AV1: Implementation, Performance and Applications

Jerome Vieron and Ian Trow

Ateme, France and UK

ABSTRACT

Now that the specification is finalised, the Alliance for Open Media has set high expectations for AV1. The prospect of a royalty free alternative to HEVC has many industry supporters, particularly those addressing streaming applications for a standard that is optimised for OTT delivery. Fundamental to AV1's success will be the transfer of the tool set to efficient software implementations that can deliver the expected performance over existing compression standards within the constraints of future server provision. Efficient implementations of AV1 will require a careful balance between utilising bespoke processor features, to offload demanding processor intensive functions, and the desire to have a platform agnostic compression standard that can be applied across a wide range of infrastructure. As with all complex compression schemes, realising an efficient implementation of AV1 is crucial if it is to succeed against rival standards. The royalty free aspect of AV1 is significant, and when combined with the changes in viewer consumption habits and the increased processor resources available subtly changes the success criteria compared with the introduction of predecessor standards.

INTRODUCTION

AV1 is an industry driven initiative by the Alliance for Open Media (AOM) with the aim of producing an interoperable and open video CODEC suitable for internet delivery of Over The Top (OTT) services.

With video currently accounting for around 73% of global internet traffic, rising to 82% by 2021(1) and the network issues associated with IP delivery infrastructure, it's clear there is an ongoing need for further video compression standards optimised for internet applications if services are to scale.

From a system-based perspective, AV1 benefits from being based upon VP9, which was developed for streaming applications. This is key to adopting a container-based approach to rationalise workflows. The emphasis on video infrastructure has fundamentally shifted from direct to home to a hybrid model where live, OTT and catch-up services need to be provided for new infrastructure commissions. AV1 has been drafted supporting the key streaming parameters essential for OTT, thereby reflecting the complex hybrid nature of IP based content delivery. Consequently, the AV1 standard aims to deliver a compression standard that can be intergrated into solutions that allow for the needs of OTT systems deployed in the cloud, while retaining the necessary legacy compatibility with existing live and file based infrastructure. The suitability of AV1 to meet the performance expectations and application needs of internet streaming, especially in comparison with competing compression standards like HEVC, is the main focus of this paper.

AV1 CURRENT STATUS

AV1 has reached its final stage of specification, as of June 25th, 2018. Conceived in September 2015 by the Alliance for Open Media, AV1 aims to be a royalty free alternative to existing standards like HEVC. Key contributors include content producers (Amazon and Netflix), web browser developers (Apple, Google, Microsoft and Mozilla) as well server hardware manufacturers (AMD, ARM, Intel and NVIDIA).

AV1 ARCHITECTURE

Functionally, AV1 appears very similar to preceding MPEG standards, see Figure 1.

