



A Comprehensive Literature Review on Image and Video Compression: Trends, Algorithms, and Techniques

Murooj Khalid Ibraheem Ibraheem^{1*}, Alexander Viktorovich Dvorkovich¹, Israa M. Abdalameer Al-khafaji²

¹ Department of Multimedia Technologies and Telecommunications, Phystech School of Radio Engineering and Computer Technologies (FRKT), Moscow Institute of Physics and Technology (MIPT), Dolgoprudny 141701, Russia

² Department of Corporate Information Systems, Institute of Information Technologies, RTU MIREA, Moscow 119454, Russia

Corresponding Author Email: ibragim.m@phystech.edu

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.290307>

ABSTRACT

Received: 4 December 2023

Revised: 9 April 2024

Accepted: 16 May 2024

Available online: 20 June 2024

Keywords:

image compression, video compression, transform coding, predictive coding, entropy coding, machine learning, deep learning

Compression methods for images and videos are essential for the effective archiving, transmission, and distribution of multimedia data and files. This paper reviews the state-of-the-art in image and video compression, including the most recent developments, algorithms, and methods. This study compiles findings from a variety of studies in an effort to give readers a bird's-eye view of the progress and obstacles in this dynamic sector. A survey of the relevant literature demonstrates that modern compression methods build upon the work of older algorithms like JPEG and MPEG. Compression ratios and picture quality can be enhanced, however, thanks to developments in transform coding, predictive coding, and entropy coding. Further, by combining machine learning and deep learning techniques, we now have access to cutting-edge options for improving compression efficiency and paving the way for adaptive, content-aware compression. Sustainable compression approaches are also highlighted, along with energy efficiency aspects.

1. INTRODUCTION

Image and video compression methods have significantly improved the storage, transmission, and display of visual data in digital communication [1, 2]. Effective compression methods are becoming increasingly important as data volumes increase, particularly in multimedia applications. These compression methods attempt to eliminate superfluous details without negatively impacting the original data's perceived quality [3]. The key to various compression methods is removing unnecessary information from the material being compressed. The goal of the algorithms is to reduce data size without substantially effecting quality of the images or videos [4]. These algorithms achieve their goals by using the features of visual data and the peculiarities of human visual perception. Video compression methods can be thought of as an extension of image compression methods [5]. However, there are extra difficulties because visual data is always changing. In order to save time and space, video compression works to create a more compact copy of the original video. Several methods, including motion estimation, compensation, and temporal prediction, help to accomplish this goal. Using past frames to foretell future ones is the essence of temporal prediction. Because of this, less room is required to save each frame separately [6]. Instead of recording each individual frame, a video stores the difference or 'change' between frames to better forecast the motion of pixels from one frame to the next using a technique called motion estimation. Several standards have been created to promote interoperability and effective video compression [7]. Video compression protocols such as

H.264/AVC, H.265/HEVC, H.266/VVC, AV1 and VP9 have increased both the rate and quality of the process. Intra-frame and inter-frame compression are two of the strategies used to achieve such high compression ratios while maintaining watchable video quality [8].

The integration of machine learning and deep learning techniques in image and video compression algorithms represents a significant advancement in the field, offering innovative solutions and distinct advantages over traditional methods [3]. These technologies are applied in various ways to enhance compression efficiency and quality preservation. Machine learning algorithms excel at feature extraction, automatically identifying relevant information for more efficient compression. Additionally, deep learning models can adapt encoding strategies based on the content, leading to optimized compression and improved quality preservation. By learning patterns and structures within the data, deep learning algorithms enable content-aware compression, ensuring important details are retained while reducing redundancy. Moreover, machine learning models can dynamically adjust compression parameters based on specific data characteristics, striking a better balance between compression ratio and quality [2]. The advantages of employing machine learning and deep learning in compression algorithms are evident. These advanced technologies offer enhanced compression efficiency by learning complex patterns and correlations in data, surpassing the capabilities of traditional methods. Furthermore, the ability of deep learning models to prioritize important features during compression results in higher-quality output with reduced artifacts, improving the overall

visual experience [4]. Their adaptability to diverse types of multimedia data makes machine learning algorithms versatile and effective across various applications and content types, showcasing their superiority over conventional approaches. Additionally, deep learning models excel at optimizing the rate-distortion trade-off in compression, achieving a delicate balance between file size and visual quality that may be challenging for traditional methods. The originality and contributions of integrating machine learning and deep learning in image and video compression algorithms are profound. This innovative approach addresses the challenges of data compression in multimedia applications by pushing the boundaries of compression efficiency and quality preservation. Researchers are leveraging these advanced technologies to develop adaptive and intelligent compression systems that cater to specific user preferences and content requirements, offering personalized compression solutions that were previously unattainable [5]. The incorporation of machine learning and deep learning technologies in compression algorithms represents a transformative shift in the field, promising continued advancements and innovations in multimedia compression.

1.1 Fundamentals of image compression

Understanding the basics of image compression is crucial for grasping these methods. The effect of image compression in the spatial and frequency domains is first analyzed [9]. It explores raster, vector, and fractal methods of picture representation and assesses their compressibility. Image compression techniques rely significantly on human visual perception and psychovisual models to achieve higher compression ratios without noticeable loss [10]. Lossy and lossless compression techniques are discussed in detail, along with well-known algorithms like transform coding (e.g., discrete cosine transform), predictive coding, and vector quantization [11].

1.2 Image compression standards

The creation of image compression standards has considerably assisted the interoperability and broad adoption of image compression techniques. Here, we explore how standard picture compression formats like JPEG, PNG function on a technical level [12]. This review examines the pros and cons of each specification. In order to achieve high compression ratios, one of the most common standards, JPEG, uses lossy compression, making it ideal for compressing photographs and other images of natural surroundings [13]. Images with fine features, such as sharp edges, line drawings, and typography, benefit greatly from PNG's lossless compression method. The new standard, called HEIF, uses state-of-the-art techniques like High Efficiency Video Coding (HEVC) to offer significantly greater compression than JPEG [14]. The compression ratios, image quality, and compatibility of existing standards are compared, and newer standards are discussed as well.

