**Source:** InterDigital Inc.

**Title: Candidate 2D video capabilities for MeCAR**

**Document for** Discussion and Agreement

**Agenda item:** 9.5 - MeCAR (Media Capabilities for Augmented Reality)

# Introduction

The MeCAR work item defines service-independent media capabilities for AR devices. Among the objectives of this work item is “to define media types and formats produced and consumed by the AR device (…), define decoding capabilities, including support for multiple parallel decoders and define encoding capabilities”. One justification is that “media capabilities typically contribute to (..) basic media applications on such AR glasses with simple rendering functionalities”

This contribution proposes to start listing encoding/decoding capabilities which may be of interest for the Edgar-1 MeCAR device, referring to video capabilities which have been referenced in TS 26.511 (“5G Media Streaming, Profiles, Codecs and Formats”).

Video profiles which have been referenced in TS 26.118 (“Virtual Reality (VR) profiles for streaming applications”) refer to a 360° immersive context, which may not be relevant for the optical see-through device type which we are focusing on in a first step.

# Possible candidates for Media Capabilities

In the next chapter, we list, for the different video profiles which may be considered as relevant for consideration in MeCAR. These candidates do not reflect the final recommendations which MeCAR will produce. Additionally the lists may be refined and updated as the work in MeCAR progresses.

## Possible 2D Video profiles candidates

2D video encoding and decoding capabilities have of course to be considered. First, a basic scenario for AR is the virtual TV-set, where a 2D video stream is displayed as an overlay to the real world (for instance on the wall of the living room). Then, 2D video encoding/decoding capabilities may be used to display 3D objects on AR glasses, using the stereoscopic effect, by feeding the two eyes buffers of the AR glasses with the appropriate 2D videos.

We refer here to video capabilities which have been referenced in clause 4.2 of TS 26.511 (“5G Media Streaming, Profiles, Codecs and Formats”). We do not refer to video profiles which have been referenced in TS 26.118 (“Virtual Reality (VR) profiles for streaming applications”). They refer to a 360° immersive context, which may not be relevant for the optical see-through device type which we are focusing on in a first step.

Lastly, the here listed candidates do not preclude of the result of the work which shall be carried on within MeCAR. Especially, if both encoding and decoding profiles are listed as they are referenced in clause 4.2 of TS 26.511, requirements for encoding and decoding capabilities may differ.

### AVC

#### Decoding

TS 26.511 has defined the following decoding capabilities for H.264 (AVC4) in its clause 4.2.1.1:

- **AVC-HD-Dec**: the capability to decode H.264 (AVC) Progressive High Profile Level 3.1 [2] bitstreams, for which the maximum VCL Bit Rate is constrained to be 14 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1,000 and 1,200, respectively.

- **AVC-FullHD-Dec**: the capability to decode H.264 (AVC) Progressive High Profile Level 4.0 [2] bitstreams.

- **AVC-UHD-Dec**: the capability to decode H.264/AVC Progressive High Profile Level 5.1 [2] bitstreams for H.264/AVC with the following additional restrictions and requirements:

- the maximum VCL Bit Rate is constrained to be 120 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively.

- the bitstream does not contain more than 10 slices per picture.

NOTE: High Profile for H.264/AVC excludes Flexible macro-block order, Arbitrary slice ordering, Redundant slices, Data partition.

#### Encoding

The corresponding encoding capabilities (in the sense that their bitstream are decodable by the decoder with the same name) have been defined:

- **AVC-HD-Enc:**

- up to 108,000 macroblocks per second;

- up to a frame size of 3,600 macroblocks;

- up to 120 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bits;

- **AVC-FullHD-Enc:**

- up to 245,760 macroblocks per second;

- up to a frame size of 8,192 macroblocks;

- up to 240 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bits;

- **AVC-UHD-Enc:**

- up to 983,040 macroblocks per second;

- up to a frame size of 36,864 macroblocks;

- up to 480 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bits.

### HEVC

#### Decoding

TS26.511 has defined the following decoding capabilities for H.265 (HEVC) in its clause 4.2.2.1:

- **HEVC-HD-Dec**: the capability to decode H.265 (HEVC) Main Profile, Main Tier, Level 3.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

- **HEVC-FullHD-Dec**: the capability to decode H.265 (HEVC) Main10 Profile, Main Tier, Level 4.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

- **HEVC-UHD-Dec**: the capability to decode H.265 (HEVC) Main10 Profile, Main Tier, Level 5.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

- **HEVC-8K-Dec**: the capability to decode H.265 (HEVC) Main10 Profile, Main Tier, Level 6.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1 with the following further limitations:

- the bitstream does not exceed the maximum luma picture size in samples of 33,554,432,

- the maximum VCL Bit Rate is constrained to be 80 Mbps with CpbVclFactor and CpbNalFactor being fixed to be 1000 and 1100, respectively.

Note that HEVC\_8K-Dec is mentioned as a reminder of TS 26.511 references. It may not be relevant in the context of AR glasses. As said in introduction to this chapter, this discussion is left for further study.

#### Encoding

The corresponding encoding capabilities (in the sense that their bitstream are decodable by the decoder with the same name) have been defined:

- **HEVC-HD-Enc**:

- up to 33,177,600 luma samples per second;

- up to a luma picture size of 983,040 samples;

- up to 120 frames per second;

- the Chroma format being 4:2:0; and

- the bit depth being 8 bits;

- **HEVC-FullHD-Enc**:

- up to 133,693,440 luma samples per second;

- up to a luma picture size of 2,228,224 samples;

- up to 240 frames per second;

- the Chroma format being 4:2:0; and

- the bit depth being either 8 or 10 bits;

- **HEVC-UHD-Enc**:

- up to 534,773,760 luma samples per second;

- up to a luma picture size of 8,912,896 samples;

- up to 480 frames per second;

- the Chroma format being 4:2:0; and

- the bit depth being either 8 or 10 bits.

0

# Proposal

We propose to include clause 2 of this contribution as a sub-clause of clause 5 of the permanent document.

# References

1. TS 26.118 “Virtual Reality (VR) profiles for streaming applications”.
2. TS 26.511 “5GMS Profiles, Codecs and Formats”.
3. TR 26.928 “Extended Reality in 5G”.
4. TR 26.998 “Support of 5G Glass-type Augmented Reality / Mixed Reality (AR/MR) devices”.