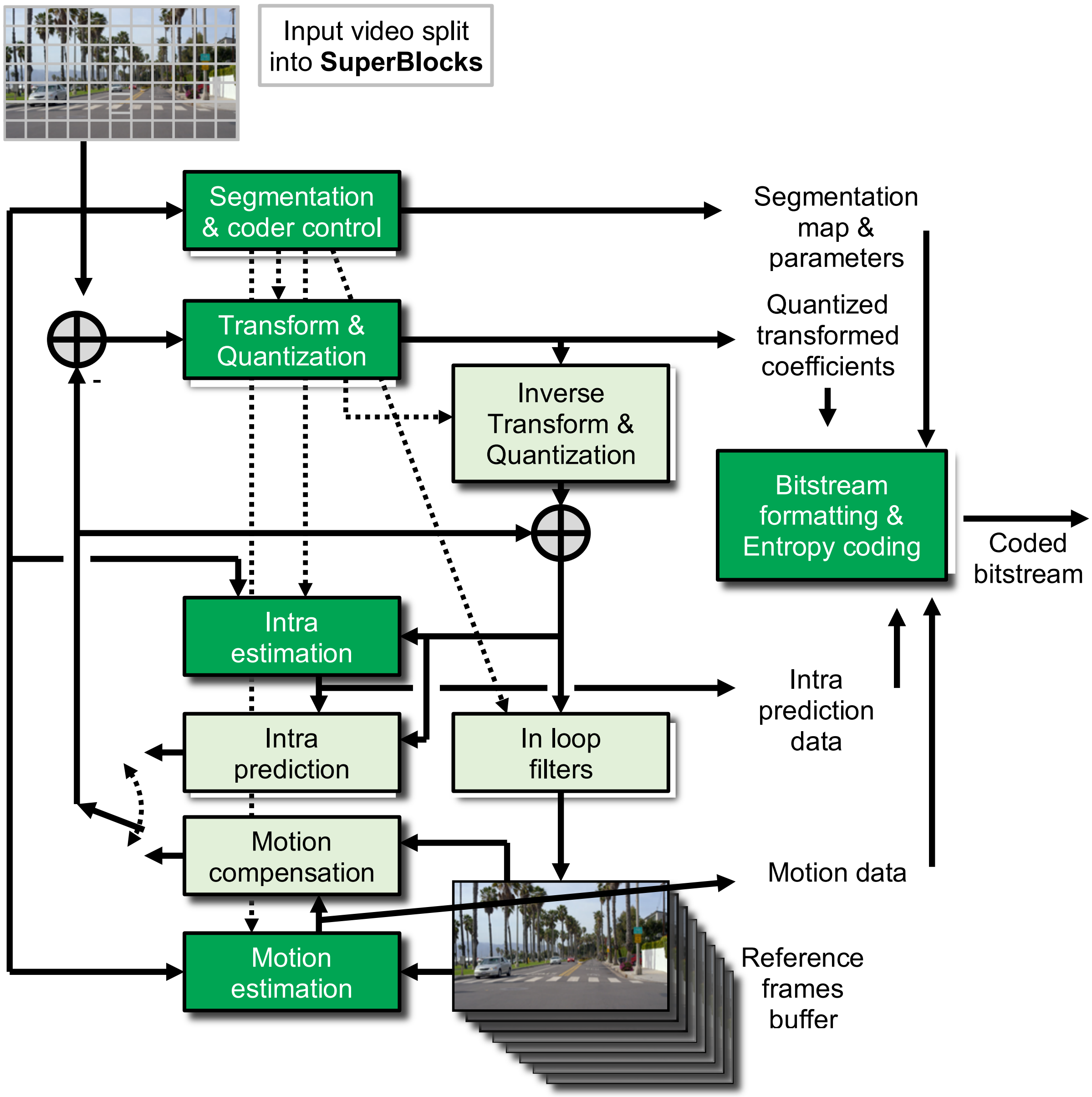

Figure 1 – AV1 Architecture

A closer analysis is required to distinguish from the architecture of the current HEVC standard, which is widely used as the most obvious benchmark for performance comparison. While not the focus of this paper, it is the royalty free aspect of AV1 that is cited as the major competitive impetus to development. However, even though the standard proports to be royalty free and offers considerable industry support to mitigate those implementing AV1 to IPR and patent challenges, the key technical parameters by which AV1 will be judged relate to processor resources to implement, bit rate savings and encoder run time for file applications. These comparisons are by no means fixed references. The Joint Video Exploration Team (JVET) are well underway in improving performance and have already produced the Joint Exploration Model (JEM) which is delivering more than 25% improvement over the HEVC hardware model.

Consistent with HEVC and the applications targeted by JEM, AV1 aims to improve suitability for internet based delivery for OTT as well as handling UHD and associated wide color gamut and high dynamic range features.

AV1 CODING STRUCTURE AND TOOLS

The coding structure of AV1 is shown in Figure 2 where differences start to become evident in terms of coding, block handling, hierarchical / recursive tree techniques, transform and prediction handling when compared to HEVC.

