ISO/IEC JTC 1/SC 29/WG 03 N1379



**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** Exploration Experiments for MPEG-I Avatar Representation Format

**Status:** Approved

**Date of document:** 2024-11-07

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 3 (with cover page)

**Email of Convenor:** young.L @ samsung . com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 N1379**

**November 2024, Kemer, TK**

|  |  |
| --- | --- |
| **Title** | **Exploration Experiments for MPEG-I Avatar Representation Format Description** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **24442** |

[1 EE1: Compression for Animation Streams 1](#_Toc181872407)

[2 EE2: Integrating on Geometry Data Components for Avatar Data 1](#_Toc181872408)

[3 EE3: Other Animation Sample Formats for Animation Stream 1](#_Toc181872409)

1. **EE1: Compression for Animation Streams**

## **1.1 Introduction**

This Exploration Experiment (EE) aims to evaluate and develop methods for compressing animation streams for avatars, focusing on both facial and body animations. The experiment will explore compression for various animation data types, including blendshapes, facial landmarks, animation controllers, and joint transforms, with the goal of establishing a set of standardized approaches that enable efficient and flexible encoding and decoding of avatar animation data across platforms.

## **1.2 Problem statement**

The following statement on Compression for Animation streams are identified:

* Compression of Facial Animation Streams: Develop and test methods for compressing diverse facial animation data types, including blendshapes, landmarks, and animation controllers, focusing on preserving expression fidelity and temporal consistency. The exploration should address various compression levels (both lossless and lossy) to accommodate different bandwidth and latency requirements for real-time applications.
* Compression of Body Animation Streams: Address compression techniques for body animation streams, with a particular emphasis on joint transforms used for skeletal motion. The goal is to preserve the accuracy and natural flow of body movements while reducing the amount of data. This includes evaluating compression mechanisms for different joint hierarchies and motion patterns and examining trade-offs between compression efficiency (both lossless and lossy) and articulation fidelity.
* Standardization of Avatar Animation Data Formats: Propose a standardized format for compressed avatar animation data that is interoperable across platforms and supports a wide range of animation components. The format should be extensible to accommodate new animation parameters as they emerge, and must support compatibility with existing frameworks, such as Mediapipe, ARKit, OpenXR, etc.
* Evaluation of Compression Techniques on Realism and Responsiveness: Establish metrics and testing methodologies to assess the impact of compression on avatar realism and responsiveness, with an emphasis on user perception. The evaluation process will examine how different compression levels affect the perceived quality of facial expressions and body movement, including potential delays, artifacts, or loss of motion fidelity. Objective and subjective testing methods will be developed to validate the compression approaches.

The EE will focus on developing compression tools for avatar animation streams including but not limited to : blendshapes, joints, controllers, landmarks, etc and support their integration into the new part related to Avatar representation format (Part 38).

## **1.3 Use cases relevant for the EE**

The work of the EE is based on the requirements defined in N00359 and use cases identified in m64008 .

The requirements defined in N00359 are currently organized into five different categories including: Avatar Representation, Coding format, Transport/Synchronisation/Carriage, Storage and Privacy and Security requirements.

The requirements related to the “Avatar Representation” are summarized as follows:

* Interoperability: allow converting to and from other avatar representation.
* 3D Scenes: to allow integration into scene description solution.
* Attributes: requirements applicable to the features supported by the avatar, such as, CGI features as textures, geometries, skeletons, accessories, semantic definitions.

The requirements related to the “Coding Format” are summarized as follows:

* General: requirements applicable to the whole format and its use
* Geometry: requirements applicable to the geometry of the base avatar
* Interaction: requirements that relate to the interaction between Avatars and between Avatars and other objects in the scene

* Animation and control: requirements on the facial and body animation

Transport/Synchronisation/Carriage: requirements related to the transport of the base avatar model and the associated animation streams

* Storage: requirements that pertain to the storage of the base avatar model
* Security and privacy: requirements on the protection of the avatar in different scenarios

The requirements and usecases explicit mention the interoperability concept to allow converting to and from other representation and the representation of all components of an avatar including but not exclusive to geometries, textures, skeletons, accessories and semantics.

