AV1 Video Codec

Introduction

The field of digital video has witnessed a relentless pursuit of efficient compression techniques to manage the ever-increasing volume of video data generated and consumed globally. Video compression, at its core, aims to reduce the number of bits required to represent video content by identifying and eliminating redundant information. This redundancy can manifest spatially within a single frame, temporally across consecutive frames, or perceptually in details that the human visual system is less sensitive to. Over the decades, a series of video coding standards have emerged, each building upon its predecessors to achieve better compression ratios and support new functionalities. These standards, including MPEG-2, H.264 (Advanced Video Coding), and H.265 (High Efficiency Video Coding), have played a crucial role in enabling the widespread adoption of digital video in various applications, from DVD and Blu-ray to online streaming and broadcasting. However, each of these standards has presented its own set of limitations, particularly in handling the demands of ultra-high resolutions and emerging video formats, and in navigating the complexities of patent licensing [1].

In this evolving landscape, the AV1 (AOMedia Video 1) codec has emerged as a next-generation solution, developed by the Alliance for Open Media (AOMedia), a consortium of leading technology companies [2]. AV1 distinguishes itself as an open and royalty-free video coding format, designed to address the shortcomings of previous standards and to meet the growing demands of the internet video ecosystem [2]. The primary design objectives of AV1 include achieving high compression efficiency, maintaining a low computational footprint for practical implementation, ensuring feasibility for hardware decoding, and being specifically optimized for internet-based video delivery [5]. Backed by major industry players such as Netflix, YouTube, and Facebook, as well as browser vendors, AV1 has seen increasing adoption, signaling its potential to become a cornerstone of future video technology [2]. A significant driving force behind AV1's development is its royalty-free licensing model, which offers a compelling alternative to patent-encumbered codecs like HEVC, where licensing complexities and costs have often hindered widespread adoption, particularly in open-source projects and web browsers [2]. This report aims to provide a comprehensive and technical analysis of the AV1 video codec, with a specific focus on the compression techniques it employs and the underlying reasons for its superior performance compared to previous generation codecs.

The continuous drive for better video compression is a fundamental aspect of the digital media landscape, directly influenced by the relentless increase in video resolution and the need to efficiently manage bandwidth and storage. As display technologies advance to support higher resolutions like High Definition (HD), Ultra High Definition (UHD), and even 8K, the sheer volume of raw video data escalates dramatically. This necessitates the development and deployment of more sophisticated compression algorithms to enable seamless streaming and cost-effective storage. AV1, with its suitability for these high resolutions [11], is a direct outcome of this trend. Furthermore, the intricate and often costly patent situation surrounding video codecs like HEVC has created a strong impetus for the development and adoption of royalty-free alternatives such as AV1 [2]. The difficulties in navigating HEVC's licensing terms, as highlighted in various sources [11], underscore the critical need for an open and accessible codec. The widespread embrace of AV1 has the potential to revolutionize how video content is delivered and consumed, allowing more creators and platforms to offer high-quality video experiences without the encumbrance of complex licensing fees. This could lead to greater innovation and accessibility within the digital video ecosystem, particularly benefiting open-source communities and emerging markets where cost sensitivity is paramount.

AV1 Codec Architecture and Technical Specifications

The AV1 video codec employs a hybrid block-based predictive coding approach, a common framework shared by many modern video compression standards [5]. In this scheme, video frames are divided into blocks, and redundancy is reduced through prediction techniques. AV1 utilizes both intra-frame prediction, where the content of a block is predicted based on the information within the same frame, and inter-frame prediction, where information from previously (and sometimes future) frames is used to predict the current block [5]. Following the prediction stage, the difference between the original block and the predicted block, known as the residual, undergoes a series of processing steps including transform coding and quantization. Transform coding converts the residual into the frequency domain, where the energy is often concentrated in fewer coefficients. Quantization then reduces the precision of these coefficients, thereby achieving lossy compression. Finally, entropy coding is applied to the quantized coefficients and other signaling information to achieve further lossless compression [5]. As an optional step, AV1 also includes film grain synthesis as an out-of-loop post-processing stage, which can enhance the perceptual quality of the decoded video, particularly for content that originally contained film grain [5].

The encoded AV1 bitstream is organized into a sequence of Open Bitstream Units (OBUs) [5]. These OBUs are the fundamental syntactic structures within the AV1 bitstream, each comprising a header and a payload [5]. The header identifies the type of information contained within the OBU and specifies the size of the payload, which holds the actual compressed video data or metadata. Various types of OBUs are defined in the AV1 specification to carry different kinds of information necessary for decoding. These include the sequence header OBU, which contains global parameters for the entire video sequence; the frame header OBU, which provides coding information specific to each frame; the tile group OBU, which contains the coded data for one or more tiles within a frame; the metadata OBU, which can carry information such as high dynamic range (HDR) metadata; and the padding OBU, which can be used to ensure proper bitstream alignment [5].

AV1 offers considerable flexibility in terms of supported color formats and pixel representations. It supports a range of chroma subsampling formats, including 4:0:0 (monochrome), 4:2:0, 4:2:2, and 4:4:4, catering to diverse content and quality requirements [5]. The codec also supports pixel representations with bit depths of 8, 10, and 12 bits per sample, allowing for accurate representation of video with varying levels of color precision, including high dynamic range (HDR) content [5]. AV1 is designed to efficiently handle high-resolution video, including High Definition (HD), Ultra High Definition (UHD), and 8K resolutions, as well as high frame rates, making it well-suited for modern video applications [2]. Furthermore, AV1 supports the ITU-R Recommendation BT.2020 color space, ensuring compatibility with wide color gamut displays and enabling a richer and more accurate color representation [2].

The fundamental unit for partitioning and processing in AV1 is the superblock [1]. A superblock is a square block of luma samples, the size of which can be either 128x128 or 64x64, as signaled in the sequence header [2]. This larger block size, compared to the macroblocks in H.264 (16x16) and the coding tree units (CTUs) in HEVC (up to 64x64), allows for more efficient coding of large, uniform areas within a video frame [2]. Superblocks can be recursively partitioned into smaller coding blocks, down to a minimum size of 4x4 pixels, using a flexible 10-way partitioning structure [1]. AV1 introduces new partitioning patterns, including T-shaped splits and horizontal or vertical splits into four stripes with a 4:1 or 1:4 aspect ratio, providing more options to adapt to the shapes of objects and regions within the video content [1]. Notably, each resulting coding block can independently select its own prediction mode (intra or inter) and transform modes, allowing for optimized coding based on the local characteristics of the video signal [20].

