**Agenda item:** 9.7

**Source:** Qualcomm Inc.

**Title: [FS\_AVATAR] Scene Management for Avatar Communications**

**Document for** Discussion andAgreement

# Introduction

In this contribution, we describe the procedure by which an Avatar communication session is established and how the scene management for such an immersive session is performed.

# Scene Management

## 2.1 Introduction

An AR call with Avatars may be setup in different ways. In its simplest form, the Avatar call is established between 2 participants. This scenario may be realized in one of the following 3 ways:

1. the 3D avatar animation and rendering is performed at the sender side. The receiver’s pose is fixed and assumed to be facing the sender’s avatar. There is no 3D scene in this scenario and the call is performed in a backwards compatible way to 2D video calls.
2. The 3D avatar animation rendering is performed at the sender side. However, the receiver’s pose is used to adjust the rendering operation, allowing the receiver to walk around the sender’s avatar and see them from different angles. This scenario assumes a simple localized scene description that only has one 3D object representing the sender’s avatar and a camera representing the receiver’s pose. No shared session is required.
3. The 3D avatars of the users are placed in a shared space to allow for more immersive experiences that support interactions between the participants. In this case, a shared scene is required and is provided by a central Scene Manager to both participants.

For multi-participant AR calls that offer a shared experience. A shared 3D scene is required. The Scene Manager is responsible for the creation and update of the scene based on input and actions from all participants in the session. The following section describes the call flow for setting up and managing the shared scene for shared space AR calls.

## 2.2 Scene Management in Calls with Avatars

The following call flow depicts the process for setting up and managing a common 3D scene for an AR call with Avatars from 2 or more participants.



The steps are described as follows:

1. In order to use Avatars in communication and shared experience sessions, the user needs to generate and upload their base avatar model:
   1. The user may use local or cloud-based avatar generation tools and services to create a personalized avatar base model
   2. The user uploads its base avatar model to a central accessible storage server that will offer download of that user’s base avatar model to authorized users.

NOTE: secure handling of based avatar models is expected, but not included in this call flow.

1. User #1 establishes or joins a communication/shared space session with User #2.
2. User #1 offers its 3D avatar for use in the session. This is done by offering a scene update that inserts a node into the scene description. The node contains a description of how user #1’s avatar can be reconstructed and animated by other participants in the session.
3. The Scene Manager adds a new node (or set of nodes) to the 3D scene graph that represents User #1’s avatar and related assets. It locally assigns the ownership of this node(s) to User #1, thus only allowing User #1 to update the status of these nodes. A camera node is also inserted and assigned to User #1. This camera is the one used by User #1 to render the 3D scene of the AR call.
4. The participants in the session receive the scene description or an update of the scene description that contains the users’ avatar descriptions.
5. User #2 downloads user #1’s base avatar model, based on the information in the scene description and in accordance with the granted level of access (i.e. which assets and at which level of detail) during that session. The access may for instance be limited to a predetermined level of detail or to a subset of the digital assets that are stored as part of the base avatar model.
6. User #2 informs User #1 on the available animation functionalities for the base avatar model.
7. Based on this information, User#1 generates animation streams and sends those to User #2
8. User #2 uses the downloaded base avatar model and the animation streams from User #1 to reconstruct and animate the 3D avatar of User #1. The avatar is then rendered as part of the scene.

Exchanging information and scene updates to add and track the user’s avatar can be performed according to the protocol specified in TS 26.264 [1] section 6.4. This protocol may need to be updated to address all needs of a shared space experience with interactivity.

# Proposal

We propose to agree the content of section 2 into the PD under a new section on session management for AR calls with avatars.

# References

[1] 3GPP TS 26.264, IMS-based AR Real-Time Communication