**3GPP TSG-SA WG4 Meeting #129-eS4-241590\_r01**

**Online, 19 – 23 August 2024**

**Source: InterDigital Canada**

**Title: [FS\_AVATAR] pCR on MPEG Avatar Representation Format**

**Agenda item: 9.8**

**Document for: Agreement**

**1. Introduction**

A new standardization effort has recently started in the MPEG WG03 (Systems) workgroup with the aim of defining a new avatar representation format. Initial information on the objectives of this new standard and the envisaged representation format can be found in the publicly available output document number WG03N1316 [1].

This avatar representation format would include a geometrical model and all associated data (e.g., blendshapes, skeleton, normals, textures, maps, metadata, etc.), as well as a streamable format for all the dynamics associated to this model (e.g., animation parameters, tracking information, contextual data, etc.). Interoperability between existing models and formats and the MPEG avatar representation format is also taken into consideration in the development of this new representation format.

**2. Reason for Change**

This document proposes to document a most recent avatar representation format standardization work.

**4. Proposal**

It is proposed to agree the following changes and to integrate them into 3GPP TS 26.813.

\* \* \* Begin Changes \* \* \* \*

6.3.4 MPEG Avatar Representation Format

The MPEG WG03 (Systems) workgroup is working on a new standard for defining a new avatar representation format. The objectives of this new standard and the envisaged representation format can be found in the publicly available output document number WG03N1316 [1].

The scope of the work started in MPEG is to develop an interchange representation format for computer generated avatars and associated containers, and an animation stream format to represent the avatar dynamics and timed-based information. This avatar representation format would include a geometrical model and all associated data (e.g., blendshapes, skeleton, normals, textures, maps, metadata, etc.), as well as a streamable format for all the dynamics associated to this model (e.g., animation parameters, tracking information, contextual data, etc.). Interoperability between existing models and formats and the MPEG avatar representation format is also taken into consideration in the development of this new representation format.

The planned timeline for the first phase of this MPEG standard is shown in Figure 1.

*Figure 1 - Phase 1 timeline for MPEG avatar representation format.*

The identified requirements and their priorities for Phase 1 of the MPEG avatar representation format standardization work are documented in [1] and are shown in the following table, where the priorities are defined as follows:

* Low – should be considered once the higher-level requirements are fulfilled or partially acceptable.
* Medium – should be considered once higher-level requirements partially acceptable.
* High – should be the first to be considered for the technical solution.

Table 1 - Requirements for Phase 1 of the MPEG Avatar Representation Format

|  |  |  |
| --- | --- | --- |
| **Category** | **Description** | **Priority** |
| **Avatar Representation** | A suitable exchange format for conversion between avatar representation formats in market | High |
| Representation format includes description of body, skeletal, etc. | High |
| Representation format does not prevent DRM protection | Medium |
| Mesh-based format for representation and animation | High |
| Integration into scene description | Medium |
| Signal coding format | High |
| **General** | Semantic and signal representation* Semantic - high level features
* Signal - low level features
 | High |
| **Geometry** | Multiple levels of detail | High |
| Reference known 3D models | High |
| **Interaction** | Avatar-avatar, user-avatar, avatar-scene interactions | Low |
| Integrates with existing interaction frameworks | Low |
| **Animation and control** | Facial and body animation | High |
|  | Storage and replay of animation streams | Low |
| Efficient animation through:* External sensors
* Pre-defined actions
 | * High
* Low-Medium
 |
| Animation through* Pre-defined actions
* External sensors
 | * Low-Medium
* High
 |
| **Transport** | Delay sensitive animation streams | High |
| Partial transport of base avatar | High |
| Uniquely identifiable avatar representation and animation streams | High |
| MPEG format as fallback | High |
| Reuse of animation streams for different formats (potentially with loss) | High |
| **Storage** | Base avatar stored in single file | High |
| Partial access | High |
| Independent component compression | High |
| Independent component encryption | High |
| Different levels of detail | High |
| Entry point to describe components and relationships | High |
| Identify format for interoperability | High |
| **Privacy/Security** | Protection of authenticity and association with user | Medium |

The aspects that will be technically addressed by MPEG related to this JSON avatar format include what the JSON avatar format should contain as information related to an avatar and serves as a baseline for the continuation of the work to add more features in the future.

The representation baseline format is expected to contain the following data:

* High-level avatar information:
* Name
* Identifier
* Age
* Gender
* Representation data:
* reference template
* mesh: one or more geometries (LoDs).
* body mappings: a list of string with associated geometries to identify the body parts
* skeletons: one or more skeleton skins
* controllers: one or more animation controllers (inc. facial blend shapes, joints, morph targets)
* Additional information:
* Model mappings: one or more mapping function (reference template to B)
* Animations: a set of predefined animations
* Processing functions
* Avatar gaze
* Parametric textures
* Hair
* Accessories

\* \* \* End of Changes \* \* \* \*

References

[1] WG03N1316, “Procedures and Test Formats for Avatar Representation Formats as part of MPEG-I”, MPEG#147, Sapporo, Japan, July 2024.