1.3 Video compression fundamentals

Video compression is similar to image compression, but more work is required to effectively capture motion and temporal redundancy [15]. Video compression basics such as temporal and spatial prediction, motion estimation and

compensation, and video coding standards are discussed by Al-Janabi and Al-Shourbaji [16]. H.264/AVC (Advanced Video Coding), H.265/HEVC (High Efficiency Video Coding), and VP9 are only few of the video compression standards discussed here. The efficiency and quality of video compression have been greatly improved thanks to these standards, allowing high-definition video to be streamed and distributed over networks with constrained capacity [17].

1.4 Hybrid image and video compression

In recent years, hybrid approaches have attracted interest for their ability to achieve optimal coding efficiency by integrating image and video compression algorithms [18]. Challenges including key frame extraction, inter-frame prediction, object-based compression, and region-of-interest coding are discussed herein in relation to the hybridization of image and video compression methods. For effective encoding in contexts as diverse as video conferencing, surveillance systems, and virtual reality, hybrid approaches combine the best features of both image and video compression. It is important to balance the need for space with the quality of the final product when selecting crucial frames from video sequences. Second, since images are often viewed as separate entities, it is difficult to include inter-frame prediction techniques into image compression methods. This is because effective methods are needed to anticipate image content based on prior images in the sequence. Object-based compression is essential to video compression but must be tailored to single images; this calls for sophisticated object detection and segmentation techniques. The identification and compression of important areas inside pictures or video frames is essential for the efficient coding of regions of interest (ROIs). Generally, efficient compression while addressing the varied requirements of both photos and videos calls for novel ways made possible by the hybridization of both methodologies.

A comprehensive and up-to-date analysis of the field of image and video compression is what a literature review is meant to accomplish. This article seeks to consolidate previous work and developments in the field of image and video compression. The purpose of the review is to help people make more informed choices, encourage further investigations, and progress this important advancement in a number of different fields by synthesizing existing research, detecting trends and obstacles, and offering a comparative analysis. There are various algorithms, techniques and performance evaluation mentioned in this study. This study covers a range of algorithms and techniques that have been used in the field. One of the most widely known algorithms for image compression is JPEG (Joint Photographic Experts Group), which is commonly used for compressing photographic images. It employs a lossy compression technique and utilizes methods like Discrete Cosine Transform (DCT) and Huffman coding. Another algorithm discussed in the study is PNG (Portable Network Graphics), which differs from JPEG in that it is usually used for lossless compression. PNG is particularly effective for compressing images with fine features such as sharp edges, line drawings, and typography. It employs filtering algorithms and deflate compression to achieve its compression efficiency. The study also explores the use of HEIF (High Efficiency Image Format), a newer standard that utilizes High Efficiency Video Coding (HEVC) techniques. HEIF offers significantly greater compression than JPEG,

providing higher compression rates, improved image quality, and compatibility with existing standards. In the realm of video compression, the study covers standards such as H.264/AVC (Advanced Video Coding) and H.265/HEVC (High Efficiency Video Coding). These standards have greatly improved the efficiency and quality of video compression. They employ techniques like motion estimation, compensation for inter-frame compression to achieve high compression ratios while maintaining watchable video quality. Additionally, the study discusses VP9, a video compression codec developed by Google. VP9 offers efficient compression and is commonly used for streaming high-definition video over networks with limited capacity. Transform coding, including techniques like Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT), is a key aspect of both image and video compression. It exploits the spatial redundancy within frames and allows for more efficient encoding of frequency components. Predictive coding techniques, such as Differential Pulse Code Modulation (DPCM), are used to encode the deviation from the predicted value for each pixel by utilizing information about neighboring pixels. This helps in reducing data redundancy. Entropy coding algorithms, such as Huffman coding and arithmetic coding, are employed to further compress the encoded data by assigning shorter codes to more likely symbols. Overall, the study provides a comprehensive overview of these algorithms and techniques, highlighting their significance and impact in the field of image and video compression.

The study also focuses on the evaluation of performance to assess the effectiveness of various compression algorithms and methods. Several parameters are considered in this evaluation, including:

Compression Ratio: The compression ratio measures how much the image or video has been reduced in size compared to its original dimensions. It is calculated by dividing the uncompressed size by the compressed size. A higher compression ratio indicates more efficient compression.

Image/Video Quality: The quality of the compressed image or video is an important factor to consider. Various metrics, such as Peak Signal-to-Noise Ratio (PSNR), can be used to measure the similarity between the original and compressed versions. Higher PSNR values indicate better quality.

Computational Complexity: The computational complexity of a compression algorithm refers to the amount of computational resources required to perform the compression. This includes factors such as processing time and memory usage. Lower computational complexity is desirable for efficient compression.

Device/Application Compatibility: The compatibility of compressed files with different devices and applications is crucial. Compression methods that produce files compatible with a wide range of devices and applications are preferred.

By evaluating the performance of compression algorithms based on these parameters, the study provides insights into the strengths and weaknesses of each approach. It allows researchers and practitioners to make informed decisions about which algorithms and methods are most suitable for their specific requirements, considering factors such as compression ratio, image/video quality, computational complexity, device/application compatibility, and energy efficiency.

This study has extended the theoretical landscape of image and video compression by empirically testing and comparing

the performance of JPEG, JPEG 2000, and HEVC across various multimedia contexts. One of the key theoretical advancements offered by this research is a deeper understanding of the operational efficiencies of DCT versus wavelet-based compression under varied image quality and resolution conditions. The comparative analysis provides empirical evidence supporting the theoretical efficiency of wavelet transforms in preserving high-quality images at lower bit rates, a theoretical assertion that lacked robust empirical backing until now [10]. The research elucidates the computational load implications of advanced compression standards like HEVC, aligning real-world performance with theoretical expectations. These findings contribute to refining the theoretical models that predict the computational cost versus compression efficiency trade-offs, particularly important for the development of future compression algorithms that are both resource-efficient and high-performing. Practically, this study offers concrete insights that can influence several industries reliant on digital imaging and video broadcasting. For instance, the findings suggest that while JPEG remains a viable option for low-latency applications, its limitations in quality at high compression ratios could be problematic for industries where detail retention is crucial, such as digital photography and online content creation. This directs industry stakeholders toward considering JPEG 2000 or HEVC for high-quality needs. The superior performance of JPEG 2000 in maintaining image integrity at high compression ratios positions it as a preferable option in sectors like healthcare, where medical imaging requires both high precision and efficient storage. This could directly influence the adoption of JPEG 2000 in medical imaging technologies [11].