## **1.4 Related (WG2) and Extracted (new) Requirements**

## **1.5 Relation to other activities (EE, requirements, etc…)**

## **1.6 Mandates**

## **1.7 Participants**

|  |  |  |  |
| --- | --- | --- | --- |
| Participant | Contact | Email | Type |
|  |  |  |  |
| Xidian Univ. | Anthony Trioux | anthony\_trioux@xidian.edu.cn | L |
| Qualcomm | Imed Bouazizi | bouazizi@qti.qualcomm.com | P |
| Interdigital | Joao Regateiro | joao.regateiro@interdigital.com | P |

(P = proponent, L = leader)

## **1.8 Information about proposed technologies**

The following contributions on Compression for Animation Stream have been submitted:

**Meeting #148**

m70195 [SD] **Avatar facial animation stream format and coding scheme**

## **1.9 Extracted from TuC**

## **1.10 Test cases**

The EE may define test cases for which the evaluation criteria will be analyzed. For instance, a first test case can be with live content while another in the on-demand content.

## **1.11 Evaluation criteria**

List of criteria that will allow to compare the different technical solutions and converge to a unique solutions. Criteria can be objective like memory efficiency, bitrate or subjective flexibility, compatibility with legacy solution, etc..

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Description** | **Evaluation** |
| Crieria #1 | Description | The technical solution should minimize/optimise … |

## **1.12 Timeline**

* 2022-01-17: MPEG #149: Refine evaluation of animation streams for compression, bitstream rate savings and subjective/objective metrics.

## **1.13 References**

[1] WG03 N01316, “Procedures and Test Scenario for MPEG Avatar Representation”, MPEG#147, Sapporo, Japan, July 2024.

[2] WG02 N00359, “Draft of MPEG-I Phase 3 requirements”, MPEG#146, Rennes, France, April 2024.

[3] WG03 N0XXXX, “WD of ISO/IEC 23090-38 Avatar representation formats”, Kemer, Turkey, November 2024.

[4] WG02 m64008 “Use Cases for MPEG Media Avatar”, Hanover, Germany, October 2023.

1. **EE2: Integrating on Geometry Data Components for Avatar Data**

## **1.1 Introduction**

Currently, there are several formats (FBX, USD or glTF) for encoding 3D content including mesh rigs, e.g., deformable 3D objects with coeficient parametrisations. None of these formats are dedicated to avatar data representation. In most cases the definition and parsing of the avatar data structure is given to the application has a task, such as parsing existing formats.

The non existing of a standardized avatar format motivate for the initiation of the current work and the combined effort to produce a format that is interoperable between application and compatible with existing frameworks.

The goal of the EE is thus to specify the integration of avatar data information into a container format that is accessible in the avatar data model. The objective is to improve interoperability, avoid the exclusive used of external limited formats and provide a complete representation of an avatar format in MPEG-I. The EE will study the integration of the data components and based on this integration and the possible identified technical gaps, the EE aims at defining the necessary data components to the data structure of the avatar representation format.

## **1.2 Problem statement**

An interoperable avatar representation format should be capable of allowing other formats to decode, understand, map and utilize its data model structures for individual applications. In the current days an avatar format is a combination of several formats (illustrated in Figure X).

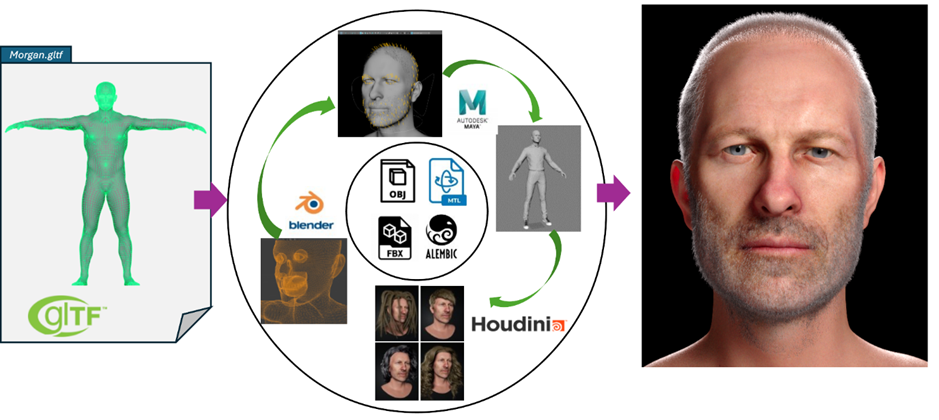


Figure 1 - Ecosystem of avatar asset creation.

