Alliance for Open Media (AOMedia) Progress Report

By Andrey Norkin, Adrian Grange, Cyril Concolato, Ioannis Katsavounidis, Hassene Tmar, Khaled Mammou, Shan Liu, and Roshan Baliga

About AOMedia

The Alliance for Open Media (AOMedia)[1] is a nonprofit organization developing media-related specifications to support media distribution in various scenarios, including but not limited to streaming over the internet and realtime communication. The alliance strives to produce royalty-free specifications to nurture the freedom to innovate in the media ecosystem, similar to how innovation was supported in the development of internet technologies. AOMedia is part of the Linux Foundation[2] and its descendant project Joint Development Foundation.[3]

The alliance was created in 2015 and currently has 43 member organizations, including 14 steering committee members. The alliance has four technical Working Groups, which address different aspects of the media ecosystem and are described in the following sections.

AOMedia has issued several media-related specifications. The specifications under development pass through several stages of maturity. The first stage is a Working Draft, which is used by a Working Group to develop the specification in a collaborative manner. When the development is finalized, the specification is promoted to the Working Group Approved Draft status based on a Working Group decision. Finally, the specification becomes the Final Deliverable based on the AOMedia Steering Committee approval. Starting from the Working Draft stage, the specifications under development are typically available publicly on Github.[4] AOMedia welcomes public feedback on its specification under development via a feedback agreement. The reference software for the specification under development is also publicly available, mostly on the AOMedia Gitlab page.[5]

Video Codec Working Group

The Video Codec Working Group (CWG) was the first AOMedia technical group. The group published the AOMedia Video 1 (AV1) video specification in 2018[6] and is currently working on the development of video compression technologies beyond AV1. Reference software for our post-AV1 work [AOMedia Video Model (AVM)] is publicly available,[7] and the v.3.0 anchor has recently been released. The v.3.0 anchor achieves similar quality (measured by YUV PSNR) to AV1 using 11.6% lower bitrate for standard dynamic range (SDR) content, and 15.7% lower bitrate for high dynamic range (HDR) content for the random-access coding configuration.

The CWG is also working to publish the AV1 film grain synthesis technology in a standalone specification that is codec-agnostic. The objective is to make the technology more widely available in applications no matter what underlying compression scheme is used. A Working Draft of the AOMedia Film Grain Synthesis 1 (AFGS1) specification is available online.[8] Signaling of the film grain parameters is expressed in terms of an ITU-T T.35 format user metadata payload.

The CWG is also actively working to improve AV1, with a future revision of the AV1 specification planned, improvements to the AV1 reference software ongoing, and subjective verification of AV1 in process.

The CWG previously worked on the AV1 Image File Format (AVIF) specification.[9] Currently, the AVIF specification is maintained by the Storage and Transport Formats Working Group (STFWG).

Storage and Transport Formats Working Group

The STFWG was created to address standardization of storage and transport formats for media data and time-aligned media-associated metadata; signaling for the combination of media, including synchronization; signaling for the protection of media; signaling for the rendering of media; and management of ITU-T T.35 AOM payloads for codec-agnostic, time-aligned media-associated metadata.

More specifically, the STFWG has been working on

- AVIF specification[9] and its reference software libavif.[10] The latest version of AVIF was approved and published on 15 April 2022. The most notable addition includes support for progressively decodable images.
- Storage of AV1 video streams in ISO Base Media File Format (ISOBMFF) and common media application format (CMAF)-compliant files.[11] The last version was published in December 2019. The group is currently working on a revision and on creating conformance test vectors.
- Carriage of ID3 timed metadata in CMAF[12] (published in April 2019) that addresses the needs for ad delivery and audience measurement applications.
- Specification for use of HDR10+ dynamic metadata with AV1 streams.[13] The specification is publicly

Digital Object Identifier 10.5594/JMI.2022.3190532
Date of publication: 25 August 2022

1545-0279/22©2022SMPTE

available. It is currently in the Working Group Approved Draft stage. The group is also working on conformance streams for this specification.