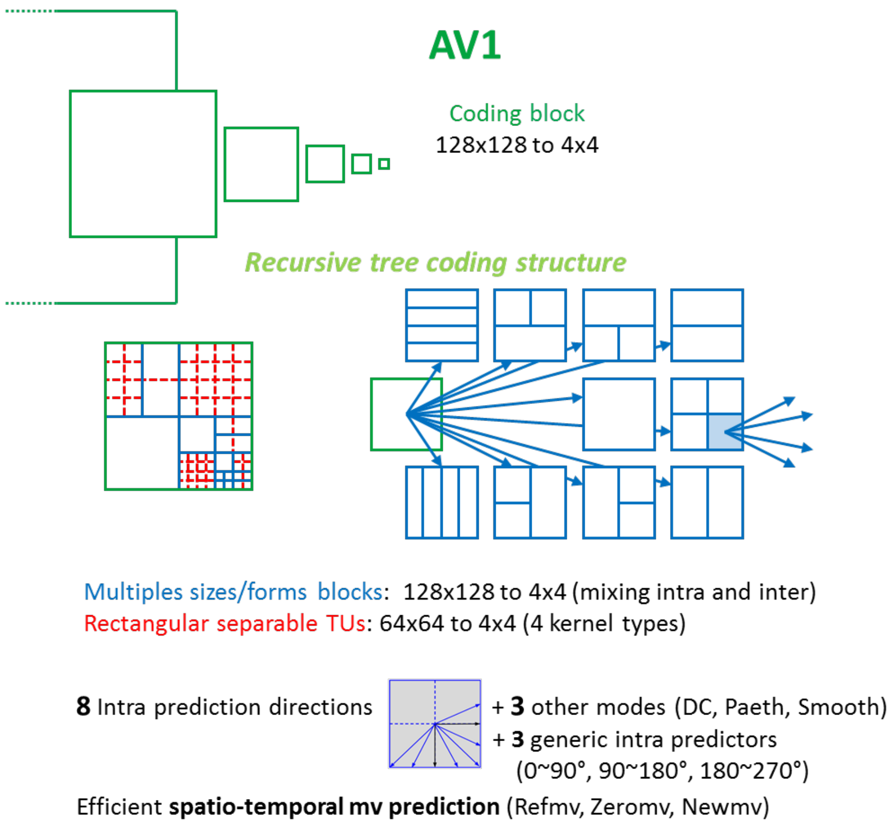

Figure 2 – AV1 Coding Structure

From initial studies of AV1, the recursive tree coding can offer bit rate gains for equivalent video quality compared with predecessor standards like HEVC. The recursive approach in AV1, along with line by line processing and enhanced loop filters results in less ringing and contouring.

An area where efficiency is currently lacking within AV1 concerns motion prediction, where the flexibility offered by the Nearestmv approach in AV1 cannot yet match the efficiency of HEVC. Many of the techniques in AV1 were considered in previous standards and are now included based on the performance now possible from servers. One such feature is the 128x128 superblocks that is undoubtedly of benefit for UHD, but of little benefit for lower resolution formats. While AV1 offers significant flexibility and allows for prediction down to 4x4 blocks, the benefit of such flexibility does come at a cost in terms of signalisation that needs to be reconciled. From modelling AV1, the availability of 4x4 blocks significantly improves subjective picture performance for lower resolution videos.

Clearly more optimisation is needed to understand what predictor combinations result in the biggest returns in terms of compression efficiency without imposing unreasonable processor or run time overheads. It is important to factor in the target application when considering the coding structure of AV1. While many of the coding techniques proposed are currently beyond consideration for live applications on current servers, much of the coding of video content targeted for OTT applications is off-line and large savings for high demand content can justify considerable resource in terms of processing and run time. An example of this concerns warped motion, a new coding tool offered within AV1, see Table 1. The processor overhead is currently very high to implement warped motion and requires more analysis is of questionable benefit from an efficiency stand point. However, such an approach could be of benefit for Off line Video On Demand (VOD) applications, where even small gains result in huge savings over CDN's. This rationale can also be applied to compound prediction (possibly mixing intra and inter prediction together, with a wedge-based blending), and the 4 transform core types that are hard to currently accommodate due to increased encoder complexity but could conceivably become viable during the lifetime of the finalised AV1 standard.