For parallel processing, an AV1 frame can be divided into rectangular regions called tiles [5]. These tiles can be decoded independently, enabling parallel decoding and potentially faster processing times, especially on multi-core architectures [2]. AV1 supports both uniform tile partitioning, where all tiles within a frame have the same dimensions, and non-uniform tile partitioning, allowing for more flexible adaptation to the video content or processing requirements [5]. Each tile in an AV1 bitstream might include an additional header that specifies its reference frame index and its position within the current frame, ensuring that each tile can be correctly decoded even when referencing other frames [5]. The tile structure in AV1 is a key architectural feature that facilitates parallel encoding and decoding, which is crucial for managing the computational complexity associated with the codec's advanced compression techniques and for enabling real-time video applications.

Intra-Frame Prediction Techniques in AV1

Intra-frame prediction in AV1 exploits the inherent spatial correlation between neighboring pixels within the same video frame to predict the content of a block [2]. This is particularly important for I-frames, which are coded without reference to other frames and serve as random access points in the video sequence. AV1 significantly expands upon the directional intra prediction modes offered by its predecessor, VP9. While VP9 provided 8 directional prediction modes, AV1 extends this to a total of 56 directional modes by introducing finer angle variations [6]. These finer angles are achieved through an angle delta parameter, ranging from -3 to +3 with a step size of 3 degrees, which is added to the nominal angles inherited from VP9 [6]. This increased granularity allows AV1 to more accurately predict edges and directional textures within a block, resulting in a smaller prediction residual and improved compression efficiency. These extended directional prediction modes are implemented using a unified directional predictor that links each pixel to a reference sub-pixel location on the edge of the block and interpolates the reference pixel using a 2-tap bilinear filter [6].

In addition to directional prediction, AV1 employs several non-directional smooth intra prediction modes. These include the DC prediction mode, which predicts all pixels in a block with the average value of the surrounding reference pixels and is effective for flat regions [6]. AV1 also introduces three new smooth prediction modes: SMOOTH_PRED, SMOOTH_V_PRED, and SMOOTH_H_PRED [6]. These modes are designed to better predict blocks with smooth gradients by using a distance-weighted average of the boundary pixels, effectively performing quadratic interpolation. SMOOTH_V_PRED and SMOOTH_H_PRED are tailored for vertical and horizontal gradients, respectively, while SMOOTH_PRED is an average of both. Furthermore, AV1 replaces the TM_PRED mode from VP9 with PAETH_PRED [2]. The Paeth predictor aims to select the best prediction from the left, top, and top-left neighboring pixels based on a simple calculation, proving particularly effective for predicting edges and corners.

AV1 also features recursive-filtering-based intra prediction through the FILTER_INTRA modes, which are designed for luma blocks and treat them as 2D non-separable Markov processes [6]. These modes use a set of pre-designed 7-tap filters that are applied to 4x2 pixel patches within the block, with each patch being predicted using one of five available filter sets [6]. Another significant intra prediction tool in AV1 is Chroma from Luma (CfL) prediction [5]. CfL is a chroma-only intra predictor that leverages the correlation between the luma and chroma components by modeling chroma pixels as a linear function of the co-located reconstructed luma pixels. Importantly, the scaling parameters for this prediction are determined based on the original chroma pixels and explicitly signaled in the bitstream, leading to more accurate chroma prediction [5].

For screen content, AV1 introduces the Intra Block Copy (IBC) mode [5]. IBC allows the current block to be predicted by copying a previously decoded block from the same frame, functioning as a form of motion compensation within a single frame. This is particularly effective for encoding screen content, which often contains repeated patterns such as text and graphics [6]. When IBC is enabled in intra-only coded frames, loop filtering is disabled, and motion vectors are restricted to integer precision [6]. AV1 also offers a Palette Mode, specifically designed for content with a limited color palette, like screen sharing or certain types of animation [6]. In this mode, the encoder identifies the unique colors (typically 2 to 8) within a block, stores them in a palette, and then encodes each pixel in the block by its index in the palette, achieving high compression ratios for such content [20].

Inter-Frame Prediction Techniques in AV1

Inter-frame prediction is a crucial aspect of video compression, aiming to reduce temporal redundancy by predicting the content of the current frame based on information from previously (and sometimes future) frames [2]. AV1 employs sophisticated motion estimation and compensation techniques to achieve this, utilizing motion vectors with sub-pixel accuracy. For luma components, this accuracy is up to 1/8 of a pixel, and for chroma components, it can reach up to 1/16 of a pixel [20]. Sub-pixel values are generated using separable interpolation filters, with AV1 adaptively selecting between SMOOTH, REGULAR, and SHARP filters independently for the horizontal and vertical directions, allowing for better adaptation to the characteristics of the motion [5].

AV1 utilizes Advanced Motion Vector Prediction (AMVP) to improve the efficiency of motion vector coding [22]. This involves analyzing motion vectors of neighboring blocks and examining motion patterns in previous frames to predict the most likely motion vector for the current block, thus reducing the number of bits needed to encode the motion vector itself. To handle more complex types of motion, such as camera pan, zoom, and rotation, AV1 introduces affine and warped motion compensation [2]. This includes both global motion compensation, which applies to the entire frame, and local motion compensation, which is applied at the block level [20]. AV1 supports similarity and affine transformations to model these complex movements more accurately.

Compound prediction is another key feature of AV1, where predictions from two different reference frames are combined to generate the final prediction for a block [5]. AV1 offers various blending modes for compound prediction, including weighted average, distance-weighted compound (where weights are based on the temporal distance of the reference frames), and wedge-partitioned prediction (where a block is split using a wedge pattern, and different predictions are used for each partition) [2]. Notably, AV1 limits the combinations of unidirectional reference frames for compound prediction to only four possible pairs, while supporting all twelve combinations in the bidirectional case [5].