This is impractical and not conformed with the claim for an interoperable format.

An interoperable format requires:

* Representation and definition of basic structures e.g., meshes, joints, textures and external file references.

## **1.3 Use cases relevant for the EE**

The work of the EE is based on the requirements defined in N00359 and use cases identified in m64008 .

The requirements defined in N00359 are currently organized into five different categories including: Avatar Representation, Coding format, Transport/Synchronisation/Carriage, Storage and Privacy and Security requirements.

The requirements related to the “Avatar Representation” are summarized as follows:

* Interoperability: allow converting to and from other avatar representation.
* 3D Scenes: to allow integration into scene description solution.
* Attributes: requirements applicable to the features supported by the avatar, such as, CGI features as textures, geometries, skeletons, accessories, semantic definitions.

The requirements related to the “Coding Format” are summarized as follows:

* General: requirements applicable to the whole format and its use
* Geometry: requirements applicable to the geometry of the base avatar
* Interaction: requirements that relate to the interaction between Avatars and between Avatars and other objects in the scene

* Animation and control: requirements on the facial and body animation

Transport/Synchronisation/Carriage: requirements related to the transport of the base avatar model and the associated animation streams

* Storage: requirements that pertain to the storage of the base avatar model
* Security and privacy: requirements on the protection of the avatar in different scenarios

The requirements and usecases explicit mention the interoperability concept to allow converting to and from other representation and the representation of all components of an avatar including but not exclusive to geometries, textures, skeletons, accessories and semantics.

## **1.4 Related (WG2) and Extracted (new) Requirements**

## **1.5 Relation to other activities (EE, requirements, etc…)**

## **1.6 Mandates**

## **1.7 Participants**

|  |  |  |  |
| --- | --- | --- | --- |
| Participant | Contact | Email | Type |
|  |  |  |  |
| Interdigital | Joao Regateiro | joao.regateiro@interdigital.com | L |
| Qualcomm | Imed Bouazizi | bouazizi@qti.qualcomm.com | P |
| Xidian Univ. | Anthony Trioux | anthony\_trioux@xidian.edu.cn | P |

(P = proponent, L = leader)

## **1.8 Information about proposed technologies**

The following contributions on XXXX have been submitted:

**Meeting #148**

m69421 **Avatar JSON Interchangeable Format**

## **1.9 Extracted from TuC**

## **1.10 Test cases**

The EE may define test cases for which the evaluation criteria will be analyzed. For instance, a first test case  can be with live content while another in the on-demand content.

## **1.11 Evaluation criteria**

List of criteria that will allow to compare the different technical solutions and converge to a unique solutions. Criteria can be objective like memory efficiency, bitrate or subjective flexibility, compatibility with legacy solution, etc..

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Description** | **Evaluation** |
| Crieria #1 | Description | The technical solution should minimize/optimise … |

## **1.12 Timeline**

* 2022-01-17: MPEG #149 : Refine data components for representation of avatar data in the avatar data model and document

## **1.13 References**

[1] WG03 N01316, “Procedures and Test Scenario for MPEG Avatar Representation”, MPEG#147, Sapporo, Japan, July 2024.

[2] WG02 N00359, “Draft of MPEG-I Phase 3 requirements”, MPEG#146, Rennes, France, April 2024.

[3] WG03 N0XXXX, “WD of ISO/IEC 23090-38 Avatar representation formats”, Kemer, Turkey, November 2024.

[4] WG02 m64008 “Use Cases for MPEG Media Avatar”, Hanover, Germany, October 2023.

