**Agenda item:** 10.7

**Source:** Qualcomm Inc.

**Title: [AvCall-MED] 3D Avatar Profile**

**Document for** Discussion andAgreement

# Introduction

In this contribution, we propose further updates to the 3D avatar profile based on the ARF specification.

# Proposed Changes

### 5.6.2 3D Avatar Format

NOTE: The following is contingent on results of evaluation of ARF in TR 26.813, which is FFS.

[An AR-MTSI client that supports 3D avatars shall support the ARF base avatar format as specified in ISO/IEC 23090-39 [11] with the following specific requirements and constraints:

The 3D avatar container shall consist of the following mandatory components:

* At least one Skeleton component that defines the hierarchical joint structure for body animation, with support for at least partial humanoid joint configurations. Inverse Bind Matrices (IBMs), shall be provided for each joint of the skeleton,
* At least one Skin component that references both the skeleton and associated meshes to enable skeletal deformation through Linear Blend Skinning (LBS),
* 3D mesh geometry data that conforms to the binary glTF (GLB) format version 2.0 or later, with support for:
	+ Vertex positions, normals, and texture coordinates
	+ Triangle-based topology
	+ Multiple levels of detail (LOD) for scalable rendering
* Skinning weight data provided as dense tensors in the format specified in Annex E of ISO/IEC 23090-39 [11], where:
	+ Weights are represented as an NM tensor (N vertices M joints)
	+ Each vertex’s weights sum to 1.0 for proper deformation
	+ Maximum of 4 joint influences per vertex for optimization
* Texture data components that conform to still image formats as defined in section 5.5, including:
	+ Diffuse/albedo textures in Recommended normal maps for surface detail enhancement
	+ Optional metallic-roughness textures for PBR rendering
	+ Optional occlusion maps for ambient lighting
* For facial animation support, 3D avatars shall include:
	+ At least one BlendshapeSet component that references the head/face mesh, containing:
		- A minimum of 50 blend shapes
		- Shape key data as meshes in GLB format, restricted to vertex positions, face information, normals, and tangents. All other information for the blend shape key is inherited from the base mesh.
	+ At least one BlendshapeSet or a mapping table shall correspond to an OpenXR-specified blendshape set.

The ARF document for 3D avatars shall specify:

* Multiple animation types in the supportedAnimations list:
	+ Body animation URNs for skeletal joint animation
	+ Face animation URNs for blend shape animation
	+ Optional hand animation URNs for detailed finger tracking
* A hierarchical asset structure supporting multiple levels of detail:
	+ Each LOD shall maintain consistent UV mapping for texture compatibility
	+ Lower LODs should preserve animation capability while reducing polygon count

The following performance and optimization recommendations apply:

* Mesh data should be optimized with:
	+ Indexed geometry to reduce vertex duplication
	+ Quantized vertex attributes where appropriate
	+ Efficient texture atlas usage to minimize draw calls
* Data items of the 3D avatar should signal no compression or protection schemes by default to ensure broad compatibility.

NOTE: Compression aspects of ARF are for FFS.

NOTE: Content protection aspects are for FFS

* 3D avatars should target a maximum of 200,000 polygons for the highest LOD on mobile devices, with appropriate reductions for lower LODs.]

# Proposal

We propose to accept the proposed updates and integrate them into the base CR.