AV1 significantly expands the number of reference frames that can be used for inter prediction, allowing a coding frame to choose any 7 frames from the 8 available frames in the decoded frame buffer [5]. These reference frames include LAST, LAST2, and LAST3 (near past frames), GOLDEN (a distant past frame), BWDREF, ALTREF, and ALTREF2 (future frames) [6]. Alternate Reference Frames (ARF) are also utilized; these are frames that are coded and stored but may not be displayed, serving as references for subsequent frames and often built using temporal filtering to retain common information while removing noise [5]. The "show existing frame" functionality allows the codec to directly use a frame already in the decoded frame buffer for display, without needing to code a new frame [5].

To address visible discontinuities at block boundaries that can arise from block-based motion compensation, AV1 employs Overlapped Block Motion Compensation (OBMC) [2]. This technique extends the block sizes used for motion compensation to overlap with neighboring blocks, and the overlapping regions are blended together to create smoother transitions.

Finally, AV1 introduces Switch frames (S-frames), a new type of inter-frame prediction that can be predicted using already-decoded reference frames from a higher-resolution version of the same video. This allows for smooth resolution switching in adaptive bitrate streaming without requiring a full keyframe [2].

Transform Coding and Quantization Processes in AV1

Transform coding is applied to the prediction residuals in AV1 to further reduce spatial correlation within the blocks. AV1 supports a wide range of transform block sizes, including both square and rectangular shapes, ranging from 4x4 up to 64x64 [5]. The rectangular transform sizes include options such as N×N/2, N/2×N, N×N/4, and N/4×N, providing flexibility in adapting to the specific characteristics of the residual signal [5]. AV1 utilizes four primary 1D transform kernels: the Discrete Cosine Transform (DCT), the Asymmetric Discrete Sine

Transform (ADST), Flip ADST, and the Identity Transform (IDTX) [6]. Each transform block can independently choose from 16 2D separable transform kernels formed by combinations of these 1D kernels applied horizontally and vertically [6]. However, for larger transform block sizes (32x32 and above), the set of available kernels is typically reduced [6].

For inter-coded blocks, AV1 employs a recursive transform block partitioning approach, allowing luma blocks to be further subdivided into transform units of multiple sizes (square, 2:1/1:2, and 4:1/1:4 from 4x4 to 64x64) up to two levels [6]. In contrast, intra-coded blocks generally use a uniform transform block size approach with similar size options [6]. Chroma components typically use the largest possible transform size available for a given coding block [6].

Following the transform stage, the transform coefficients are quantized using a Quantization Parameter (QP), which typically ranges from 0 to 255 (though some sources mention 0 to 63) [5]. Quantization is a lossy process that reduces the precision of the transform coefficients, thereby achieving compression. The QP value controls the strength of the quantization, with higher QP values generally leading to greater compression but also more loss of detail and lower visual quality [15]. AV1 features Quantization Matrix Selection (QMS), supporting 15 sets of predefined quantization weighting matrices [2]. These matrices are based on contrast-sensitive functions and are applied according to the quantization level, with higher quantization levels often using flatter matrices [20]. AV1 also allows for per-superblock changes in the quantization parameter, providing a mechanism for sub-frame rate control and region-of-interest based rate control [6].

Entropy Coding Methods Employed by AV1

AV1 adopts a multi-symbol adaptive arithmetic coding method for entropy coding, which differs from the binary Context-Adaptive Binary Arithmetic Coding (CABAC) used in HEVC [2]. This system, often referred to as the AV1 Entropy Coder (EC), can encode multiple symbols at once and handle a wider range of symbol probabilities as it is non-binary [2]. AV1 utilizes customized context models for different types of data, although the context modeling in AV1 is generally considered simpler than that in CABAC [6]. Probability models for syntax elements are updated based on the frequency of the symbols encountered, with the update rate adapting to the symbol's appearance count within a frame [5]. Notably, probability models can also be inherited from reference frames [5]. For coding the quantized transform coefficients, AV1 employs a level map based system [5]. In this approach, the coefficients are decomposed into a series of symbols, such as the sign bit and the magnitude range. The probability model used for encoding these symbols is conditioned on the values of previously coded coefficients within the same transform block, allowing for efficient exploitation of statistical dependencies.

Loop Filtering Techniques for Quality Enhancement

AV1 incorporates three optional in-loop filter stages to enhance the quality of the reconstructed video frame: the Deblocking Filter (DBF), the Constrained Directional Enhancement Filter (CDEF), and the Loop Restoration Filter (LRF) [5]. The Deblocking Filter (DBF) is applied across the boundaries of transform blocks to reduce the blocking artifacts that can arise from the quantization process [5]. The DBF uses Finite Impulse Response (FIR) filters with lengths that depend on the minimum transform block sizes on either side of the boundary [39]. The application of the filter can be conditioned on a flatness metric to prevent the smoothing of actual edges in the video content [39]. AV1 supports up to four filter levels per frame for deblocking, with separate levels for the luma and chroma components, and the filter level can even change on a per-superblock basis [20].

The Constrained Directional Enhancement Filter (CDEF) is a non-linear deringing filter designed to reduce ringing artifacts that can occur around sharp edges due to the quantization of transform coefficients [2]. CDEF operates on 8x8 blocks and identifies the dominant direction of edges within each block by minimizing the squared error [45]. It then applies a non-linear low-pass filter along the identified direction and, to a lesser extent, along directions rotated by 45 degrees [45]. The strengths of the primary and secondary filters used in CDEF are signaled explicitly in the frame header [45]. CDEF effectively combines a constrained low-pass filter with a deringing filter to reduce coding artifacts while preserving image detail [20].

The final in-loop filter in AV1 is the Loop Restoration Filter (LRF), which is applied after CDEF to further enhance the quality of the reconstructed frame by attempting to recover some of the information lost during compression [2]. LRF is applied to Loop Restoration Units (LRUs), which can be of sizes 64x64, 128x128, or 256x256 pixels [39]. For each LRU, the encoder can choose to bypass filtering, apply a 7x7 separable Wiener filter with encoder-determined parameters, or use a self-guided filter with parameters derived from least-square regression [6]. The Wiener filter in AV1 is designed to be separable and symmetric, with the sum of its coefficients constrained to be 1 [46]. The Self-Guided Restoration Filter (SGRPROJ) uses a guide image, often the reconstructed image itself, to apply filtering that is a function of the spatial characteristics (variance) of the immediate neighborhood of the pixel being filtered [46].

