**3GPP TSG-SA WG4 Meeting #128S4-241117**

**Jeju, Korea, 20 - 24 May 2024**

**Source: InterDigital Canada**

**Title: [FS\_AVATAR] Updates to MPEG-I Scene Description Mapping**

**Agenda item: 9.8**

**Document for: Agreement**

1. Introduction

A description of an implementation of the avatar reference architecture based on MPEG-I Scene Description (ISO/IEC 23090-14) was presented in S4-240759 during the SA4#127-bis-e meeting and documented in the permanent document of the FS\_AVATAR study. During the discussions on the contribution, there was a request to provide some more details on how this mapping can work. This contribution provides a processing flow example for an Avatar AR Call.

1. Proposed Updates

6.2.1 Processing Flow Example for an Avatar AR Call

This section provides a high-level example of the processing procedures on the UEs participating in an avatar AR call, where the base avatar for the user on UE1 has been loaded by UE2. The procedure is illustrated in Figure X.

Figure X - Processing flow example using MPEG-I Scene Description.

The media function loads a reference avatar. UE1 sends data for avatar generation to the media function. The media function processes the received data and uses them to personalize the reference avatar and generate the base avatar. The media function then stores the generated base avatar in the avatar storage for future access.

UE1 retrieves the base avatar model for the user. Based on the avatar representation format for the avatar model, it may, for instance, train a neural network to perform the extraction of avatar animation data from the captured video of the user. These data are then transmitted to the media function/UE2 using a transport format (e.g., RTP) to drive the animation of the base avatar.

The media function/UE2 initializes a MAF pipeline that supports the required ISO/IEC 23090-14 extensions (e.g., MPEG\_animation\_timing, MPEG\_accessor\_timed, MPEG\_buffer\_circular, and MPEG\_node\_avatar) to allow streaming of timed metadata associated with these extensions. The MAF creates the bindings for the circular buffers that will receive the animation data generated by UE1. The presentation engine on UE2 then reads the animation data from the circular buffers and applies them to the base avatar and finally renders the animated avatar.

1. Proposal

It is proposed to update clause 6.2 in the permanent document of the FS\_AVATAR study with processing flow described in section 2 of this contribution.