|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligencewww.mpai.community |

|  |
| --- |
| **Public Document** |
| **N274** | 2021/06/09 |
| **Source** | Requirements |
| **Title** | MPAI Application Note #10 - MPAI-MCS – Mixed Reality Collaborative Spaces |
| **Target** | MPAI Members |

**Proponent:** Adam Sobieski (Phoster)

**Description:**

New technologies are emerging which equip developers to deliver mixed-reality collaborative space (MCS) scenarios where biomedical, scientific, and industrial sensor streams and recordings are to be viewed.

A state-of-the-art MCS software development kit is Microsoft Mesh [1]. Related software include Adobe Aero, ApertusVR, Campfire, GatherInVR, Lobaki, STRIVR, Vectory Web AR and Cesium.

Microsoft Mesh “provides a cross-platform developer SDK so that developers can create apps targeting their choice of platform and devices – whether AR, VR, PCs, or phones. Today it supports Unity alongside native C++ and C#, but in the coming months, Mesh will also have support for Unreal, Babylon, and React Native” [1].

Microsoft Mesh “supports most 3D file formats to natively render in Mesh-enabled apps, solving the challenge of bringing in users’ existing 3D models for collaboration” [1]. It appears that new efforts can facilitate viewing 3D live streams and recordings from biomedical, scientific, and industrial sensors and devices in MCS applications (see examples).

**Comments:**

Artificial intelligence can be utilized throughout MCS systems for immersive presence, spatial maps (e.g. Lidar scans, inside-out tracking), rendering, and multiuser synchronization.

Artificial intelligence can be utilized in applications of supervised machine learning for extracting, viewing and recording of biomedical, scientific, and industrial sensor streams in MCS applications. Artificial intelligence can enhance sensor and device data with descriptors.

**Examples:**

1. A science teacher connects a digital microscope to their Mixed-Reality (MR) device – or to another computing device – to stream photorealistic 3D digital content to students in a MCS.

Students may view, for example, a living cell as it is being streamed with descriptors as the teacher explains.

1. The teacher can adjust physical and software controls of the digital microscope while immersed, without having to physically touch the digital microscope which may be at a different location.
2. A science student uses their MR device to browse and interact with a large collection of photorealistic recordings from digital microscopes and other scientific sensors and devices.
3. Medical students study together in a MCS while viewing recordings from biomedical sensors and devices.
4. In an industrial environment, a person involved in training with computer vision algorithms in a MCS for industrial inspection scenarios, viewing data from multiple sensors and algorithms as foods, parts, or products glide on a conveyor belt.

**Requirements:**

1. Capture and stream digital representation of biomedical, scientific, and industrial objects preserving their 3D nature.
2. Extract and stream the required descriptors from the objects.
3. Present the object(s) and their descriptors in a MCS allowing users to interact with, zoom, rotate, and move (virtually or physically) the object and to create and view sections of the object’s interior by intersecting with a plane.
4. Store the objects and descriptors so that a user may later be able to perform the actions with the recorded data.

**Object of standard:**

1. Data formats for
	1. raw sensor data
	2. processed/compressed sensor data (i.e., suitable for streaming and storage)
	3. descriptors of sensor data in a form suitable for streaming and storage, e.g., compressed
	4. two-way data to enable users to remotely control sensor devices
2. Streaming protocols (TBD)
3. Raw data types include, but are not limited to:
	1. imagery, video
	2. light-field imagery, light-field video
	3. RGB-D imagery, RGB-D video
	4. point-cloud imagery, point-cloud video
	5. 3D meshes, and mesh-based animations
	6. volumetric data.

The proposed standard will enable interoperability between (1) biomedical, scientific, and industrial sensors and devices, (2) live streams and recordings from such sensors and devices, and (3) their presentation in MCS systems.

**Benefits:**

With new standards, manufacturers of biomedical, scientific, and industrial sensors and devices and developers of software for their interoperability will have a clear view of how to make their sensors, devices, and systems interoperable with MCS systems.

With new standards, developers of MCS systems will have a clear view of how to make their systems interoperable with live-streaming and recorded sensor data.

**Bottlenecks:** TBD.

**Social aspects:**

Facilitating scientific visualization scenarios in MCS will accelerate scientific progress, industrial operation and advance STEM education.

**Success criteria:**

A success criterion is the utilization of new standards by biomedical, scientific, and industrial sensor and device manufacturers, developers of interoperability software for them, and the developers of MCS systems.

**References:**

[1] <https://techcommunity.microsoft.com/t5/mixed-reality-blog/microsoft-mesh-a-technical-overview/ba-p/2176004>

[2] <https://www.khronos.org/openxr/>

[3] <https://immersiveweb.dev/>

[4] <https://www.khronos.org/gltf/>

[5] <https://github.com/KhronosGroup/glTF/issues/1238>

[6] <https://github.com/KhronosGroup/glTF/issues/1238#issuecomment-736221220>

**Note:**

glTF [4] is a popular format for 3D resources. The topic of streaming 3D animations was recently discussed in a glTF GitHub issue [5]. There, Don McCurdy stated that “one case that is already available is to put different animations into different .bin files, downloading each animation when it is needed. Could be used for breaking an animation into chronological chunks, or lazy-loading animations in a game that aren't needed until the player unlocks them.” He continued, indicating that one would “need something considerably more complex than glTF to have one application reading from a file at the same time as another application is writing arbitrary data into it. That feels more like the role of an interchange format, perhaps. But I imagine someone could define an extension that allows open-ended buffers for animation samplers, allowing animation to stream in without fundamentally changing the structure of the scene.” [6]

Streaming of 3D data is being considered in several environments.