Specific Tools and Features Contributing to AV1's High Compression Efficiency

While the Adaptive Loop Filter (ALF) is mentioned in the query, the AV1 specification primarily refers to the Loop Restoration Filter (LRF) as the adaptive loop filtering tool beyond deblocking and CDEF. The LRF is indeed adaptive, as it allows the encoder to select between a Wiener filter, a self-guided filter, or no filtering at all on a per-LRU basis, based on the characteristics of the content [39]. This adaptability enables targeted quality enhancement based on the specific types of artifacts present in different regions of the frame.

The Constrained Directional Enhancement Filter (CDEF), as previously discussed, plays a crucial role in improving visual quality by reducing ringing artifacts along edges without excessive blurring. Its directional nature allows for more precise filtering, contributing to a cleaner and sharper final image, which in turn enhances the perceived compression efficiency.

Film Grain Synthesis (FGS) is another significant feature that contributes to AV1's high compression efficiency [5]. By removing the naturally occurring film grain from the video before encoding and allowing the decoder to synthesize it algorithmically based on parameters signaled in the bitstream, AV1 can achieve substantial bitrate savings, potentially up to 90% for the film grain component [7]. This approach preserves the artistic intent of the content, as film grain is often a deliberate creative choice, without incurring the high bitrate cost of encoding the inherently unpredictable grain pattern directly [5]. FGS in AV1 uses an auto-regressive (AR) model to generate the grain samples at the decoder [5].

Quantization Matrix Selection (QMS), as detailed earlier, allows AV1 to apply different levels of quantization to different frequency components based on their perceptual importance [2]. This perceptual optimization enables AV1 to achieve better visual quality at a given bitrate by preserving the details that the human eye is most sensitive to while more aggressively quantizing less perceptible frequencies.

Intra Block Copy (IBC), as discussed in the context of intra prediction, provides significant efficiency gains for specific types of content, particularly screen content with repeated textures and patterns [5]. By allowing a block to be predicted from a previously decoded block within the same frame, IBC can drastically reduce the prediction error and the number of bits required to encode such content.

Comparison and Contrast with Previous Generation Codecs (H.264 and HEVC)

AV1 exhibits several key differences in its compression techniques compared to H.264 (AVC). H.264 utilizes a smaller maximum macroblock size of 16x16 pixels, limiting its ability to efficiently code large uniform regions compared to AV1's superblocks which can be up to 128x128 [2]. AV1 also features a significantly expanded set of intra prediction modes (56+) compared to the more limited set in H.264, allowing for more accurate spatial prediction [2]. In inter prediction, AV1 incorporates more advanced tools like affine motion compensation and compound prediction, whereas H.264 primarily relies on translational motion compensation [2]. Furthermore, AV1 offers more flexible transform block sizes and a richer set of transform kernels. For entropy coding, AV1 employs multi-symbol arithmetic coding, while H.264 uses either CABAC or CAVLC [6]. Finally, AV1 includes more sophisticated in-loop filtering techniques, such as CDEF and Loop Restoration, in addition to a deblocking filter, offering more comprehensive artifact reduction compared to H.264's deblocking filter alone [16].

Compared to HEVC (H.265), AV1 continues the trend of increasing complexity for improved efficiency. AV1's maximum superblock size of 128x128 is larger than HEVC's CTUs which go up to 64x64 [2]. AV1 further expands the number of available intra prediction modes compared to HEVC's already substantial set (33-35 modes) [2]. AV1 also allows for a more extensive use of reference frames in inter prediction [5]. Notably, AV1 introduces tools like Intra Block Copy (IBC) and Palette Mode, which are not standard features in HEVC [6]. While both codecs use advanced entropy coding, AV1 employs multi-symbol arithmetic coding as opposed to HEVC's more complex CABAC [6]. Both AV1 and HEVC include a deblocking filter, but AV1 also adds CDEF and Loop Restoration filters, similar to HEVC's Sample Adaptive Offset (SAO) and Adaptive Loop Filter (ALF) [6].

Feature	H.264 (AVC)	HEVC (H.265)	AV1

Max Superblock/CTU Size	16x16	64x64	128x128
Intra Prediction Modes	Limited	More (33-35)	Significantly More (56+)
Inter Prediction	Basic	Advanced	Highly Advanced
Motion Compensation	Translational	Advanced	Warped, Affine, OBMC
Reference Frames	Limited	More	Extensive (Up to 7)
Entropy Coding	CAVLC/CABAC	CABAC	Multi-Symbol Arithmetic
In-Loop Filters	Deblocking	Deblocking, SAO, ALF	Deblocking, CDEF, Loop Restoration
Specialized Modes	None	None	IBC, Palette Mode
Royalty Status	Encumbered	Encumbered	Royalty-Free

Analysis of Compression Performance Benchmarks

Numerous studies and reports have benchmarked the compression performance of AV1 against other codecs, including H.264, VP9, and HEVC. Netflix's internal testing indicated that AV1 achieved approximately 25% greater compression efficiency compared to VP9 [8]. Facebook's evaluations demonstrated even more significant gains, with the AV1 reference encoder achieving 34% to 50% higher data compression than VP9 and x264 [2]. Moscow State University's codec comparison studies have consistently shown AV1 delivering superior visual quality at lower bitrates compared to both HEVC and VP9 [11]. Bitmovin also concluded that AV1 offers superior video quality compared to HEVC and VP9 when played at the same bitrate settings [19].

In terms of bitrate savings, various sources report that AV1 can provide around 30% bitrate reduction compared to HEVC for equivalent visual quality [3]. Some studies have even shown more substantial savings, with AV1 outperforming HEVC by as much as 43.9% at UHD 2160p resolution and by 37.81% at 720p [19]. A 2020 study by the University of Waterloo also corroborated these findings, demonstrating bitrate savings for AV1 over HEVC and VP9 for 4K video [2]. A more recent report from Moscow State University in 2023 indicated that AV1 could provide the same apparent visual quality as HEVC while consuming 12% less data [19]. These benchmark results, using both objective metrics like PSNR, VMAF, and SSIM, as well as subjective evaluations, consistently point to the superior compression performance of AV1.