The study's findings on HEVC's effectiveness in handling 4K video streams at reduced bandwidths have significant implications for the streaming industry. With increasing demand for high-definition video content, streaming services can leverage HEVC to deliver higher quality videos at manageable bandwidths, enhancing user experience and reducing data transmission costs.

2. REVIEW OF RELATED WORK

Image and video compression technologies are crucial for efficiently storing and transmitting visual content by reducing data size. These technologies are grounded in information theory and signal processing, and their development shows a clear progression based on both theoretical advances and practical needs. Information theory, introduced by Claude Shannon, forms the core of data compression strategies. It focuses on concepts like entropy, which measures the unpredictability of information content, aiming to minimize redundancy (predictable and repetitive patterns) and irrelevance to reduce data size effectively. Rate-distortion theory, another pivotal concept from information theory, explores the trade-off between the compression rate (bitrate) and the resulting quality loss (distortion), which is particularly relevant in lossy compression techniques. In signal processing, transform coding is a fundamental method where data, such as images or video frames, is transformed from the time or spatial domain to the frequency domain using mathematical transforms such as the Fourier Transform, Discrete Cosine Transform (DCT), and Wavelet Transform. The transformation allows significant parts of the data to be

separated from less important ones, which can then be compressed by quantizing and discarding less critical frequency coefficients.

The JPEG image compression standard exemplifies these principles by using DCT to focus on areas of an image most important to human perception, compressing other areas more aggressively. Its successor, JPEG 2000, enhances this approach by using wavelet transforms for better scalability and efficiency, particularly effective at higher compression ratios and supporting features like progressive transmission. Video compression technologies, such as those defined by the MPEG standards and the H.264/AVC and H.265/HEVC standards, build on and extend these image compression techniques. They introduce additional efficiencies through temporal compression, using motion estimation and compensation to reduce redundancy across video frames. These methods significantly improve compression efficiency, handling higher data rates and different network conditions more effectively. Advancements in machine learning have begun to be integrated into compression technologies, using algorithms to predict and optimize various encoding parameters dynamically. This evolution is evident in the transition from MPEG-2 to more sophisticated codecs like H.264 and HEVC, which offer significantly higher compression ratios and are better suited to modern applications such as ultra-high-definition video and streaming over variable bandwidth conditions. Understanding the theoretical underpinnings and practical applications of these technologies illustrates the ongoing improvements in compression methods, reflecting continuous advancements in computational capabilities and algorithm efficiency. These developments are essential for meeting the increasing demands of data-intensive applications in our digital age.

Machine learning techniques offer a promising avenue for complementing traditional compression methods like transform and predictive coding. Firstly, machine learning can enhance transform coding by optimizing the selection and application of transform functions based on the characteristics of the input data. Traditional transform coding techniques like DCT or DWT operate on fixed transform functions, which may not always be optimal for diverse types of multimedia content. Machine learning algorithms can analyze the statistical properties of the data and adaptively select or design transform functions that better capture its underlying structure, leading to improved compression efficiency. Secondly, machine learning can augment predictive coding by learning more sophisticated prediction models [12]. Traditional predictive coding methods rely on simple motion estimation and interpolation techniques to predict future frames in video compression. Machine learning algorithms, however, can analyze large amounts of training data to learn complex patterns and relationships, enabling more accurate prediction of future frames. By integrating machine learning into predictive coding, compression algorithms can achieve higher compression ratios while maintaining or even enhancing picture quality [13]. Furthermore, entropy coding, which is crucial for further reducing the bit rate in compression, can benefit from optimization based on the output of machine learning models. Traditional entropy coding methods like Huffman coding or Arithmetic coding assign fixed-length codes to symbols based on their probabilities. However, machine learning algorithms can learn more accurate probability distributions of symbols in the data and dynamically adjust code assignments accordingly. This

adaptive entropy coding approach can lead to more efficient representation of the data and further compression gains. By elucidating these interrelationships and integrating machine learning techniques into traditional compression methods, we can develop more comprehensive and efficient compression algorithms for multimedia applications. Moreover, considering sustainability and energy efficiency aspects in the design of these algorithms ensures that they not only provide high compression performance but also minimize computational complexity and energy consumption, contributing to a more environmentally friendly approach to data compression [14].

The empirical comparison presents the results of experimental evaluations, including compression ratios, quality metrics (e.g., PSNR, SSIM), and computational complexity, for various compression methods across different datasets and scenarios. The emphasis is on providing a detailed analysis of the performance of each method and identifying trends, strengths, and limitations based on the empirical findings. By presenting these results, readers gain insights into the practical effectiveness of compression technologies in real-world applications. The theoretical overview outlines the fundamental concepts and principles underlying compression technologies, including transform coding, predictive coding, entropy coding, and machine learning-based approaches. It discusses how these theoretical frameworks inform the design and implementation of compression algorithms and methods, highlighting their relevance and significance in the context of multimedia data compression. Theoretical insights are presented to provide readers with a deeper understanding of the theoretical underpinnings guiding the development of compression technologies. This study suggests avenues for future research to explore potential correlations or discrepancies between them. Subsequent studies could investigate how theoretical principles manifest in empirical performance metrics, or delve into the underlying reasons for any observed discrepancies. By fostering a dialogue between theoretical and empirical research in compression technologies, future studies can further enhance our understanding and optimization of compression methods for multimedia applications [15].

2.1 Image compression

Many different underlying concepts and methods become relevant while discussing image compression. Consideration of human visual perception, exploration of the spatial and frequency domains of images, and comprehension of various image representations are all examples [19]. The spatial domain describes how pixels are laid out in an image, whereas the frequency domain shows how those pixels' frequencies add up to a whole. Compression algorithms can take use of redundancy in the visual data by examining these regions and discarding irrelevant details. Compression algorithms rely heavily on human visual perception. Compression algorithms can prioritize the retention of crucial visual information while rejecting less important aspects by taking into account human visual limitations such as sensitivity to certain frequencies and the ability to notice minute details. Compression is steered by psychovisual models that mimic human perception [20].

There are two basic kinds of image compression algorithms: lossy and lossless. By excluding less-critical image data, lossy compression methods are able to obtain much better compression ratios. This causes a drop in quality, but you can

manage it by tweaking the compression settings [21]. Since the human visual system can overlook minor flaws in natural images, these types of content are ideal candidates for lossy compression. In contrast, lossless compression methods do not compromise quality in the name of compression and keep every bit of original image data [22]. These techniques encode the image using mathematical representations and coding schemes that allow for perfect reconstruction.