- Carriage of AV1 video streams in Moving Picture Experts Group (MPEG)-2 transport streams.[14] The specification is currently at the Working Draft stage and is publicly available for feedback.
- Bitstream and container formats for immersive audio content[15] for applications such as internet audio streaming, multicasting/broadcasting services, file download, gaming, communication, and extended reality. The specification is at the Working Draft stage and publicly available for feedback.

Software Implementation Working Group

Founded in July 2020, the Software Implementation Working Group (SIWG) is focused on delivering practical open software implementations of coding standards published by AOMedia. The group is also tasked to help other working groups within AOMedia to evaluate and optimize coding tools in practical software implementations to ensure the successful adoption of new coding specifications.

SIWG's first deliverable was to build, maintain, and enhance an open-source practical AV1 encoder codebase that addresses various use cases of SIWG members and the video industry at large. The codebase used for this purpose is SVT-AV1,[16] and its release 1.0 on 22 April 2022[17] marked the completion of many rounds of quality and speed improvements. As shown in **Fig. 1**, SVT-AV1 performance spans three orders of magnitude in terms of computational complexity, bridging the performance gap that existed previously among different generations of video encoders, from AVC to VP9 to HEVC to AV1 reference software, libaom. Offering excellent BD-rate performance across a wide range of speed settings with almost linear multithreading scaling, SVT-AV1 demonstrates that AV1 is a competitive option to use in video encoding pipelines and products for video-on-demand and live video use cases. SVT-AV1 is also integrated with FFmpeg[18] and GStreamer.[19]

In the process of developing and optimizing SVT-AV1, SIWG member companies have published benchmarking methodologies and results against other open-source software implementations of video coding standards in various conferences.[20,21]

Volumetric Visual Media Working Group

With the rapid development of 3D acquisition technologies (e.g., LiDAR, photogrammetry, and general 3D modeling software) and wide adoption of augmented reality (AR), virtual reality (VR), 3D printing, autonomous driving, 3D mapping, and several other immersive media applications, volumetric media content has become ubiquitous. Addressing technical challenges related to efficient storage, transmission, and realtime playback of such volumetric content is essential for enabling immersive applications.

Recognizing such needs, AOMedia created, in February 2022, the Volumetric Visual Media Working Group (VVMWG), with the goal of developing specifications for efficient coding and streaming of volumetric