Computational Complexity of AV1 Encoding and Decoding

While AV1 offers significant advantages in compression efficiency, it generally comes at the cost of increased computational complexity, particularly for encoding [3]. Early implementations of AV1 encoding were notably slower than those of older codecs, with some tests showing encoding times being 4 to 10 times longer than VP9 [1]. AV1 decoding can also be more resource-intensive compared to codecs like H.264 and HEVC [2].

To address these computational demands, various strategies have been employed to optimize the performance of AV1 encoders and decoders. One key technique is the use of tiles, which allows for parallel processing of different parts of a frame, significantly speeding up both encoding and decoding, especially on multi-core processors [2]. The development of optimized encoder implementations like SVT-AV1 (Scalable Video Technology for AV1) by Intel and AOMedia has focused on achieving high performance and scalability on multi-core systems, making AV1 more practical for real-world applications [11]. Similarly, software decoders like dav1d, developed by VideoLAN and FFmpeg, have undergone extensive optimizations for various CPU architectures (x86, ARM, etc.), improving decoding speed and reducing resource consumption [9]. Furthermore, hardware acceleration for AV1 encoding and decoding is becoming increasingly prevalent in newer GPUs and devices, which offloads the computationally intensive tasks from the CPU, leading to significant performance improvements [2].

AV1 encoders often provide different speed/quality presets, allowing users to balance the trade-off between compression efficiency and computational complexity based on their specific needs and the available processing resources [54]. For real-time applications like video conferencing, faster presets with lower complexity might be preferred, while for video-on-demand services or archival purposes, slower presets that maximize compression efficiency might be more suitable, even if they require longer processing times.

AV1 Profiles and Levels

The AV1 specification defines three profiles: Main (Profile 0), High (Profile 1), and Professional (Profile 2) [2]. The Main profile supports 8 or 10 bits per sample and 4:0:0 (monochrome) or 4:2:0 chroma subsampling. It is intended for general-purpose applications. The High profile builds upon the Main profile by adding support for 4:4:4 chroma subsampling, catering to applications requiring higher color fidelity. The Professional profile extends the capabilities further, supporting 8, 10, and 12 bits per sample, as well as all chroma subsampling formats (4:0:0, 4:2:0, 4:2:2, and 4:4:4), making it suitable for professional video production and high-end applications.

AV1 also defines a series of levels, ranging from 2.0 to 6.3, which specify the maximum capabilities of an AV1 decoder in terms of picture size, display rate, decode rate, and bitrate [2]. These levels ensure interoperability by setting performance limits for AV1 streams. For example, Level 4.0 supports resolutions up to 1920x1080 at 30 frames per second with a maximum bitrate of 12 Mbps for the Main profile and 30 Mbps for the High profile [2]. Higher levels, such as Level 5.1, support resolutions up to 3840x2160 at 60 frames per second with higher bitrates [2]. The highest currently defined levels, like 6.2 and 6.3, can support resolutions up to 7680x4320 at 120 frames per second, requiring significantly higher bitrates [2]. These defined profiles and levels allow content creators to target specific device capabilities and network conditions, ensuring a consistent decoding experience.

seq_level_idx	**Level**	**MaxPicSize (Samples**	**MaxHSize (Samples**	**MaxVSize (Samples**	**MaxDisplayRate (Hz)**	**MainMbps (Mbit/s)**	**HighMbps (Mbit/s)**	**Example**

		)	)	)				
0	2.0	147456	2048	1152	4,423,680	1.5	-	426×240 @30fps
1	2.1	278784	2816	1584	8,363,520	3.0	-	640×360 @30fps
4	3.0	665856	4352	2448	19,975,680	6.0	-	854×480 @30fps
5	3.1	1065024	5504	3096	31,950,720	10.0	-	1280×720 @30fps
8	4.0	2359296	6144	3456	70,778,880	12.0	30.0	1920×1080@30fps
9	4.1	2359296	6144	3456	141,557,760	20.0	50.0	1920×1080@60fps
12	5.0	8912896	8192	4352	267,386,880	30.0	100.0	3840×2160@30fps
13	5.1	8912896	8192	4352	534,773,760	40.0	160.0	3840×2160@60fps
16	6.0	35651584	16384	8704	1,069,547,520	60.0	240.0	7680×4320@30fps
17	6.1	35651584	16384	8704	2,139,095,040	100.0	480.0	7680×4320@60fps
18	6.2	35651584	16384	8704	4,278,190,080	160.0	800.0	7680×4320@120fps

Standardization Process of AV1

The AV1 video codec was developed under the auspices of the Alliance for Open Media (AOMedia), a consortium founded in 2015 with the goal of creating open, royalty-free technologies for multimedia delivery [2]. AOMedia comprises a wide range of industry leaders, including semiconductor firms, video-on-demand providers, content producers, software developers, and web browser vendors. AV1 was developed as a successor to Google's VP9 codec and was intended to compete with the highly efficient but patent-encumbered HEVC standard [2].

The initial version of the AV1 bitstream specification was released on March 28, 2018 [2]. This was followed by the release of a validated version 1.0.0 in June 2018 and a version 1.0.0 with Errata 1 in January 2019 [2]. The collaborative and open nature of AOMedia's standardization process aims to foster innovation and encourage widespread adoption of AV1 by removing the barriers associated with proprietary codecs. The involvement of numerous major technology companies ensures that AV1 is designed to meet the diverse needs of the industry.