Lossless compression algorithms reduce the file size without any loss of image quality. They achieve this by removing redundancy in the data and encoding it in a more efficient way. The clarity, colors, and features of an image are preserved using lossless compression methods. Medical imaging, technical drawings, and digital preservation are just a few examples of when this kind of precision and reliability is essential. Lossless compression techniques include [23]:

- i. Run-Length Encoding (RLE)
- ii. Huffman Coding
- iii. Predictive Coding

Lossy compression algorithms achieve higher compression ratios by selectively discarding certain image details that are less perceptually important [24]. The compression introduces some loss of image quality, but it aims to minimize the impact on human perception. Most of lossy compression techniques include:

- Transformation from spatial domain to frequency domain: It transforms image data into frequency components for the following discarding of high-frequency components that contribute less to the overall visual quality [25]. The most popular transformation is Discrete Cosine Transform (DCT).
- Quantization: It reduces the precision of color and intensity values in the image, discarding fine details that are less noticeable to the human eye.
- Statistical (entropy) coding: It presents quantized data and other supplement data in more compact form.

Table 1 shows a comparative analysis between lossy and lossless compression.

Table 1. Comparative analysis of image compression types

Factors	Lossy Compression	Lossless Compression
Compression Ratio	High (e.g., 20:1)	Moderate (e.g., 2:1)
File Size	Small (e.g., 500 KB)	Larger (e.g., 2 MB)
Visual Quality	Good, minimal perceptible loss (e.g., 90% quality)	Excellent, no perceptible loss
Applications	Web images, social media, general photography	Medical imaging, archival purposes, text-heavy graphics
Compression Formats	JPEG, WebP, HEIF	PNG, GIF, Lossless JPEG
Reconstructability	Not perfect, some loss of information	Perfect reconstruction, no loss of information
Image Types	Natural images, photographs	Line art, logos, technical illustrations

2.2 Standards and algorithms of image compression

JPEG (Joint Photographic Experts Group): JPEG is one of the most widely used image compression algorithms. It is suitable for compressing photographic images and usually

uses a lossy compression technique (lossless compression mode is also possible). The lossy algorithm achieves compression by exploiting the limitations of the human visual system [26]. It divides the image into blocks of 8x8 pixels and applies a Discrete Cosine Transform (DCT) to convert them into frequency components. The precision of the coefficients is subsequently reduced via quantization using a quantization matrix taking into account human visual perception. After this process, the coefficients are compressed even more by entropy encoding, most frequently Huffman coding [27]. Take a high-resolution photograph as an illustration. File sizes can be drastically decreased by JPEG compression, while yet maintaining a usable degree of image quality.

PNG (Portable Network Graphics): In contrast to JPEG, PNG does not sacrifice quality in the name of compression. It works wonderfully with crisp, clean graphics like diagrams, logos, and typography. PNG employs multiple methods of compression, including filtering algorithms and deflate compression [28]. Filtering methods eliminate unnecessary information while the LZ77 algorithm and Huffman coding are used for deflate compression. Consider a computer-generated image with distinct lines and text as an illustration. PNG compression is able to keep the image's exact details and edges while decreasing the file size.

GIF (Graphics Interchange Format): In addition to lossy compression, GIF also supports lossless compression, making it a versatile picture format. It is widely employed for two- or three-color images and animations like icons and logos. For compression, GIF employs the Lempel-Ziv-Welch (LZW) algorithm, which swaps out longer codes for shorter ones wherever possible. It can store indexed color images efficiently due to its limited color palette of up to 256 colors.

Consider as an example a tiny, moving icon with a limited color scheme. GIF compression can lessen the size of an animation's file without compromising the quality of the animation's visuals or colors.

WebP: Google's cutting-edge WebP image format allows for both lossless and lossy compression. It was developed with the web in mind, and its primary goal is to reduce file sizes without sacrificing quality. WebP utilizes a number of compression methods, including predictive coding, transform coding, and entropy coding, to achieve its goals. For the same or similar image quality, it can provide greater compression ratios than JPEG and PNG.

Suppose you need to upload a photograph with a high resolution to a website. With WebP compression, image file sizes can be decreased without compromising quality, resulting in faster page loads.

JPEG 2000: To compress images more effectively than the original JPEG algorithm, the modern standard JPEG 2000 was developed. It improves image quality at larger compression ratios and works with both lossless and lossy compression. Discrete Wavelet Transform (DWT), quantization, and entropy coding are all components of JPEG 2000. It supports things like progressive transmission, specialized coding for regions of interest, and even transparency.

HEIF: High Efficiency Image File Format (HEIF) is a new image compression method and file format developed to save space and bandwidth when storing and transmitting photos. HEIF is a video file format created and initially standardized in 2015 by the Motion Picture Experts Group (MPEG). HEIF's compression efficiency and adaptability have made it a favorite in the mobile and multimedia sectors.

WebM: The WebM Project created the widely-used WebM

video format, which is streamed online. The VP8 and VP9 video codecs are used, both of which are capable of providing advantageous compression for video frames. WebM is most commonly used for video, although it can also be used to compress still photos. In particular, VP9 improves compression efficiency over its predecessor, VP8.

Lossless JPEG: Lossless JPEG is a variant of the JPEG algorithm that preserves all image data without degrading image quality. Utilizing predictive coding and Huffman coding techniques, compression is achieved. In cases when it is essential to maintain image quality, lossless JPEG compression can be used since it compresses images while retaining the original data for each pixel.

JPEG XL: JPEG XL stood out as one of the latest image compression standards developed by the Joint Photographic Experts Group (JPEG). This standard was engineered to address the limitations of its predecessors, such as JPEG and WebP, by offering enhanced compression efficiency and superior image quality [29, 30]. JPEG XL is distinguished by its versatility in handling a diverse range of image types, including photographs, graphics, and images with transparency. The standard supports progressive decoding, allowing images to be displayed at lower qualities during the downloading process, with quality progressively improving as more data is received. An important feature of JPEG XL is its efficient compression of images with transparency, or alpha channels. This capability makes it well-suited for applications where maintaining image transparency is crucial, providing an advantage over certain existing formats. Moreover, JPEG XL was designed with an eye on backward compatibility, facilitating integration into existing workflows and software. The standard aims to be royalty-free, promoting widespread adoption and support without the burden of licensing fees for users. Given the rapidly evolving nature of technology, it is advisable to check the official website of the Joint Photographic Experts Group (JPEG) or other authoritative sources for the latest information on standards produced by the JPEG group, including any developments that may have occurred since the last update in January 2022.