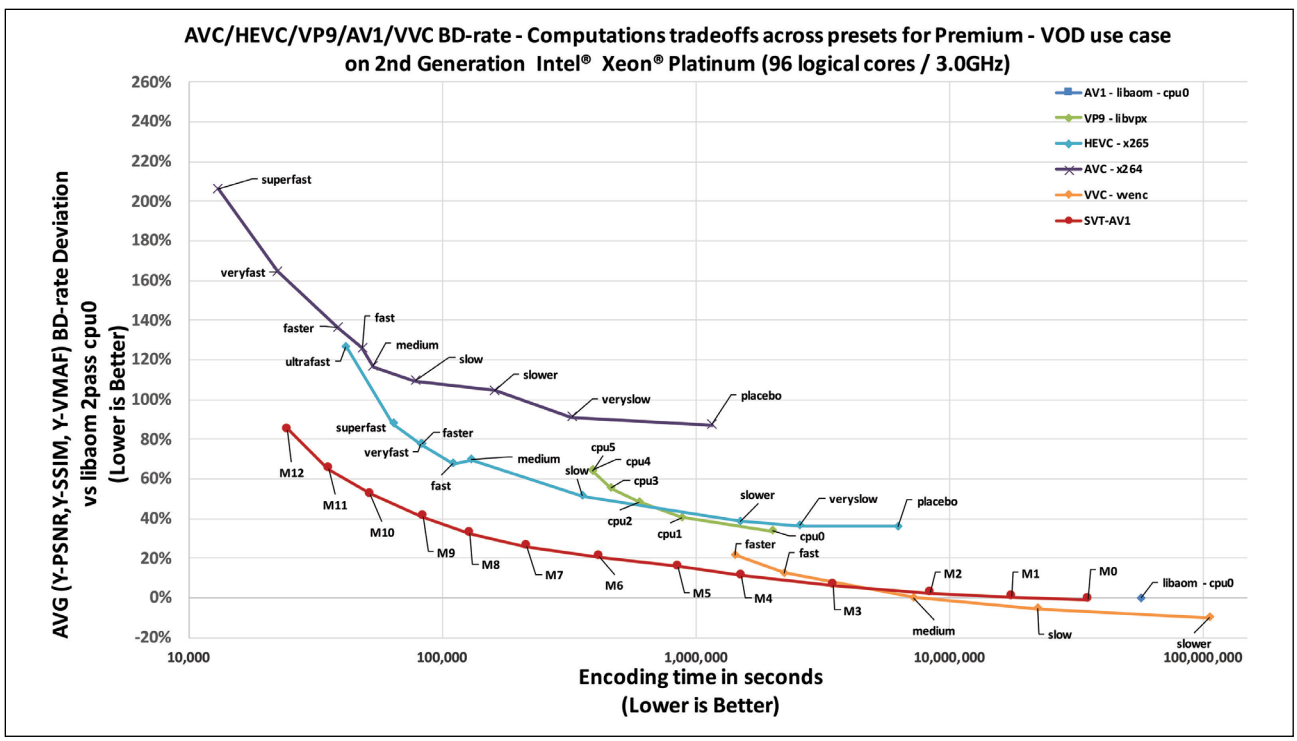

FIGURE 1. BD-rate versus encoding time plot of SVT-AV1 compared to other open encoders—libaom, libvps, x265, x264, and vvenc.

visual media data for immersive applications. The group is mainly focused on developing compression technologies that could be integrated with any scene graph representation.

Since its creation, the VVMWG has been gathering requirements and use cases and preparing a call for proposals for static polygonal mesh compression, which will be published later this year. The corresponding standard is expected to be completed by the end of 2024. Additional specifications for point cloud and 3D animation compression will be developed in parallel and published later.

AV1 Adoption

There is increasing support for AV1 throughout the video industry, and it is currently used by a number of video services, including both linear and nonlinear workflows, such as those from Facebook, iQIYI, Twitch, and Vimeo. YouTube began serving AV1 in high volume at the end of 2018 and has more recently added hardware-based encoders. YouTube's second-generation data center Video Coding Unit (VCU) [application-specific integrated circuit (ASIC) in 10k+ server clusters] supports AV1 hardware encoding up to 8K and is already in production.[22] Netflix began serving AV1 to Android devices in 2020 and is streaming AV1 content to various devices including smart TVs, streaming sticks, smartphones, and game consoles.

AV1 is gaining traction for realtime communications with Cisco rolling out AV1 support in Webex, and Google introducing AV1 in Duo at resolutions up to QVGA.

AV1 support is being added to cloud services. In particular, A1 Alibaba Cloud Video Cloud supports AV1 encoding; Amazon AWS Elemental MediaConvert has supported AV1 since 18 March 2020; ATEME's TITAN service and Bitmovin's "Bitmovin Encoding" service support AV1 encoding.

Allegro DVT, Chips&Media, and VeriSilicon all license their own 8K AV1 hardware design, which may be licensed for integration into custom ASICs, and NETINT Technologies provides an off-the-shelf ASIC that supports 4K live (linear) AV1 encoding.

Several manufacturers have included support for AV1 hardware decoding in their new TV sets. Samsung (4K and 8K QLED models) and LGE (4K and 8K OLED, and 4K nanocell models) added support from 2020, as did Toshiba and Insignia (FireTVs) for their 43–70 in. 4K models. In addition, TCL Europe's X915 series 8K TV and Sony's 2021 and newer 4K and 8K Android TVs support AV1 hardware decoding.