Conclusion

The AV1 video codec achieves its high compression efficiency through a comprehensive suite of advanced coding tools and architectural features. Its flexible block partitioning, ranging from large superblocks down to small coding units, allows for efficient adaptation to various content complexities. The extensive set of intra and inter prediction techniques, including finer directional modes, smooth predictors, recursive filtering, chroma from luma prediction, intra block copy, palette mode, affine and warped motion compensation, compound prediction, and the use of multiple reference frames, all contribute to reducing spatial and temporal redundancy with greater accuracy than previous generation codecs. AV1's versatile transform coding, with its range of block sizes and selectable kernels, further compacts the energy of the prediction residuals. The quantization process, enhanced by quantization matrix selection, allows for perceptually optimized bitrate reduction. Finally, AV1's entropy coding, using a multi-symbol arithmetic coder, efficiently encodes the quantized coefficients and other signaling information. The in-loop filtering techniques, including the deblocking filter, constrained directional enhancement filter, and loop restoration filter, work to improve the visual quality of the decoded video by reducing various coding artifacts. Unique features like Intra Block Copy for screen content and Film Grain Synthesis for preserving artistic intent further enhance AV1's capabilities.

Compared to previous generation codecs like H.264 and HEVC, AV1 incorporates numerous advancements and innovations across its entire coding pipeline. It utilizes larger block sizes, offers a significantly greater number of prediction modes, employs more sophisticated motion compensation techniques, provides more flexible transform options, and features advanced in-loop filtering. Notably, AV1's royalty-free licensing model stands in stark contrast to the patent-encumbered nature of H.264 and HEVC, making it a particularly attractive option for widespread adoption.

Benchmark studies consistently demonstrate that AV1 offers superior compression efficiency, achieving comparable or better visual quality at significantly lower bitrates than its predecessors. While the computational complexity of AV1 encoding and decoding is generally higher, ongoing optimizations in software implementations like SVT-AV1 and dav1d, coupled with increasing hardware support, are making AV1 more practical for a wide range of applications. The defined profiles and levels within the AV1 specification ensure interoperability and allow for targeting specific device capabilities and performance requirements. The collaborative and open standardization process under AOMedia further strengthens AV1's position as a key codec for the future of internet video. With its strong technical foundation and growing industry support, AV1 holds significant promise to become a dominant video codec, enabling high-quality video delivery across diverse platforms and devices without the burden of licensing fees.

Works cited

1. Complexity and compression efficiency analysis of libaom AV1 video codec - PMC, accessed March 22, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC10161165/
2. AV1 - Wikipedia, accessed March 22, 2025, https://en.wikipedia.org/wiki/AV1
3. AV1 - The Powerful Next Generation Video Codec - Bitmovin, accessed March 22, 2025, https://bitmovin.com/av1/
4. AV1 Codec Explained: The Path Forward for Online Video - Flussonic, accessed March 22, 2025, https://flussonic.com/blog/news/av1-codec-explained/
5. A Technical Overview of AV1 - arXiv, accessed March 22, 2025,

https://arxiv.org/pdf/2008.06091
6. An Overview of Core Coding Tools in the AV1 Video Codec - Jean-Marc Valin, accessed March 22, 2025, https://www.jmvalin.ca/papers/AV1_tools.pdf
7. AV1 Codec: The Future of Video Compression - Ateme, accessed March 22, 2025, https://www.ateme.com/av1-codec-the-future-of-video-compression/
8. AV1 at Netflix: Redefining Video Encoding for a New Era of Streaming, accessed March 22, 2025, https://aomedia.org/av1-adoption-showcase/netflix-story/
9. The State of AV1 Playback Support [2024 Update] - Bitmovin, accessed March 22, 2025, https://bitmovin.com/blog/av1-playback-support/
10. Meta Backs AV1 for Better User Engagement Results | Alliance for Open Media, accessed March 22, 2025, https://aomedia.org/av1-adoption-showcase/meta-story/
11. AV1 Codec - Complete guide for video application devs - ImageKit, accessed March 22, 2025, https://imagekit.io/blog/av1-codec/
12. What is the AV1 encoding codec and what do we use it for? - Bunny.net, accessed March 22, 2025, https://bunny.net/academy/streaming/what-is-av1-codec-and-what-is-it-used-for/
13. Why AV1 Matters – The Game-Changer in Video Streaming - Visionular, accessed March 22, 2025, https://visionular.ai/the-game-changer-in-video-streaming-why-av1-matters/
14. AV1 Video Encoding (AOMedia Video 1) - Library of Congress, accessed March 22, 2025, https://www.loc.gov/preservation/digital/formats/fdd/fdd000541.shtml?loclr=blogsig
15. Performance Analysis of AV1 video codec - Research India Publications, accessed March 22, 2025, https://www.ripublication.com/ijaer23/ijaerv18n3_03.pdf
16. AV1 is a Copy Cat Codec - Sisvel, accessed March 22, 2025, https://www.sisvel.com/insights/av1-is-a-copy-cat-codec/
17. AV1 decoder model - Andrey Norkin, accessed March 22, 2025, https://norkin.org/research/av1_decoder_model/index.html
18. av1-spec/06.bitstream.syntax.md at master · AOMediaCodec/av1 ..., accessed March 22, 2025, https://github.com/AOMediaCodec/av1-spec/blob/master/06.bitstream.syntax.md
19. AV1 vs HEVC: Which Codec is Best for You? - Gumlet, accessed March 22, 2025, https://www.gumlet.com/learn/av1-vs-hevc/
20. Technical-Overview-Of-AV1-Spec/README.md at master - GitHub, accessed March 22, 2025, https://github.com/QuPengfei/Technical-Overview-Of-AV1-Spec/blob/master/README.md
21. AV1 Encoding: Opening a new era of open source video encoding - Tencent MPS, accessed March 22, 2025, https://mps.live/blog/details/what-is-av1
22. What is AV1? - Murat Demirci, accessed March 22, 2025, https://muratdemirci.com.tr/en/av1/
23. arxiv.org, accessed March 22, 2025,