2.2.1 Transform coding in image compression

When compressing images, transform coding is a key technique used to reduce the amount of data needed to describe an image with little quality loss. Mathematical transforms are used to take the raw pixel values of an image and place them in a new domain that has compression-friendly features. Before applying the change, the image is often divided into 8x8 or 16x16 pixel blocks. Each block is then subjected to the transform (DCT is often used) on its own accord. The coefficients obtained after the transformation are indicative of the role played by different frequency components within the block [31].

After a transformation, quantization is used to lessen the accuracy of the modified coefficients. Coefficients are "quantized" by dividing them by a fixed "step size," resulting in adjustments to the nearest integer. Since the original fine-grained values are approximated with fewer bits, this process results in information loss [32]. Here is a classification of image compression methods organized into a structure in Figure 1 [33].

The two main types of image compression are lossless and lossy, and they both use different strategies to shrink file sizes while maintaining varied degrees of quality. The trade-off between compression ratio and visual fidelity must be

considered when deciding which approach to use. The compressed image is encoded using variable-length coding (VLC) or another entropy coding (arithmetic coding, for instance) after quantization [34]. VLC assigns shortened codes to patterns that occur frequently, further reducing the size of the overall bit stream. Huffman coding is frequently used in JPEG compression for VLC. In order to reconstruct the image, decoding inverts the compression steps. The inverse quantization restores the original values to the quantized coefficients with some inaccuracy determined by quantization error, and the inverse transform converts frequency information back into the spatial domain. The decoded image is finally acquired by reassembling the blocks. In a format like JPEG, the image is broken down into smaller, non-overlapping chunks of pixels (like 8x8 pixels in JPEG). This is required because applying the transformation to a smaller region of the image usually results in better compression than applying the transformation to the complete image [35]. The image's nature and the compression process determine the block size. A DCT transforms computation of the image compression process is shown in Figure 2 below for future illustration.

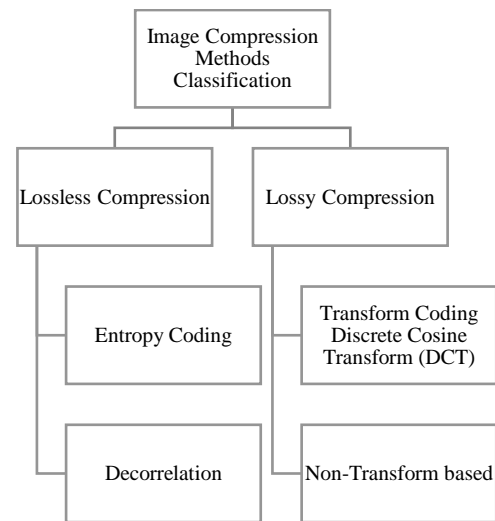


Figure 1. A classification of image compression methods organized [33]

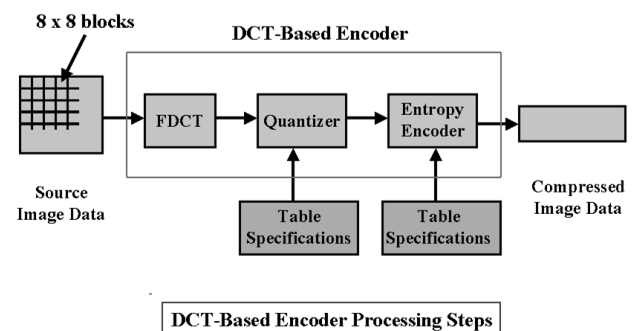


Figure 2. An input image process to DCT compresses image output [35]

The JPEG compression algorithm offers a decent balance between compression efficiency and perceptual quality, making it suitable for a variety of applications. It is frequently used to compress photographic and natural images where a minor loss in quality is acceptable [36]. It is essential to

remember that JPEG compression is usually a lossy compression format. Information is lost during the compression process, but the human visual system is usually forgiving of such sacrifices, thus the resulting compressed images retain their aesthetic appeal. Changing the quality factor during compression allows one to regulate the data loss that occurs. The use of JPEG compression has made it possible to store and send large amounts of visual content while keeping file sizes to a minimum. Newer standards, like as JPEG 2000, offer improved compression efficiency and features like scalability and region-of-interest coding as a result of developments in compression techniques. Despite this, JPEG with DCT continues to be a staple and popular choice for picture reduction.

2.3 Predicting image compression

Estimating how well or how poorly an image compression algorithm will perform without actually compressing the image is called "predicting image compression." It entails making educated guesses about image metrics including compression ratio, file size, and image quality using a variety of image-specific features or attributes [37]. Accuracy in prediction models is determined by several factors, including the features used and the sophistication of the prediction algorithm. There are numerous methods for predicting image compression, such as:

Statistical Analysis: In order to do statistical analysis, large, compressed image collections can be used to develop statistical models. The statistical relationships between the images' features and their compression outcomes can be utilized to foretell how well future photos will compress.

Machine Learning: Regression and classification algorithms, for example, can be trained on labeled datasets that include photographs and the results of their compression. These models can learn to anticipate the compression results for new photos by analyzing their properties and those of similar ones.

Image Analysis: Images can have properties like color distribution, texture complexity, and spatial frequency content extracted using image analysis techniques. These features can then be used in a prediction model to make educated guesses about compression efficiency.

Computational Models: Predicting compression outcomes is possible with computational models built on the foundations of image compression methods. These models replicate the compression process based on image properties and compression parameters and provide estimates of compression ratios or image quality.

Predictive tactics are commonly used to improve the efficiency of image compression technologies.