AV1 decoding is supported in many smartphones. MediaTek began supporting 4K at 60 frames/sec AV1 hardware decoding in their 2019 Dimensity 1000 system on chip (SoC), which is used in phones such as Oppo Reno3 5G, Vivo iQOO Z1, Xiaomi Redmi K30 Ultra, and LG Velvet. Samsung's S21 and S22 families of smartphones support 8K AV1 hardware decoding: Samsung Galaxy S21, S21 Plus, S21 Ultra S22, S22 Plus, and S22 Ultra.

AV1 hardware is also available on newer PCs. Intel's Tiger Lake and Rocket Lake CPUs support 8K AV1 hardware decoding. Among the graphics cards, Nvidia's GeForce GTX 3000 series and AMD's Radeon RX 6000 series have support for AV1 hardware decoding, and Intel's Arc GPUs support both AV1 encoding and decoding.

The 2020 Roku Ultra 4K streaming device supports 4K at 60 frames/sec AV1 hardware decoding.

AV1 software-based decoding is available in the following web browsers: 1) Mozilla Firefox, 2) Google Chrome, 3) Opera, and 4) Microsoft Edge. Additionally, Google Chrome supports fixed-function hardware decoding when used on Windows with an Intel Tiger Lake CPU or an Nvidia 30xx series GPU. Mozilla Firefox supports fixed-function hardware decoding when used with Intel Gen 11+, AMD RDNA 2 excluding Navi 24, or Nvidia GeForce 30 series.

AV1 decoding is currently supported in operating systems such as Microsoft Windows, BSD, Linux, ChromeOS, and AndroidOS.

AVIF Adoption

AVIF is an image compression format based on the AV1 intraframe coding that offers superior compression efficiency compared to earlier image compression standards, such as JPEG, GIF, and PNG. AVIF has been gaining support in browsers. According to Ref. 23, AVIF is supported in newer versions of desktop Firefox, Chrome, and Opera. On the mobile side, AVIF is supported in Android Browser, Opera Mobile, Chrome, and Firefox for Android, as well as in the Samsung internet browser.

The following content delivery network (CDN)/services support AVIF: 1) Cloudinary, 2) Cloudflare, 3) Akamai (experimental), 4) Imgix, 5) Adobe AEM Dynamic Media, and 6) ImageEngine.

References

1. Alliance for Open Media. [Online]. Available: https://aomedia.org/
2. Linux Foundation. [Online]. Available: https://www.linuxfoundation.org/
3. Joint Development Foundation. [Online]. Available: https://www.jointdevelopment.org/
4. AOMedia Specifications. [Online]. Available: https://github.com/AOMediaCodec/
5. AOMedia Reference Software. [Online]. Available: https://gitlab.com/AOMediaCodec/
6. AV1 Specification, Version 1.0.0 with Errata, 2019-01-08. [Online]. Available: https://aomediacodec.github.io/av1-spec/av1-spec.pdf
7. AOMedia Video Model (AVM). [Online]. Available: https://gitlab.com/AOMediaCodec/avm
8. AOMedia Film Grain Synthesis 1 (AFGS1), Working Draft. Accessed: Jun. 6, 2022. [Online]. Available: https://aomediacodec.github.io/afgs1-spec/
9. AV1 Image Format (AVIF), v.1.1.0, 2022-04-15. [Online]. Available: https://aomediacodec.github.io/av1-avif/

