Video Compression: Trends, Applications, and Innovations

Advanced Topics - Bo Fu

Abstract

The explosive growth of video data has made compression, encoding, and decoding critical for efficient storage, transmission, and processing. These technologies are widely applied in streaming and real-time communication. This paper examines industry trends and needs, reviews current solutions with their advantages and limitations, and proposes a context-aware adaptive encoding method to address existing challenges while meeting future demands.

1. Introduction

In recent years, global video traffic has grown exponentially. According to Statista, online video traffic is expected to account for over 80% of global internet traffic by 2025 [1]. This trend has driven the demand for efficient video compression technologies [2]. Platforms like Netflix and YouTube rely on advanced compression to enhance user experience and reduce operational costs. Additionally, real-time communication software such as Zoom depends on low-latency compression technology, while some advanced applications such as autonomous driving and telemedicine have put forward even higher requirements. This paper explores these industry demands, evaluates current implementations, and discusses potential improvements.

1.1 Current Needs

With the proliferation of video technology, various industries have specific requirements for compression and encoding:

High Definition and Visual Quality

The growing adoption of UHD formats like 4K and 8K has raised the bar for visual quality expectations. High-definition video with vibrant colors, sharp details, and minimal artifacts is essential for streaming services, cinematic productions, and VR/AR experiences. Lossless or visually lossless compression becomes crucial for industries like healthcare and scientific research, where even minor quality degradation could lead to misinterpretations.

High Compression Rate

The exponential growth of video content, including ultra-high-definition (UHD) formats such as 4K and 8K, has made high compression rates a top priority for many industries. Streaming platforms like Netflix, YouTube, and Amazon Prime require efficient compression to minimize bandwidth usage while maintaining video quality. For instance, a video broadcast requires approximately 32 Mbps using AVC, but newer standards like HEVC reduce this to around 15 Mbps with comparable quality [3]. High compression rates also help reduce storage costs, especially in cloud environments where video archives grow rapidly.

Low Latency

Low latency is vital in real-time communication services like Zoom, cloud gaming, and telemedicine platforms. These services demand encoding and decoding with minimal delays to ensure smooth interactions. For example, as a highly interactive multimedia application, cloud gaming requires reliable and low-latency network communication [4]. Similarly, latency becomes critical in telemedicine, where delays can affect the performacne of the telemedicine.

1.2 Industry Trends

Emergence of Advanced Standards

Video compression standards are evolving rapidly to meet modern demands by employing more adaptability, more refined temporal and spatial prediction models with better distortion metrics [5]. H.266/VVC and AV1 are the latest standards designed to deliver significant improvements in compression efficiency. For instance, H.266 claims a 50% bitrate reduction compared to H.265 for the same visual quality, making it ideal for UHD and VR content [6]. AV1, developed by the Alliance for Open Media, offers comparable efficiency but is royalty-free, making it attractive for companies seeking cost-effective solutions [7].

AI-Driven Compression

The development of neural networks (learning-based), especially convolutional neural networks (CNN) has risen the attention of researchers in the compression field [2]. In the past few years, neural video codec has impoved the performance of video compression by changing extraction and utilization of context. For example, a new model DCVC-DC which efficiently utilizes the Diverse Contexts to further boost compression ratio outperforms traditional codec in both RGB and YUV420 colorspaces [8].

2. Current Solutions

2.1 AVC (Advanced Video Coding)

AVC, or H.264, remains one of the most popular and widely supported video codecs. It offers a balance between compression efficiency and computational requirements, making it suitable for a broad range of devices and applications. However, it struggles with the growing demand for ultra-high-definition video and modern applications requiring higher efficiency [3].

2.2 HEVC (High Efficiency Video Coding)

HEVC, also known as H.265, is a widely adopted video compression standard developed to replace AVC (H.264). It achieves approximately 50% better compression efficiency compared to AVC, making it ideal for 4K and HDR content. HEVC is used extensively in streaming platforms like Netflix and broadcasting services. Despite its efficiency, HEVC is associated with high computational complexity and licensing costs, which limits its adoption in open-source projects [3].