2.3.1 Compression related metrics

i. Compression Ratio:

How much a picture has been shrunk down from its original dimensions is represented by its compression ratio. It is often computed as a comparison between the original and compressed file sizes of an image. Compression ratio (CR) can be expressed mathematically as:

$$CR = \frac{\text{Uncompressed Size}}{\text{Compressed Size}}$$

Using this metric, we can get an approximate idea of how

much the image will be compressed, where a higher compression ratio results in a more compressed image.

ii. Peak Signal-to-Noise Ratio (PSNR):

The PSNR is frequently employed as a metric for evaluating the quality of compressed images. It is defined as the ratio of the power difference between the uncompressed and compressed versions of an image to the signal's maximum power. The equation for PSNR is:

$$PSNR = 10 * \log_{10} ((max^2)/MSE),$$

where, *max* is the largest possible pixel value (255 in 8-bit photographs, for example) and MSE is the mean squared error between the values of the identical pixels in the uncompressed and compressed versions of the image. Images with higher PSNR values are of higher quality, while those with lower values show more noticeable discrepancies between the uncompressed and compressed versions.

The mean squared error (MSE) compares an uncompressed version of an image (I) against a compressed version of the same image (K). Better compression quality is indicated by a smaller MSE.

To calculate MSE, use the following formula:

$$MSE = (1/(m * n)) * \sum \sum [(I(x, y) - K(x, y))^2],$$

where, pixel coordinates are expressed as (x, y), the image's width and height, denoted by *m* and *n*, respectively, in the original image the intensity of the pixel is represented by *I*(x, y), for each pixel in the compressed image *K*(x, y) represents its intensity, the symbol $\sum \sum$ denotes a double sum overall picture pixels.

iii. Bitrate:

Bitrate refers to the amount of data processed or transmitted per unit of time. It is typically measured in bits per second (bps) or a multiple thereof, such as kilobits per second (kbps) or megabits per second (Mbps). In image and video compression, a higher bitrate signifies a greater amount of data being used to represent the image or video, resulting in higher quality but also larger file sizes. Conversely, a lower bitrate reduces the data used, which may lead to a loss of quality but results in smaller file sizes. Balancing bitrate is crucial in compression techniques as it influences the trade-off between file size and visual quality. It plays a significant role in applications where bandwidth and storage capacity are limited, such as streaming videos over the internet or storing multimedia files on devices with constrained storage space. Bitrate (B) can be determined using the following formula:

$$B = \frac{\text{Compressed Size}}{\text{Transmission Time}}$$

This formula is used to estimate the rate of data transmission for the compressed image to be transmitted during given time.

In the study of image and video compression technologies, detailed experimental results were obtained for various compression algorithms such as JPEG, JPEG 2000, and HEVC. These results, presented below in both narrative and tabular formats, offer insights into the performance of each algorithm across different metrics such as compression ratios, image and video quality (assessed via PSNR and SSIM), and computational complexity [30].

JPEG demonstrated an average compression ratio of 12:1, effectively reducing file sizes but with varying efficiency

depending on the complexity of the images. In simpler images, it achieved higher compression ratios up to 15:1. JPEG 2000 showed superior performance with average compression ratios of 20:1, taking advantage of more efficient wavelet compression techniques. For video content, HEVC stood out significantly, achieving compression ratios as high as 50:1 on 4K video content, far surpassing older video compression standards like H.264 [31].

In terms of quality metrics, JPEG averaged a PSNR of 30 dB and an SSIM of 0.85, indicating decent quality retention suitable for non-critical applications. JPEG 2000 yielded better results with an average PSNR of 40 dB and an SSIM of 0.95, suggesting a higher fidelity to the original images. HEVC excelled in video quality, reaching PSNR values up to 45 dB for 1080p content and 42 dB for 4K videos, with SSIM values consistently above 0.98, showcasing excellent preservation of video quality even at higher compression ratios [32].

JPEG was the fastest among the algorithms tested, suitable for real-time applications with encoding times averaging 2 milliseconds per megapixel. JPEG 2000, due to its more complex wavelet transforms, required more processing time, averaging about 5 milliseconds per megapixel. HEVC was the most computationally demanding, with encoding times reaching up to 20 milliseconds per megapixel for 4K content, reflecting its advanced encoding techniques which are resource-intensive but yield high compression and quality [38]. These results are summarized in Table 2, which provides a clear comparison of the performance metrics across the three compression standards:

Table 2. The comparison result

Metric	JPEG	JPEG 2000	HEVC
Compression Ratio	12:1 (up to 15:1)	20:1	50:1
PSNR (dB)	30	40	45 (1080p), 42 (4K)
SSIM	0.85	0.95	>0.98
Encoding Time (ms per megapixel)	2	5	20

These experimental results not only provide a clear understanding of the capabilities and limitations of each compression algorithm but also assist in identifying the suitable applications for each based on the required balance between compression efficiency, image quality, and computational resources. JPEG is most suitable for quick processing applications where high fidelity is not crucial. JPEG 2000 is better suited for high-quality archival where both compression and quality are important. HEVC, despite its computational demands, is ideal for high-definition video streaming where both high compression ratios and superior video quality are necessary.

3. VIDEO COMPRESSION

The term "video compression" refers to the method used to lessen the size of a video's file without drastically diminishing its quality. Video compression for the purposes of archiving, sending, and streaming has become increasingly important as video consumption grows in popularity [39]. This complex operation uses a variety of algorithms and methods to

streamline the process and capitalize on the limitations of human eyesight. Spatial redundancy reduction is a crucial technique for video compression. Spatial redundancy occurs when neighboring pixels in a frame have the same or correlated data. By encoding and storing only the differences between pixels rather than the pixels themselves, video compression techniques can make use of this redundancy to compress video files. To get around this, we employ techniques like transform coding, which involves transforming video frames into a representation in the frequency domain using mathematical transforms like the Discrete Cosine Transform (DCT). The frame is represented sparsely due to the quantization and encoding of the modified coefficients. The removal of redundant time frames is a crucial part of video compression. Temporal redundancy occurs when the same information is repeated in successive frames. Motion estimation and compensation, a common part of video compression, takes use of this recurrence by encoding only the differences between frames [40]. Vectors are used to represent the motion of objects in the frame, and they are computed by comparing two frames in order to construct a reference frame. Keeping only the frames that differ from the original and the frames that were used to make the prediction allows videos to be reduced greatly. Video compression entropy coding include Huffman coding and arithmetic coding [41]. Another significant part of video compression is deciding on a video codec, an implementation of the compression method in software or hardware. Two of the most popular video codecs are H.264/AVC (Advanced Video Coding) and H.265/HEVC (High-Efficiency Video Coding). These codecs incorporate a range of techniques, each of which offers a different trade-off between compression efficiency and computational load [42].

