms | 10-3 | GBR | 0.5~1.5 Mbps（depending on the complexity of the model) |