AV2 Draft Final Deliverable

Release Notes

Introduction

This document is intended to be used in conjunction with the released AV2 Draft Final Deliverable specification and corresponding AV2 Reference Software (AVM). Its purpose is to highlight known issues that will be addressed prior to the publication of the Final AV2 specification and reference software, and to make implementers aware that work on such issues is on-going. Issues are labeled as [spec] and/or [software], to indicate which aspect(s) of this release are impacted. A number of editorial improvements are planned for the final release to ensure correctness, and for the consistent use of language, naming, and formatting.

The AVM reference software, tagged “v13.0”, is available at:

https://gitlab.com/AOMediaCodec/avm/-/tree/research-v13.0.0?ref_type=tags

As reference software, AVM does not support every conceivable use case, and it has not been fully optimized for production deployment.

1. Terminology

Draft Specification

draft of “AV2 specification v1.0.0”.

Reference Software

reference software (AVM v13.0) for the Draft Specification.

2. Coding tools

2.1. [Spec, software] Restricted switch frames: The current specification does not describe how Quantization Matrices and Film Grain modeling are expected to be handled when Restricted Switch Frame OBUs are encountered in a bitstream.

2.2. [Spec, software] Coding block sizes: There is a known issue in the current specification where content using the YUV 4:2:2 chroma format is allowed to use a 4x64 block size. This will be corrected in the final specification, such that 4x64 and 64x4 coding block sizes are disallowed in all chroma formats.

2.3. [Software] Quantization matrices: There is a known issue in the Draft Software where the subsampling process used to generate 4x32 and 32x4 quantization matrices does not match the Draft Specification, which follows the AV1 subsampling scheme. The final version of the software will be corrected to match the Draft Specification.

3. Multi-layer Support

AV2 will support advanced multilayer features and capabilities that are currently not fully described in the Draft Specification or exercised in the Reference Software. The following aspects will be further clarified in the final version of the specification and reference code:

3.1. [Spec] Semantics for several multilayer related features, such as the semantics of the OPS OBU, are not well defined or incorrect or missing.

3.2. [Spec] The current specification does not clearly describe how different layers, carrying data annotated with different types of information, such as view IDs, auxiliary information, atlas segments etc., can be associated and used for different applications (e.g. in the context of a stereoscopic application that includes transparency or other types of information such as depth). This will be clarified in the final specification. Examples for applications such as sub-region coding using layers, will also be included.

3.3. [Spec, software] Sub-bitstream extraction based on operating point selection is missing from both the Draft Specification and the Draft Software.

4. Temporal units and random access support

4.1. [Spec] The definitions of a temporal unit and the random access points need refinement. In particular, the random access point definition is not complete. Additional requirements beyond the presence of certain OBU types (e.g. a CLK or OLK OBU) are necessary. For example, the availability of any corresponding HLS information (such as an SH OBU) either within the same temporal unit or through external means is also essential for such definitions. More specifically, the definition of the following random access points is missing:

- Closed Loop Key Frame Random Access
- Open Loop Key Frame Random Access
- Random Access Switch Frame

4.2. [Spec] Activation of certain HLS OBUs, such as SH, LCR, and MFH OBUs, is not clearly defined.

5. Profile/Level Capabilities

5.1. [Spec] Compatibility of bitstreams tagged with different “Sequence” level profile indication syntax elements is currently not specified in the Draft Specification. For example, bitstreams tagged with seq_profile_idc = 0, can be decoded by any decoders that are capable of Main_420_10 decoding. However, bitstreams with 0 < seq_profile_idc < 4, chroma_format_idc = 0, and seq_max_mlayer_cnt = 1 are also decodable by a Main_420_10_IP0 decoder. This needs to be more clearly described in the specification with a well defined profile “compatibility” section.

5.2. [Spec, software] The Draft Specification and Draft Software do not properly address the fact that only profiles capable of 8 and 10 bit support are defined. The Draft Specification and the Draft Software assume that there is support for up to 12 bits, and incorrectly set variables for bit_depth_idc > 1 to states relevant to 12 bit decoding. This needs to be addressed. In particular, The Draft Specification should enforce that bit_depth_idc shall not take values greater than 1, and that other values are reserved for future versions of this specification. The related sections of the specification and / or code modules of the reference SW may need to be adjusted.

5.3. [Spec, software] The Draft Specification does not include conformance constraints for the presence of the MSDO and LCR OBUs for different multilayer bitstream configurations and profiles.

5.4. [Spec, software] Some level related constraints on frame buffer management when both Global IBC and in-loop filtering are enabled are not correctly reflected in the Draft Specification. Such constraints result in allocating one frame slot in the decoder buffer pool to hold the reconstructed frame before the in-loop filter is applied.

5.5. [Spec, software] The 4:2:2 and 4:4:4 profiles will be indicated with separate seq_profile_idc values.

6. Metadata

6.1. [Spec] It might be appropriate to move the specific metadata payload types into a separate section (e.g. an Annex) and not include them in sections 5 and 6 with essential decoding information. Such a section should also describe how metadata persistence and inheritance should be handled, which is currently missing.

6.2. [Spec, software] Inheritance of information across embedded layers for Content Interpretation (CI) OBUs is currently not described

6.3. [Spec, software] Issues were identified with the Buffer Removal Timing Metadata.

6.4. [Spec] The syntax of the ITU-T T.35 metadata is incomplete since it does not clearly identify the payload as syntax (it is identified as such in a note, which is not appropriate).

6.5. [Spec, software] The temporal point info metadata is planned to be changed in the final version of the specification.

6.6. [Spec, software] New informative metadata types may be added in the final version of the specification.

7. Decoder model

7.1. [Spec, software] There is ongoing discussion whether decoder model conformance needs to be required to simultaneously account for both models specified in the specification rather than just one model. This may result in changes in the final version of the specification and the reference software.

7.2. [Spec, software] Due to the changes in random access and layer handling in the AV2 design compared to that of AV1, as well as how the output process is handled, additional changes to the details of the decoder model may need to be made in the final version of the specification.

8. Other items:

8.1. [Spec, software] The AV2 specification defines numerous syntax elements and variables that are associated with elements in the CICP specification (ISO/IEC 23091-2 and ITU-T H.273). These references are sometimes incorrect, i.e. use of ISO/IEC 23091-4 throughout, which is only a technical report. In addition the CICP related tables should likely be removed or edited and instead the CICP tables should be directly referenced. This can ensure that if any new values are introduced in the CICP specification, those will be immediately supported without requiring publication of a new version of the specification. Edited tables that provide only samples of values that are useful could still be retained.

8.2. [Spec, software] There are known differences between the Draft Specification and the Draft Software related to syntax elements show_frame and showable_frame. These changes will be aligned and further clarified in the final specification and reference software. Such changes will improve readability and the understanding of how such features can be used and/or implemented.

8.3. [Spec, software] OBU extensibility is not properly reflected in the Draft Specification and will be addressed in the final version of the specification.