Encoder run time and compression efficiency have been identified as key areas in recent performance comparisons between AV1, JEM, VP9 and HEVC encoders. The conclusion being that AV1 lags behind both JEM and the HM reference model (2). The use of reference models provides an indication of the ultimate performance potential of a compression standard.

AV1 CODING TOOLS

One hundred and seventy new tools were considered for AV1, from which 116 are currently selected. Proposals for all aspects of compression encoding shown in the AV1 architecture in Figure 1 have been implemented. Table 1 itemises the major tools that are different to those in previous compression schemes and require examination.

While the evaluation of the AV1 toolset is still in progress, early analysis indicates the tools most likely to deliver significant gains, along with those tools yielding modest gains and likely to be disabled for applications where optimisation in terms of processor footprint and encoder runtime are required (i.e. live streaming), see Table 2.

Tool	Description
Ext_intra	Generic directional intra predictor to enhance the 8 directional predictors available originally
Motion_var	Overlapped block prediction strategy exploiting neighboring temporal predictors, thus refining the motion model
Warped_motion Global_motion	New motion predictors which consist in geometrical transform
Ext_Inter	Use weighted compound prediction with variable weights inside the block
Dual_Filter	4 interpolation filters in horizontal and vertical directions (three 8-tap and one 12-tap)
Cdef	Deringing filter that takes into account the directions of edges and patterns being filtered
Ref_mv	Adapts the candidate list according to the number of available reference motion vectors
Reference_buffer	Loss detection and recovery method
Ext_tx/ Rect_tx	New transform types (DCT, ADST, IDTX, FlipADST, Rect)-Separable kernels
Tiles_group Independent_row_tiles	Independent group of tiles. Inside this group row of tiles could be independent or not
Delta_q + Ext_Delta_q	Signals delta quantizer at the superblock level + delta loop filter as well but only when delta-q is enabled
Ext_refs	Extends the number of references to six and provides more flexibility on bi-prediction

Table 1 – Key Coding Tool Introductions for AV1

The focus of AV1 investigations has to date concerned identifying which are the most promising tools in terms of coding gain. However, both run time and processing overhead are key considerations and areas of concern for those wishing to realise AV1 as a software CODEC, especially when compared to alternate compression strategies. This aspect forms a large part of the effort to realise an AV1 coder and will certainly limit any implementation in terms of the number & scope of tools utilised.

However, it should be noted that while new tools have been added at every stage of the encoding process (partitioning, intra/inter prediction, transform, quantization, loop filtering and entropy coding) not all of those proposed were focused on coding gain. An example being the frame level context adaptive arithmetic coder, which aims to improve on throughput from a hardware perspective rather than just outright coding efficiency.

From early evaluations using AOM common test sequences, it is evident that the AV1 tool set currently performs well for off-line encoding of movie content. This ties in with the 3 quality levels modes (real-time, good and best), where good and best modes are dedicated to file encoding. This application allows for the use of features such as two-pass encoding.

AV1 Tool	% Gain	
segment_globalmv + palette_throughput + ext_comp_refs	0.14%	Mean
	0.01%	Min
	0.30%	Max
ext_partition_types good add a lot to complexity	1.76%	Mean
	1.04%	Min
	2.65%	Max
tx64x64	0.17%	Mean
	0.12%	Min
	0.26%	Max
filter_intra	0.45%	Mean
	0.36%	Min
	0.71%	Max
ext_skip	0.03%	Mean
	-0.20%	Min
	0.24%	Max
aom_qm objective vs PSNR subjective	0.00%	Mean
	-0.01%	Min
	0.00%	Max
ext_partition 128 by 128	0.89%	Mean
	0.57%	Min
	1.12%	Max
ext_intra_mod	0.20%	Mean
	0.12%	Min
	0.35%	Max

Table 2 – AV1 Tool Coding Gains

AV1 PERFORMANCE

Four operating points are defined for evaluating AV1 performance which are defined by combining two parameters, namely low / high latency and constant / unconstrained Quantization Parameter (QP). To date, while the AV1 CODEC has quality levels for real-time coding, the emphasis has been on off-line encoding allowing two-pass encoding and high latency mode. To apply AV1 to real-time applications will require a specifically optimized subset of AV1 tools and is the current focus of on-going evaluations. Currently, AV1 benchmarking shows the standard to improve on VP9 in terms of bit-rate saving, but this gain has been achieved by at the expense of encoder run time, taking 35 times longer than the original VP9 sequence (2).

