**Title:** [IBACS] Transporting motion data for avatar animation

**Source:** Nokia Corporation1

**Document For:** Agreement

**Agenda item:** 10.6

# Introduction

There are two types of 3D video calls being considered for AR calls in IBACS and IRTCW. The current document considers a bidirectional 3D video call that can be realized using either of the following:

* 3D video of user (e.g., V3C encoded video) that is captured in real-time
* 3D avatar of user (e.g., a glTF representation) that is shared and then animated in real-time.

We assume realistic representation for the avatar use case, as well, i.e., the 3D model of the user is created using real-time capture and then animated using captured motion data.

3D modelling of real-time objects can be done from a series of images of the object captured from different angles. For an AR call, these images may be provided by the user by capturing while moving the camera or by capturing with multiple cameras. The 3D model generation and skinning process is described in S4-221357.

The MRF can assist UEs in creating a humanoid 3D model of themselves that can be animated using captured transformation data that is sent to the UE as motion signals. The following contribution presents a solution for carrying motion signals for animating 3D models that is independent of the method used for capturing and creating the motion signals.

# Possible paths for motion signals

## 2.1 Motion signal transport over RTP

We refer to the call flow in S4-221357 shown below. In step 14, UE1 is sending a stream of images to the MRF for 3D model generation. In step 19, after the 3D model has been created and delivered to the UEs, UE1 is sending motion signals to UE2. In this case, it is possible to send the motion signals to the MRF as a RTP header extension (HE) of the image/video stream shown in step 14. Note that the images may not be required continuously once the 3D model has been generated. The UE can then send the motion signals as the RTP HE and pack only dummy data of minimal length in the RTP packet payload.



The motion signals will consist of a name/number to identify the joint and a motion vector (e.g., translation, rotation). The name/number is the same as the one used in the skeletal rig of the 3D model representation received in step 17. To save bits, only the segments (joint and children) that have transformation can be included in the RTP HE. However, in this case losing a packet can lead to lost information that is not detected. Further study is needed to handle such scenarios.

## 2.2 Motion signal transport over data channel

Motion signals can be transported over data channel when an active RTP stream is not associated with it. For example, in the call flow in S4-221357 shown below, there is no related continuous media stream flowing towards UE2 from MRF. The model delivered in step 16 is likely only delivered once or sporadically updated. Therefore, the motion signals can be delivered over the data channel in this case. A grouping parameter can be used in the SDP to indicate the relationship between the motion signals and the 3D model.



# Motion signals

## 3.1 Motion signals format

A skinned 3D model representation shall consist of a hierarchical representation of joints. The representation can be, e.g., glTF, and shall include for each joint at least a joint:segment mapping and a unique joint name or joint number that will be used as joint ID. If both joint name and number are present, the number shall be used as the joint ID in motion signalling.

A motion signal for a joint may consist of a translation matrix and a joint matrix. For each joint, a sender may signal a rotation. For the root joint, the sender may additionally signal a translation. Both translation and rotation shall be with reference to the initial position of the 3D object as established by SDP or scene description signalling. Applications should follow industry practices for animating a 3D object; the method used for animation is out of the scope of this document. Joint name, number, translation and rotation are defined below:

Joint\_Name : String up to 32 characters

Joint\_Number : 16 bit unsigned number

Translation : Float [3]

Rotation : Float [4]

Scale : Float

Motion signals can use RTP HE or data channel depending on the use case. Note that motion signals should be sent only for joints that have a transformation.

NOTE: The possibility for RTP packets with empty/dummy payload needs to be checked. The preference for using a datachannel or RTP HE for sending the motion signalling is a choice of the application.

# Proposal

The proposal is to include the text in section 2 and 3 to the permanent document of IBACS.