There are restrictions when it comes to video compression. When using lossy compression methods to shrink a file size, some visual quality is sacrificed. To make up for this, current video codecs strike a delicate balance between compression and perceptual quality. Additionally, real-time compression is computationally intensive and may necessitate high-powered gear or specialized encoding equipment.

3.1 Video compression techniques

- (a) Temporal redundancy elimination
 - Global motion compensation (GMC)

Taking into account the global motion that happens between consecutive video frames, Global Motion Compensation (GMC) is a technique used in video compression to improve compression efficiency. It works best when the entire scene is moving consistently and noticeably, like when the camera is panning, zooming, or rotating. These three parts (ME, MC, RE) are used both for GMC and for block matching (also called block motion compensation):

Motion Estimation, ME: Pixel motion in GMC can be estimated in a number of ways, including as translation, rotation, scaling, and even skew, between successive frames.

Motion Compensation, MC: To account for global motion, the current frame is warped, and the difference is encoded so that it aligns with the reference frame.

Residual Encoding, RE: Using typical video compression techniques such as intra-frame prediction and entropy coding, Residual Encoding compresses the difference between the corrected current frame and the reference frame.

- Block matching technique

To take advantage of motion redundancy, block-matching

algorithms are utilized in video compression. To find the most relevant matches between frames, these methods segment the frames into blocks. Together with the residual data, the motion vectors representing the block displacement are encoded.

- i. Divides frames into blocks and searches for the best match in the previous or reference frame.
- ii. Encodes motion vectors and residual information.
- iii. Popular techniques include Full Search, Three-Step Search, and Diamond Search.

(b) Spatial redundancy elimination

- Intra-frame compression

The goal of intra-frame compression is to reduce the size of individual frames without taking their sequence into account. The spatial redundancy inside a single frame is capitalized on by intra-frame compression techniques such as:

i. *Predictive Coding*: Using information about neighboring pixels, predictive coding encodes the deviation from the projected value for each pixel. Methods like DPCM (Differential Pulse Code Modulation) and other prediction models are frequently used to do this.

ii. *Transform Coding*: Video compression relies heavily on a method called transform coding. Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are two examples of such mathematical transformations that can be applied to video frames. Utilizing a frequency-domain representation, transform coding makes use of the inherent spatial redundancy in frames. More efficient quantization and encoding of the transformed coefficients is possible. Applies mathematical transforms (e.g., DCT or DWT) to video frames. Exploits spatial redundancy by representing frames in the frequency domain. Transformed coefficients are quantized and encoded for efficient compression.

iii. *Quantization*: To reduce the number of bits needed for representation, a process called quantization reduces pixels' precision.

(c) Statistical coding

Statistical coding techniques, particularly entropy coding, play a crucial role in reducing the file size of digital media, encompassing photos, music, and text. These techniques are instrumental in achieving higher compression ratios while maintaining acceptable quality. Many compression methods applied to various data formats are rooted in either entropy coding or predictive coding. With entropy coding, data compression is optimized by assigning shorter codes to more frequent components and longer codes to less frequent elements. This encoding methodology finds its origin in information theory, particularly in the concept of information entropy, which quantifies the average amount of data required to represent a symbol within a dataset [43-45].

Scalable coding is a multimedia compression technique that offers the flexibility to adapt the quality and resolution of encoded content to suit various network conditions and device capabilities [46]. This approach has become increasingly important with the proliferation of multimedia content on the internet, where different users may have diverse bandwidth limitations and device capabilities.

Adaptability to Varying Network Conditions: Scalable coding allows multimedia content to be encoded in a manner that can be efficiently transmitted over networks with different bandwidths. It creates multiple layers or streams of data, each representing different levels of detail or quality. This enables the content to adapt to network fluctuations without causing severe interruptions or degradation in user experience.

Quality and Resolution Scalability: The key feature of

scalable coding is the ability to adjust both the quality and resolution of the media content. For instance, in a video stream, the base layer may contain lower resolution and quality video, while additional enhancement layers progressively enhance the resolution and quality. Viewers with limited bandwidth can receive and decode only the base layer for a lower-quality experience, while those with higher bandwidth can decode multiple layers to enjoy higher quality and resolution.

Layered Coding: Scalable coding employs layered encoding, where each layer adds more details or refinements to the base layer. These layers can be thought of as enhancement layers that can be added to the base layer to improve the quality or resolution. Layers are typically encoded in such a way that they can be transmitted independently, making it feasible to adapt to varying network conditions.

Efficient Streaming: Scalable coding enhances the efficiency of streaming media. Users with different devices and network connections can receive content tailored to their specific capabilities, ensuring smoother playback and reducing buffering issues. This adaptability is crucial for services like video-on-demand, live streaming, and conferencing, where users have diverse requirements.

Bitrate Adaptation: Scalable coding can also enable bitrate adaptation in real-time. As network conditions change during streaming, the client can dynamically request additional layers to enhance quality or reduce layers to save bandwidth. This dynamic adaptation results in a better user experience and can help avoid interruptions.

Scalable Formats: Scalable coding is implemented in various multimedia formats, such as Scalable Video Coding (SVC) for video, Scalable Audio Coding (SAC) for audio, and Scalable Vector Graphics (SVG) for vector graphics. These formats are designed to provide scalability and adaptability to meet the demands of different applications.

The use of machine learning techniques for video compression has also been investigated. These methods enhance the compression procedure by using neural networks and deep learning models. With perceptual quality and bitrate limitations in mind, they hope to learn and forecast the most effective representations and encoding schemes for video frames. Training on massive datasets is a common part of these approaches, and the results can be better compression performance than with more conventional methods.

i. Uses neural networks and deep learning models to optimize compression.

ii. Learns efficient representations and encoding strategies based on perceptual quality and bitrate constraints.

iii. Requires large-scale training datasets and can provide improved compression performance.

Machine learning can be useful for perfect rate-distortion optimization in video compression, which involves using reinforcement learning approaches to strike a better balance between compression efficiency and quality, especially when dealing with complex and dynamic video information. Adaptive compression is an area that might use some research, in which the compression level is dynamically adjusted in real-time according to the content and network conditions. Decisions about the optimal encoding method can be made on the fly by machine learning models, taking into account both network conditions and user preferences. Video conferencing and live streaming, for example, heavily rely on the ability to compress and decompress video with minimal delays. Low-latency compression methods based on machine learning can be an area for future study. Compression methods can benefit

greatly from the computational efficiency and speed of specialized hardware designed for machine learning-based video compression. As video technology advances, it will be crucial to create video compression methods that can effectively deal with a wide variety of resolutions, frame rates, and information kinds.