https://arxiv.org/pdf/2008.06091#:~:text=A%20superblock%20is%20the%20largest,own%20prediction%20and%20transform%20modes.
24. Parallel encoding to speed up AV1 encoders or x265 [Archive] - Doom9's Forum, accessed March 22, 2025, http://forum.doom9.org/archive/index.php/t-185071.html
25. imagekit.io, accessed March 22, 2025, https://imagekit.io/blog/av1-codec/#:~:text=Intra%2Dframe%20prediction&text=Rather%20than%20looking%20at%20other,pixels%20based%20on%20specified%20angles.
26. Advancements in intra prediction techniques beyond AV1 codec - SPIE Digital Library, accessed March 22, 2025, https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12674/126740K/Advancements-in-intra-prediction-techniques-beyond-AV1-codec/10.1117/12.2681783.full
27. Advancements in intra prediction techniques beyond AV1 codec - ResearchGate, accessed March 22, 2025, https://www.researchgate.net/publication/374444372_Advancements_in_intra_prediction_techniques_beyond_AV1_codec
28. ADVANCEMENTS IN INTRA PREDICTION TECHNIQUES BEYOND AV1 CODEC - SPIE Digital Library, accessed March 22, 2025, https://www.spiedigitallibrary.org/proceedings/Download?urlId=10.1117%2F12.2681783
29. Advancements in intra prediction techniques beyond AV1 codec - NASA ADS, accessed March 22, 2025, https://ui.adsabs.harvard.edu/abs/2023SPIE12674E..0KY/abstract
30. Decoding the Buzz Over AV1 Codec | Ittiam Blog, accessed March 22, 2025, https://www.ittiam.com/decoding-buzz-av1-codec/
31. AV1: Improving RTC Video Quality at Scale, accessed March 22, 2025, https://atscaleconference.com/av1-improving-rtc-video-quality-at-scale/
32. Understanding Video Inter-Frame Compression Techniques - FastPix, accessed March 22, 2025, https://www.fastpix.io/blog/understanding-video-inter-frame-compression
33. Hardware implementation of AV1 fraction motion estimation - SPIE Digital Library, accessed March 22, 2025, https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13545/1354508/Hardware-implementation-of-AV1-fraction-motion-estimation/10.1117/12.3060028.full
34. AV2 vs AV1: Next-Gen Video Codec Comparison - FastPix, accessed March 22, 2025, https://www.fastpix.io/blog/av2-vs-av1-a-comprehensive-comparison-of-next-gen-video-codecs
35. rzkn/SVT-AV1 - ALTREF Pictures - Gitee, accessed March 22, 2025, https://gitee.com/rzkn/SVT-AV1/blob/master/Docs/Appendix-Alt-Refs.md
36. Advanced transform coding methods for next-generation video coding beyond AV1 - SPIE Digital Library, accessed March 22, 2025,

https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13137/1313717/Advanced-transform-coding-methods-for-next-generation-video-coding-beyond/10.1117/12.3033114.full
37. muratdemirci.com.tr, accessed March 22, 2025, http://muratdemirci.com.tr/en/av1/#:~:text=AV1%20adopts%20a%20different%20entropy,coder%2C%20not%20a%20binary%20one.
38. AV1 In-loop Filters: DBF, CDEF, and SLRF. SLRF is the focus of this work and it includes SSNWF and DSGF. - ResearchGate, accessed March 22, 2025, https://www.researchgate.net/figure/AV1-In-loop-Filters-DBF-CDEF-and-SLRF-SLRF-is-the-focus-of-this-work-and-it-includes_fig1_371632753
39. A Method for In-Loop Video Coding Restoration - MDPI, accessed March 22, 2025, https://www.mdpi.com/2079-9292/13/12/2422
40. Generalized deblocking filter for AVM - Andrey Norkin, accessed March 22, 2025, https://norkin.org/pdf/PCS_2022_Deblock_AV2.pdf
41. Deblocking Loop Filter Appendix - AliveTeam/SVT-AV1 - GitHub, accessed March 22, 2025, https://github.com/AliveTeam/SVT-AV1/blob/master/Docs/Appendix-DLF.md
42. AV1 allows three optional in-loop filter stages including a deblocking... - ResearchGate, accessed March 22, 2025, https://www.researchgate.net/figure/AV1-allows-three-optional-in-loop-filter-stages-including-a-deblocking-filter-a-CDEF_fig23_349650510
43. Did you know that SVT-AV1 uses linear models to determine deblocking levels? - Medium, accessed March 22, 2025, https://medium.com/@luc.trudeau/did-you-know-that-svt-av1-uses-linear-models-to-determine-deblocking-levels-5294beed12d0
44. Docs/svt-av1-encoder-design.md · master - GitLab, accessed March 22, 2025, https://gitlab.com/AOMediaCodec/SVT-AV1/-/blob/master/Docs/svt-av1-encoder-design.md
45. THE AV1 CONSTRAINED DIRECTIONAL ENHANCEMENT FILTER (CDEF) Steinar Midtskogen Cisco Systems Inc. Lysaker, Norway - Jean-Marc Valin, accessed March 22, 2025, https://www.jmvalin.ca/papers/cdef_icassp2018.pdf
46. Docs/Appendix-Restoration-Filter.md · master - SVT-AV1 - GitLab, accessed March 22, 2025, https://gitlab.com/AOMediaCodec/SVT-AV1/-/blob/master/Docs/Appendix-Restoration-Filter.md
47. High-Throughput Hardware Design for the AV1 Decoder Switchable Loop Restoration Filters | Journal of Integrated Circuits and Systems, accessed March 22, 2025, https://jics.org.br/ojs/index.php/JICS/article/view/667
48. Help me understand SVT-AV1 parameters - Reddit, accessed March 22, 2025, https://www.reddit.com/r/AV1/comments/s7yyf9/help_me_understand_svtav1_parameters/
49. D3D12DDI_VIDEO_ENCODER_, accessed March 22, 2025, https://learn.microsoft.com/en-us/windows-hardware/drivers/ddi/d3d12umddi/ne-d3d12umddi-d3d12ddi_video_encoder_av1_restoration_type_0095
50. Netflix Is Now Streaming High-Quality AV1 Video to Select Smart TVs and Game