2.3 VVC (Versatile Video Coding)

VVC, also known as H.266, is the latest standard succeeding HEVC. It offers a 50% bitrate reduction over HEVC with comparable quality, making it highly efficient for applications like 8K video, virtual reality (VR), and cloud gaming. However, VVC's adoption faces challenges due to its even higher computational complexity and licensing uncertainties [6].

2.4 VP9

VP9, developed by Google, is a royalty-free alternative to HEVC. It offers comparable efficiency to HEVC and is supported on platforms like YouTube [3]. VP9 is particularly favored in open-source projects due to its free licensing. However, its compression efficiency and hardware support lag slightly behind newer standards like AV1 [7].

2.5 AV1

AV1, developed by the Alliance for Open Media (AOMedia), is an open-source codec designed to provide high compression efficiency with no licensing fees. It outperforms both HEVC and VP9, offering up to 30% better compression efficiency. However, AV1 encoding is computationally expensive, requiring hardware acceleration for real-time performance. Its adoption is growing rapidly, with support from major tech companies like Netflix, YouTube, and Microsoft [7].

2.6 DL-Based Video Compression

With the extensive use of Artificial Intelligence, AI, such as Deep Learning (DL) have emerged as a best-of-breed alternative for performing different tasks have also been used in the option of improving video compression in the last years, with the primary objective of reducing compression ratio while preserving the same video quality [9][10].

3. Critical Analysis

3.1 AVC

- **Pros**:
  - Low computational requirements and widespread support.
  - Ideal for standard definition (SD) and high definition (HD) content.
  - Mature ecosystem with extensive hardware and software support.
- **Cons**:
  - Lacks efficiency for UHD and modern formats.
  - Aging standard compared to HEVC and AV1.

3.2 HEVC

- **Pros**:
  - High compression efficiency (50% better than AVC).
  - Widespread adoption in streaming and broadcasting.
  - Supports advanced features like 4K and HDR.
- **Cons**:
  - High computational complexity.
  - Expensive licensing fees limit open-source use.
  - Requires hardware acceleration for real-time encoding.

3.3 VVC

- **Pros**:
  - Exceptional compression efficiency (50% better than HEVC).
  - Designed for 8K, VR, and other emerging applications.
- **Cons**:
  - Extremely high computational and memory requirements.
  - Expensive licensing fees limit open-source use.
  - Uncertainty around licensing limits adoption.

3.4 VP9

- **Pros**:
  - Royalty-free licensing makes it attractive for open-source projects.
  - Comparable efficiency to HEVC for HD and 4K content.
- **Cons**:
  - Limited hardware support compared to HEVC.
  - Marginally outperformed by AV1 in compression efficiency.

3.5 AV1

- **Pros**:
  - Royalty-free, with 30% better efficiency than HEVC.
  - Supported by major platforms and devices.
  - Strong adoption momentum in streaming and cloud applications.
- **Cons**:
  - Encoding is computationally expensive and requires dedicated hardware.
  - Slower adoption in real-time applications due to processing overhead.

3.6 DL-based Video Compression

- **Pros**:
  - Potential for breakthroughs in specific scenarios [8].
  - Highly content adaptivity to complex scenes and textures [9].
  - Could improve coding efficiency by leveraging samples from far distance [9].
- **Cons**:
  - Computationally expensive for real-time applications.
  - Requires significant computational resources for training [11].
  - The performance is still limited now [9].

4. Proposed Solution

4.1 Hybrid Compression Framework

To address the limitations of current solutions, this paper proposes a **Hybrid Compression Framework** that combines traditional video compression standards (e.g. HEVC, AV1) with AI-driven approaches (e.g. CNNs and DNNs). The framework aims to achieve high compression rates and definition.

