**Source**[[1]](#footnote-2)**:** InterDigital Communications, Nokia Corporation

**Title: [MeCAR] V3C Streaming Technologies and Considerations for MeCAR**

**Document for** Agreement

**Agenda item:** 9.5

# Introduction

A new section on volumetric video support was added to the MeCAR Permanent document during the SA4#122 meeting, with the Visual Volumetric Video-based Coding (V3C), ISO/IEC 23090-5 [1], standard being one possible codec to provide this support.

A number of use cases documented in TR 26.928 and TR 26.998, from which the MeCAR work item was established, involve streaming volumetric content (e.g., use cases 17 and 18 in TR 26.998). Therefore, within the context of MeCAR, in addition to defining operation points for the volumetric video codec, it is also necessary to define profiles that provide constraints on the packaged V3C media and associated descriptors to address different use case scenarios and device types.

This document presents a brief overview of one of the existing technologies and standards for the delivery of V3C media, namely the ISO/IEC 23090-10 [2] specification, and provides some relevant considerations for MeCAR. Note that in this document mainly focuses on the IF-4 interface, i.e., the interface between the Media Access Function and the 5G System for user plane data, as defined in section 4.1 of the permanent document. The following figure demonstrates the client architecture as documented in the MeCAR permanent document and is repeated here for convenience.



Figure Client architecture from MeCAR permanent document

# Storage and Transport of V3C

The ISO/IEC JTC1/SC29 MPEG Systems workgroup (WG3) has developed a standard to support efficient and interoperable storage and transport of compressed visual volumetric media content with 6 degrees of freedom. The standard defines how to carry V3C bitstreams in an ISOBMFF container and is based on and dereived from the ISO/IEC 14496-12 standard [3] and extends the Dynamic Adaptive Streaming over Hypertext Transfer Protocol (HTTP) (DASH) [4] and MPEG Media Transport (MMT) [5] frameworks by adding additional signaling to enable delivery of V3C-coded content over a network. A comprehensive overview of the generic V3C carriage concepts and related functionalities defined by the specification is presented in [6]. A V3C streaming demo based on the technologies defined in this standard was presented during the SA4#122 meeting and described in S4-230217.

## Encapsulation File Format

The ISO/IEC 23090-10 specification is derived from ISO/IEC 14496-12 and specifies how boxes defined in ISO/IEC 14496-12 should be used for storing V3C-coded content. It also defines new boxes required to store a V3C bitstream in an ISOBMFF container and to support various functionalities, such as recommended viewports and spatial partial access. The specification introduces three methods for storing V3C-coded content in ISOBMFF, where each mode defines a number of sample entries that impose certain constraints on the track(s) supported by that mode and a specific sample format for these tracks in addition to other mode-specific tools. The supported storage modes are: single-track storage, multi-track storage, and non-timed storage.

A single-track encapsulation mode represents the V3C bitstream in ISOBMFF as one track, V3C bitstream track. V3C bitstream track is identified by a sample entry with type ‘v3e1’ or ‘v3eg’. This encapsulation mode is intended for direct ISOBMFF encapsulation without any additional pre-processing or de-multiplexing of the ingested V3C bitstream. While simple, this encapsulation mode does not support partial access and a client is not able to select only a subset of the V3C components for playback.

The multi-track encapsulation mode stores the V3C bitstream in the ISOBMFF file as several tracks, where each track represents either part of or a complete V3C component. This is the preferred mode for streaming applications since a number of independent encoders can run in parallel and the resulting bitstreams can be stored into an ISOBMFF-compliant file or set of files as separate tracks. This provides a flexibility where the extraction and direct processing of each V3C component by their respective decoder becomes much easier without the need to reconstruct the V3C bitstream.

Finally, the non-timed encapsulation mode represents the V3C bitstream in ISOBMFF as items, as defined in [3], where each item represents part of or a complete V3C component. This mode is well suited, for example, to store a V3C bitstream encoded with a still picture profile that has less strict decoding requirements and limited coding tool support.

To signal to a parser which storage mode is used and what type of functionality needs to be supported to play the file, a number of ISOBMFF brands are defined by ISO/IEC 23090-10. A brand might indicate the type of encoding used, how the data of each encoding is stored, constraints and extensions that are applied to the file, the compatibility, or the intended usage of the file.