Consoles, accessed March 22, 2025, https://www.audioholics.com/hdtv-formats/netflix-av1
51. Bringing AV1 Streaming to Netflix Members' TVs, accessed March 22, 2025, https://netflixtechblog.com/bringing-av1-streaming-to-netflix-members-tvs-b7fc88e42320
52. Facebook: Tests Show AV1 Streaming Performance is Exceeding Expectations - Next TV, accessed March 22, 2025, https://www.nexttv.com/blog/facebook-tests-show-av1-streaming-performance-exceeding-expectations
53. AV1 beats x264 and libvpx-vp9 in practical use case - Engineering at Meta - Facebook, accessed March 22, 2025, https://engineering.fb.com/2018/04/10/video-engineering/av1-beats-x264-and-libvpx-vp9-in-practical-use-case/
54. How Meta brought AV1 to Reels - Engineering at Meta - Facebook, accessed March 22, 2025, https://engineering.fb.com/2023/02/21/video-engineering/av1-codec-facebook-instagram-reels/
55. AMD AVC vs HEVC vs AV1 - Which Recording/Streaming Codec should you use?, accessed March 22, 2025, https://www.youtube.com/watch?v=FzZKyXHP_d0
56. MSU Codec Study Full HD 2020 Results - Visionular, accessed March 22, 2025, https://visionular.ai/msu-codec-study-aurora1-av1-encoder-results-2020/
57. 3 Takeaways From Moscow State University's 2019 Codec Comparison - Streaming Media, accessed March 22, 2025, https://www.streamingmedia.com/Articles/Editorial/Featured-Articles/3-Takeaways-From-Moscow-State-Universitys-2019-Codec-Comparison-135294.aspx
58. AV1 Beats VP9 and HEVC on Quality, if You've Got Time, says Moscow State - Reddit, accessed March 22, 2025, https://www.reddit.com/r/hardware/comments/9t5sps/av1_beats_vp9_and_hevc_on_quality_if_youve_got/
59. The Visionular AV1 Encoder is Optimized for All Use Cases | Alliance for Open Media, accessed March 22, 2025, https://aomedia.org/av1%20adoption/visionular-av1-encoder-optimized/
60. Performance Comparison of VVC, AV1 and HEVC on 8-bit and 10-bit content - FastVDO, accessed March 22, 2025, https://fastvdo.com/pubs/spie/spie2018/perfCompAV1-spie2018.pdf
61. Build and Run the AV1 codec - Arm Learning Paths, accessed March 22, 2025, https://learn.arm.com/learning-paths/servers-and-cloud-computing/codec1/libaom/
62. AV1 for Dummies - Codec Wiki, accessed March 22, 2025, https://wiki.x266.mov/blog/av1-for-dummies
63. SVT-AV1: open-source AV1 encoder and decoder | by Netflix Technology Blog, accessed March 22, 2025, https://netflixtechblog.com/svt-av1-an-open-source-av1-encoder-and-decoder-ad295d9b5ca2
64. Fast AV1 Encoding Is Finally Here - Visionular, accessed March 22, 2025,

https://visionular.ai/fast-av1-encoding-is-finally-here/
65. SVT-AV1 3.0 Released With Faster CPU-Based AV1 Encoding - Phoronix, accessed March 22, 2025, https://www.phoronix.com/news/SVT-AV1-3.0-Released
66. psy-ex/svt-av1-psy: The Scalable Video Technology for AV1 (SVT-AV1 Encoder and Decoder) with perceptual enhancements for psychovisually optimal AV1 encoding - GitHub, accessed March 22, 2025, https://github.com/psy-ex/svt-av1-psy
67. Docs/CommonQuestions.md · master · Alliance for Open Media / SVT-AV1 - GitLab, accessed March 22, 2025, https://gitlab.com/AOMediaCodec/SVT-AV1/-/blob/master/Docs/CommonQuestions.md
68. SVT-AV1 Encoding Guide - GitHub Gist, accessed March 22, 2025, https://gist.github.com/dvaupel/716598fc9e7c2d436b54ae00f7a34b95
69. Aomedia Releases Optimized Av1 Encoder Svt Av1 Ver_2_0 | Alliance for Open Media, accessed March 22, 2025, https://aomedia.org/blog%20posts/aomedia-releases-optimized-av1-encoder-svt-av1-ver_2_0/
70. AV1 Encoder Optimization from the Decoder's Point of View - Visionular, accessed March 22, 2025, https://visionular.ai/av1-encoder-optimization-from-the-decoders-point-of-view/
71. How to improve AV1 software decoding speed to the max: an updated 2022 decoding and encoding guide for the future! - Reddit, accessed March 22, 2025, https://www.reddit.com/r/AV1/comments/v2kguu/how_to_improve_av1_software_decoding_speed_to_the/
72. Dav1d - Christian Külker, accessed March 22, 2025, https://christian.kuelker.info/en_US/Quick-Guide/DevOps/Video/dav1d.html
73. Optimizing media pipelines using Armv8.x and Armv9.x features - Arm Developer, accessed March 22, 2025, https://developer.arm.com/documentation/110065/0100/Optimization-case-studies/libdav1d
74. Releases · VideoLAN / dav1d - GitLab, accessed March 22, 2025, https://code.videolan.org/videolan/dav1d/-/releases
75. Dav1d 1.4.2 Provides More AVX2 & AVX-512 Performance Optimizations - Phoronix, accessed March 22, 2025, https://www.phoronix.com/news/dav1d-1.4.2-AV1-Decoder
76. dav1d 1.3 Yields Lower Memory Use, More Arm NEON Optimizations - Phoronix, accessed March 22, 2025, https://www.phoronix.com/news/dav1d-1.3-Released
77. Improving Video Quality and Performance with AV1 and NVIDIA Ada Lovelace Architecture, accessed March 22, 2025, https://developer.nvidia.com/blog/improving-video-quality-and-performance-with-av1-and-nvidia-ada-lovelace-architecture/
78. My general suggestions for AV1 encoding settings - Reddit, accessed March 22, 2025, https://www.reddit.com/r/AV1/comments/1fhzvhb/my_general_suggestions_for_av1_encoding_settings/
79. How AV1 video encoding is transforming network video for the future | Axis

Communications, accessed March 22, 2025, https://newsroom.axis.com/en-us/blog/av1-codec-video-surveillance
80. How the AV1 Codec Compresses Video for Efficient and High-Quality Delivery - Cloudinary, accessed March 22, 2025, https://cloudinary.com/guides/video-formats/open-source-and-royalty-free-av1-compresses-video-efficiently-and-effectively
81. AV1 Video Codec | Alliance for Open Media, accessed March 22, 2025, https://aomedia.org/specifications/av1/