4.2 Key Features of the Hybrid Compression Framework

4.2.1. Content-Adaptive Encoding

The framework uses AI-based scene analysis to determine the optimal encoding settings dynamically. For example, scenes with high motion or complex textures are processed using AI-driven encoders (e.g. CNN-based residual modeling), while static or simple scenes use traditional codecs like HEVC or AV1. This hybrid approach ensures that computational resources are allocated efficiently, optimizing compression rates without compromising quality.

4.2.2. Hardware-Aware Deployment

The framework integrates hardware acceleration for both traditional and AI-driven components. For example:

- **Traditional codecs**: Leverage hardware encoders like NVIDIA NVENC or Intel Quick Sync.
- **AI-based components**: Use GPU-optimized libraries like TensorRT or PyTorch for inference acceleration.

By combining these technologies, the solution minimizes latency and maximizes throughput for real-time and offline applications.

4.3 Advantages of the Hybrid Compression Framework

- **Improved Compression Efficiency**: By leveraging AI for complex scenes and traditional codecs for simpler ones, the framework achieves superior compression rates compared to using either approach alone.
- **Low Latency**: Hardware acceleration and content-adaptive encoding reduce latency, making the solution viable for live streaming and real-time applications.

4.4 Challenges

While the proposed framework offers several advantages, it faces challenges:

- **Increased Development Complexity**: Integrating AI models with traditional codecs requires careful optimization and resource management.
- **Training Data and Model Generalization**: AI components rely heavily on high-quality training data, and ensuring generalization across diverse content remains a challenge.
- **Hardware Dependency**: Effective deployment of AI-driven components depends on the availability of hardware accelerators, which may limit adoption in low-resource environments.
- **Overhead of Video Determination**: The AI-based scene analysis may introduce additional overheads, which could increase the consumption and latency of video compression.

5. Citations

1. IDC, & Statista. (June 7, 2021). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 (in zettabytes) [Graph]. In Statista. Retrieved November 18, 2024, from https://www.statista.com/statistics/871513/worldwide-data-created/
2. Hoang, T. M., & Zhou, J. (2021). Recent trending on learning based video compression: A survey. Cognitive Robotics, 1, 145-158.
3. Joy, H. K., & Kounte, M. R. (2020). A comprehensive review of traditional video processing. Advances in Science, Technology and Engineering Systems Journal, 5(6), 272-279.

4. Alhilal, A., Braud, T., Han, B., & Hui, P. (2022, April). Nebula: Reliable low-latency video transmission for mobile cloud gaming. In Proceedings of the ACM Web Conference 2022 (pp. 3407-3417).
5. Yu, L., & Wang, J. P. (2010). Review of the current and future technologies for video compression. Journal of Zhejiang University SCIENCE C, 11(1), 1-13.
6. Lee, M., Song, H., Park, J., Jeon, B., Kang, J., Kim, J. G., ... & Sim, D. (2023). Overview of versatile video coding (H. 266/VVC) and its coding performance analysis. IEIE Transactions on Smart Processing & Computing, 12(2), 122-154.
7. Liu, Y. (2018). AV1 beats x264 and libvpx-vp9 in practical use case. Engineering at Meta.
8. Li, J., Li, B., & Lu, Y. (2023). Neural video compression with diverse contexts. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 22616-22626).
9. Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., & Wang, S. (2019). Image and video compression with neural networks: A review. IEEE Transactions on Circuits and Systems for Video Technology, 30(6), 1683-1698.
10. Yasin, H. M., & Ameen, S. Y. (2021, December). Review and evaluation of end-to-end video compression with deep-learning. In 2021 International Conference of Modern Trends in Information and Communication Technology Industry (MTICTI) (pp. 1-8). IEEE.
11. Bidwe, R. V., Mishra, S., Patil, S., Shaw, K., Vora, D. R., Kotecha, K., & Zope, B. (2022). Deep learning approaches for video compression: a bibliometric analysis. Big Data and Cognitive Computing, 6(2), 44.