Compared with the Hardware Model (HM) HEVC encoder, AV1 has been shown to currently offer equivalent objective performance. Rate-distortion (measured in PSNR) curves for two sequences are shown in Figure 3 and Figure 4.

Early subjective testing confirms the findings in the objective test results. Additionally, while dependent on the test content type, AV1 was found to produce visually more pleasing smoother results. The HEVC encoder was found to be sharper but did suffer from significant quantization noise.

Figure 5 illustrates AV1 preserving sharp edges where HEVC encoding HD at 1Mbps shows ringing around edges. In contrast though, Figure 6 shows excessive smoothing by AV1.

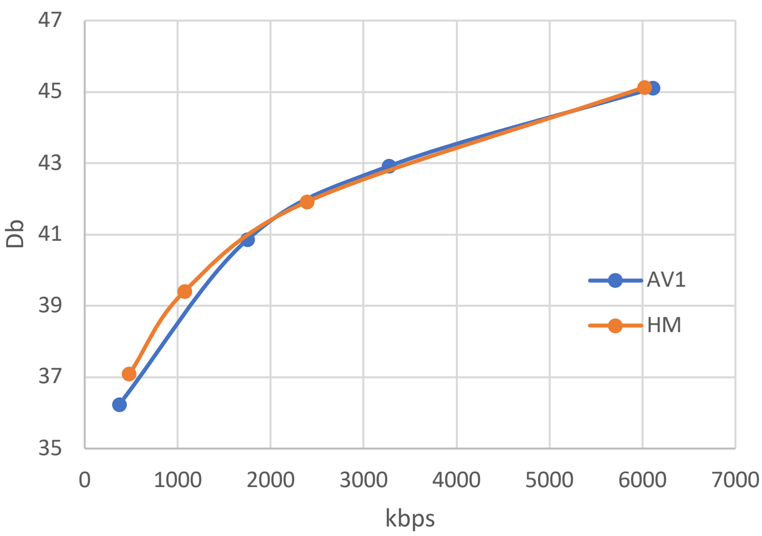

Figure 3 : Comparison between AV1 and HEVC. Kong action movie HD 1080p24 sequence, PSNR versus rate.

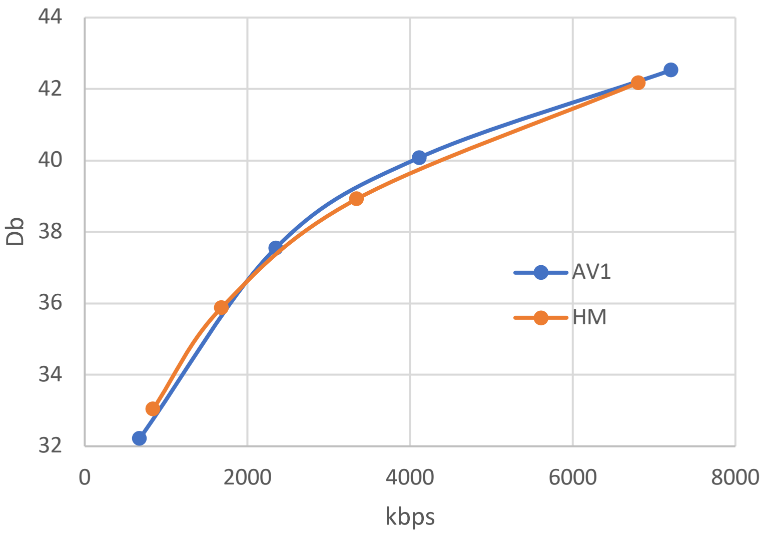

Figure 4 : Comparison between AV1 and HEVC. Netflix foodMarket TV documentary HD 720p60 sequence, PSNR

Analysis of the results suggests that improved spatial-temporal rate allocation could address the shortcomings demonstrated. However, it is clear that AV1 requires significant work to take the finalised standard to a stage where it can meet the needs of practical applications in terms of objective performance.