1. **EE3: Other Animation Sample Formats for Animation Stream**

## **1.1 Introduction**

This Exploration Experiment (EE) aims to develop structures for animation streams for avatars, focusing on both facial and body animations. The experiment will explore animation sample to be streamed and transmitted for various animation data types, including blendshapes, facial landmarks, animation controllers, and joint transforms, with the goal of establishing a set of standardized approaches that enable efficient and flexible encoding and decoding of avatar animation data across platforms.

## **1.2 Problem statement**

The following statement for Animation samples are identified:

* Standardization of Avatar Animation Data Formats: Propose a standardized format for avatar animation data that is interoperable across platforms and supports a wide range of animation components. The format should be extensible to accommodate new animation parameters as they emerge, and must support compatibility with existing frameworks, such as Mediapipe, ARKit, OpenXR, etc.

## **1.3 Use cases relevant for the EE**

The work of the EE is based on the requirements defined in N00359 and use cases identified in m64008 .

The requirements defined in N00359 are currently organized into five different categories including: Avatar Representation, Coding format, Transport/Synchronisation/Carriage, Storage and Privacy and Security requirements.

The requirements related to the “Avatar Representation” are summarized as follows:

* Interoperability: allow converting to and from other avatar representation.
* 3D Scenes: to allow integration into scene description solution.
* Attributes: requirements applicable to the features supported by the avatar, such as, CGI features as textures, geometries, skeletons, accessories, semantic definitions.

The requirements related to the “Coding Format” are summarized as follows:

* General: requirements applicable to the whole format and its use
* Geometry: requirements applicable to the geometry of the base avatar
* Interaction: requirements that relate to the interaction between Avatars and between Avatars and other objects in the scene

* Animation and control: requirements on the facial and body animation

Transport/Synchronisation/Carriage: requirements related to the transport of the base avatar model and the associated animation streams

* Storage: requirements that pertain to the storage of the base avatar model
* Security and privacy: requirements on the protection of the avatar in different scenarios

The requirements and usecases explicit mention the interoperability concept to allow converting to and from other representation and the representation of all components of an avatar including but not exclusive to geometries, textures, skeletons, accessories and semantics.

## **1.4 Related (WG2) and Extracted (new) Requirements**

## **1.5 Relation to other activities (EE, requirements, etc…)**

## **1.6 Mandates**

## **1.7 Participants**

|  |  |  |  |
| --- | --- | --- | --- |
| Participant | Contact | Email | Type |
|  |  |  |  |
| Qualcomm | Imed Bouazizi | bouazizi@qti.qualcomm.com | L |
| Xidian Univ. | Anthony Trioux | anthony\_trioux@xidian.edu.cn | P |
| Interdigital | Joao Regateiro | joao.regateiro@interdigital.com | P |

(P = proponent, L = leader)

## **1.8 Information about proposed technologies**

The following contributions on XXXX have been submitted:

**Meeting #148**

**m70326 [SD][Avatar] MPEG Avatar Animation Stream**

**m70195 Avatar facial animation stream format and coding scheme**

Interactivity in Scene Description

## **1.9 Extracted from TuC**

## **1.10 Test cases**

The EE may define test cases for which the evaluation criteria will be analyzed. For instance, a first test case  can be with live content while another in the on-demand content.

## **1.11 Evaluation criteria**

List of criteria that will allow to compare the different technical solutions and converge to a unique solutions. Criteria can be objective like memory efficiency, bitrate or subjective flexibility, compatibility with legacy solution, etc..

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Description** | **Evaluation** |
| Crieria #1 | Description | The technical solution should minimize/optimise … |

## **1.12 Timeline**

* 2022-01-17: MPEG #149: Refine Animation Samples for avatar animation

## **1.13 References**

[1] WG03 N01316, “Procedures and Test Scenario for MPEG Avatar Representation”, MPEG#147, Sapporo, Japan, July 2024.

[2] WG02 N00359, “Draft of MPEG-I Phase 3 requirements”, MPEG#146, Rennes, France, April 2024.

[3] WG03 N0XXXX, “WD of ISO/IEC 23090-38 Avatar representation formats”, Kemer, Turkey, November 2024.

[4] WG02 m64008 “Use Cases for MPEG Media Avatar”, Hanover, Germany, October 2023.