10. libavif—Software Library for Encoding and Decoding AVIF Files. [Online]. Available: https://github.com/AOMediaCodec/libavif
11. AV1 Codec ISO Media File Format Binding, v1.2.0. [Online]. Available: https://aomediacodec.github.io/av1-isobmff/v1.2.0.html
12. Carriage of ID3 Timed Metadata in the Common Media Application Format (CMAF), v1.0.0. [Online]. Available: https://aomediacodec.github.io/id3-emsg/
13. HDR10+ AV1 Metadata Handling Specification. Accessed: Jun. 7, 2022. [Online]. Available: https://aomediacodec.github.io/av1-hdr10plus/
14. AV1 Specification for Carriage Inside MPEG-2 TS. [Online]. Available: https://aomediacodec.github.io/av1-mpeg2-ts/
15. Immersive Audio Container. [Online]. Available: https://aomediacodec.github.io/iac/
16. SVT-AV1 Software. [Online]. Available: https://gitlab.com/AOMediaCodec/SVT-AV1
17. SVT-AV1 Release 1.0.0. [Online]. Available: https://gitlab.com/AOMediaCodec/SVT-AV1/-/releases/v1.0.0
18. FFmpeg. [Online]. Available: https://www.ffmpeg.org/
19. Gstreamer. [Online]. Available: https://gstreamer.freedesktop.org/
20. F. Kossentini et al., "The SVT-AV1 Encoder: Overview, Features and Speed-Quality Tradeoffs," *Proc. SPIE 11510, Applications of Digital Image Processing XLIII, 1151021*, 21 Aug. 2020. https://doi.org/10.1117/12.2569270
21. P-H Wu et al., "Towards Much Better SVT-AV1 Quality-Cycles Tradeoffs for VOD Applications," *Proc. SPIE 11842, Applications of Digital Image Processing XLIV, 118420T*, 1 Aug. 2021. https://doi.org/10.1117/12.2595598
22. P. Ranganathan et al., "Warehouse-Scale Video Acceleration: Co-Design and Deployment in the Wild," *ASPLOS 2023 Abstract*. [Online]. Available: https://asplos-conference.org/abstracts/asplos21-paper388-extended_abstract.pdf)
23. *Can I Use? AVIF Support in Browsers*. Accessed: Jun. 10, 2022. [Online]. Available: https://caniuse.com/avif

About the Authors

Andrey Norkin is co-chair of the Alliance for Open Media's (AOMedia) Royalty-Free Video Codec Working Group (CWG). He actively contributed to the development of the AV1 video codec and is currently working on video compression technologies beyond AV1. He is a senior research scientist at Netflix and has also been involved in the development of encoding techniques for over the top (OTT) video streaming.

Adrian Grange is a member of Google's Open Codecs team and was an active contributor to the VPx and AV1 video codecs. After helping to form the Alliance for Open Media (AOMedia), he chaired the Codec Working Group (CWG) throughout the AV1 development, and is currently co-chair of the ongoing effort to design "next-generation" coding tools. He also runs the Chrome University Research Program with Google.

Cyril Concolato is a senior software engineer working on studio and streaming technologies with Netflix. He is chair of AOM's Storage and Transport Format Working Group (STFWG) and editor of various AOM specifications such as the carriage of AV1 in ISO Base Media Files or the AV1 Image Format.

Ioannis Katsavounidis is a research scientist with the Video Infrastructure team with Meta. Before that, he worked for Netflix, Cidana, and InterVideo, and was an associate professor in the Electrical and Computer Engineering Department, University of Thessaly, Volos, Greece.

Hassene Tmar joined eBrisk Video in 2013, where he developed video quality improvement algorithms for HEVC-based real-time encoders. As a technical product manager at Intel, he led a team to open source the Scalable Video Technology codecs (SVT-HEVC, SVT-VP9, and SVT-AV1), optimized such encoders on Intel Xeon CPUs, and provided extensive technical support to Intel's customers for next-generation video transcoding products based on the open-source SVT encoders. Then, he joined Meta as a technical program manager for media and calling, helping the Video Processing Infrastructure Team.

Khaled Mammou is a senior engineer at Apple working on designing 3D compression solutions. He is co-chair of the Alliance for Open Media (AOMedia) Volumetric Visual Media Working Group (VVMWG). He has been actively contributing to 3D standardization activities since 2005, with a track record of several published standards, patents, books, and research articles.

Shan Liu is currently a distinguished scientist and general manager with Tencent. She has been actively contributing to various international standards over the past decade and served as coeditor of H.265/HEVC SCC and H.266/VVC. She also leads global research and development teams to serve a hundred million users.

Roshan Baliga is a group product manager with the Open Codecs Team at Google. He primarily works on video codecs and video hardware.