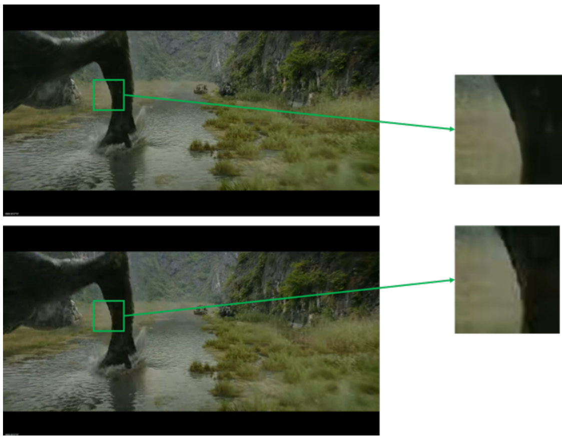

Figure 5 AV1 (top) vs HEVC (bottom)

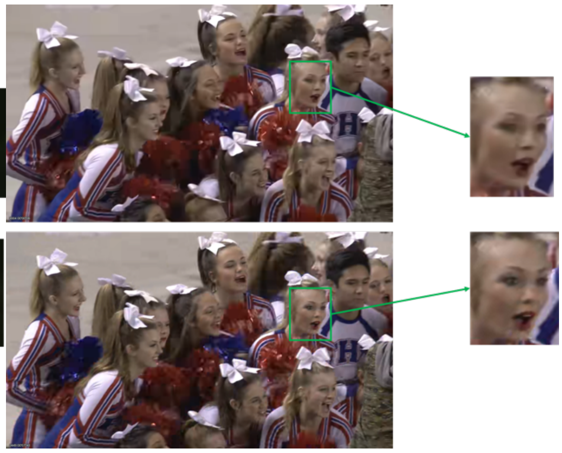

Figure 6 AV1 (top) vs HEVC (bottom)

CONCLUSIONS

Results demonstrate AV1 to have equivalent performance to HEVC in terms of both objective and subjective video quality test results.

For AV1 to be a technical success, further work is needed to improve run time encoding and processor resource requirements to allow practical software implementations to be developed. In particular, a detailed analysis of the tool set is required to evaluate which tools justify inclusion in practical software encoders. With individual tools typically yielding modest coding gains and the complex computational needs of the new AV1 tools, this is now a focus for on-going developments. This will be crucial for AV1 to widen its appeal and support real-time encoding, essential for live streaming applications currently targeted by OTT applications. AV1 is showing promise encoding offline movie content and, while demanding in terms of encoder run time and processor resource, it is anticipated that performance improvements in underlying servers coupled with the significant CDN savings of popular distributed content, will enable widespread AV1 adoption.

REFERENCES

1. Cisco Visual Networking Index: Forecast and Methodology, 2016-202. Cisco Systems Inc. June 2017 https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf
2. Grois, D. Nguyen, T. and Marpe, D., 2017. Performance Comparison of AV1, JEM, VP9 and HEVC Encoders. Proceedings of SPIE Vol. 10396 2017 SPIE
3. Akyazi, Pinar; Ebrahimi, Touradj: “Comparison of Compression Efficiency between HEVC/H.265, VP9 and AV1 based on Subjective Quality Assessments” 10th International Conference on Quality of Multimedia Experience (QoMEX), Sardinia, Italy, May 29 - June 1, 2018.

ACKNOWLEDGEMENTS

The authors would like to thank Elie Mora and Thomas Guionnet for their help in writing this paper.