The single-track encapsulation mode is identified by the brand ‘v3st’. Multi-track encapsulation is identified by the ‘v3mt’ and ‘v3mp’ brands. The ‘v3mt’ brand informs the parser that the file contains V3C content stored using a basic multi-track storage mode, while the brand ‘v3mp’ indicates a multi-track storage mode with additional features present, such as spatial partial access or recommended viewports. The non-timed encapsulation mode is identified by the brand ‘v3nt’.

## Streaming Support

The ISO/IEC 23090-10 specification supports delivering V3C content using MPEG-DASH [4] as well as MMT [5]. In the case of MPEG-DASH, the standard defines how to signal V3C content in the Media Presentation Description (MPD) for both the single-track and multi-track encapsulation modes, including defining V3C-specific DASH descriptors, and defines restrictions on the DASH segments generated for the content.

### DASH Single Track Mode

In this mode, the V3C content is represented with a single Adaptation Set in the MPD with one or more Representations. The only constraint on the Representations of this Adaptation Set is that the codec used for encoding a given V3C video component must be identical across all Representations. There is no requirement, however, that all the V3C video components in one Representation must be encoded using the same codec.

### DASH Multi-track Mode

As with encapsulation, the multi-track mode provides more flexibility over the single-track mode by enabling adaptation across several dimensions as each V3C component is represented by its own Adaptation Set. By separating the V3C video component bitstreams into multiple Adaptation Sets, a streaming client can prioritize or completely drop some components or maps when making adaptation decisions. Moreover, V3C video component representations can be encoded using different video codecs or different bitrates to allow for efficient adaptive bitrate streaming. In addition, the multi-track mode defines number of constrains on the segments of the Adaptation Sets describing V3C components.

# V3C Delivery Considerations for MeCAR

The objectives of the MeCAR work item include the following:

* Specify encapsulations into RTP, ISOBMFF and CMAF
* Specify the relevant codec-level parameters for session setup and negotiation of the media delivery and provide instantiations for SDP and DASH MPD

Moreover, as indicated in section 6.5 of the MeCAR premanent document, a MeCAR device at the minimum shall support basic delivery formats and may also support decoding of streaming formats.

It has been demonstrated through a number of implementaitons that Type 3 and Type 4 MeCAR devices have sufficient processing power to support he decoding and rendering of V3C objects. Therefore, in addition to specifying V3C codec operation points, the MeCAR specification would also need to define V3C media profiles that include the specification of elementary stream constraints based on the identified operation points as well as file format encapsulation and signalling and constraints on the elements present in a delivery manifest file such as the DASH MPD.

# Proposal

We propose to include Sections 2 and 3 in the MeCAR permanent document as the basis for further work and to agree on defining volumetric video media profiles that address the requirements of relevant use cases.

# References

[1] ISO/IEC 23090-5:2021, Information Technology — Coded Representation of Immersive media — Part 5: Visual Volumetric Video-Based Coding (V3C) and Video-Based point Cloud Compression (V-PCC).

[2] ISO/IEC 23090-10:2021, Information Technology — Coded Representation of Immersive media — Part 10: Carriage of Visual Volumetric Video-Based Coding Data.

[3] ISO/IEC 14496-12:2020, Information Technology — Coding of Audio-Visual Objects — Part 12: ISO Base media File Format.

[4] ISO/IEC 23009-1:2019, Information Technology – Dynamic Adaptive Streaming over HTTP (DASH) – Part 1: Media Presentation Description and Segment Formats.

[5] ISO/IEC 23008-1:2017, Information Technology — High Efficiency Coding and media Delivery in Heterogeneous Environments — Part 1: MPEG media Transport (MMT).

[6] Ilola L, Kondrad L, Schwarz S and Hamza A (2022) An Overview of the MPEG Standard for Storage and Transport of Visual Volumetric Video-Based Coding. Front. Sig. Proc. 2:883943. doi: 10.3389/frsip.2022.883943.

1. Contact: Ahmed Hamza, InterDigital Communications, Canada. Email: Ahmed.Hamza@InterDigital.com [↑](#footnote-ref-2)