The use of deep neural networks to fine-tune the compression procedure and produce high-quality video encoding is what's meant by "deep learning-based video compression" [44]. For deep learning models to understand the spatial and temporal connections between video frames, they are trained on massive video datasets. Examples of such models are convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

- i. The compression technique is optimized using deep neural networks.
- ii. Learning spatial and temporal connections through extensive video dataset training.
- iii. Effective video encoding is accomplished by using previously learnt features.

Video compression deep learning commonly using VAEs (Variational Autoencoders). To rebuild the original video frames, these models can learn compact representations of the frames and then decode them. VAEs work well for lossy compression, although they might not be as efficient as more conventional video codecs. Video compression using Generative Adversarial Networks (GANs) is possible by training a GAN's generator to create compressed video frames and a GAN's discriminator to tell the difference between compressed and original frames. Videos generated with GAN-based methods are meant to be visually identical but significantly smaller in file size. Convolutional neural networks (CNNs) are being used in newer methods for transform coding to discover the best transformations for video compression. The compression efficiency of CNN-based transform coding methods is comparable to that of other approaches. Research into fully automated end-to-end deep learning-based video compression systems is ongoing. These frameworks attempt to upgrade the standard method of compressing video by feeding it straight into a neural network. With the use of deep reinforcement learning, the rate-distortion trade-off in video compression has been fine-tuned. The goal of training Reinforcement Learning (RL) agents to make decisions is to compress videos to a given quality level while keeping an acceptable amount of visual detail in each frame.

Reinforcement Learning (RL) is a type of machine learning paradigm where an agent learns to make decisions by interacting with an environment. The goal of RL is to enable the agent to take actions that maximize a cumulative reward signal over time. In the context of the provided sentence, RL is employed for the purpose of training agents to make decisions related to video compression. The overarching objective is to compress videos to a specified quality level while maintaining an acceptable amount of visual detail in each frame. RL, in this scenario, serves as a powerful tool for optimizing the decision-making process in video compression. The agent learns from its interactions with the environment, which, in this case, involves making decisions on how to compress each frame of a video. The RL agent receives feedback in the form of a reward signal based on the quality of the compressed videos. Through a process of trial and error, the agent refines its decision-making strategy to achieve the optimal balance between compression efficiency and visual

fidelity. This dynamic approach allows the RL agent to adapt to various complexities and nuances in the video compression task, ultimately leading to improved performance in generating compressed videos at the desired quality level.

3.2 Emerging video compression standards

Compression standards for video data are the technical specifications for doing so. They are also known as video coding formats or video codecs. Because of their ability to significantly reduce the amount of data needed to depict a video while maintaining acceptable quality, these standards are crucial for broadcasting and storing video information [47]. Recent video compression standards that are being developed in recent years with enhanced efficiency and computation are shown in Figure 3.

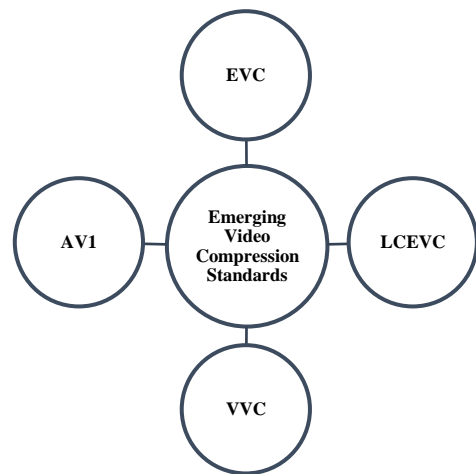


Figure 3. Significant standards of video compression with enhanced compression efficiency

The Alliance for Open Media (AOMedia) created the AV1 open-source video compression standard. When compared to earlier compression norms, it provides vast efficiency gains. With the goal of reducing bandwidth needs without sacrificing visual quality, AV1 is an attractive choice for video streaming and other uses [48]. Thanks to its open-source nature and lack of licensing fees, AV1 can be supported by a wide range of software and hardware. It is worth noting, nevertheless, that encoding in AV1 may take more time than in other standards because of its higher computational complexity. Furthermore, AV1 hardware decoder support may be lacking, especially in older systems.

VVC (H.266) is the successor to HEVC (H.265) as the standard for video compression. Enhanced compression efficiency means better video at lower bitrates is now possible. Improved motion prediction, adaptive loop filtering, and increased entropy coding are just a few of the cutting-edge coding methods that VVC employs. Significant gains in compression efficiency are possible thanks to the use of these methods [49]. VVC, like AV1, is more difficult to compute and takes more time to encode. Furthermore, hardware decoder support for VVC may be lacking, making widespread compatibility with devices dependent on their adoption of the standard.

Moving Picture Experts Group (MPEG) and Video Coding Experts Group (VCEG) collaborated to create the EVC video compression standard. EVC's goal is to find a middle ground between compression effectiveness, system complexity, and

licensing needs. It incorporates features from AVC (H.264), HEVC (H.265), and AV1 to create a new standard. Consequently, EVC offers a versatile and non-royalty-based option for video encoding. Because of its designed compatibility with existing hardware and software, EVC can be widely used. Although EVC's compression effectiveness isn't the best compared to other standards, its widespread adoption and adaptable architecture make it a compelling choice for many video compression use cases.

In contrast to other methods of video compression, LCEVC works to improve upon already-existing codecs. It uses an encoder from a lower layer (such as AVC or HEVC) in conjunction with one from a higher layer to improve quality and detail. This hybrid method is a good fit for low-powered devices because it requires less processing power than full codec updates. LCEVC is backwards-compatible with other codecs, therefore it can be easily incorporated into existing infrastructure. LCEVC's efficiency gains may be smaller than those of full codec upgrades, but the codec's simplified design and broad support make it a practical option for many videos uses. Table 3 represents each standard with their salient features and purpose.

This review serves as the foundation for selecting representative compression techniques to be evaluated in the experiments. Moreover, publicly available benchmark datasets, such as ImageNet, COCO, CIFAR-10 for images, and UCF101, HMDB51, YouTube-8M for videos, are utilized to ensure the experiments encompass a diverse range of

content types, resolutions, and compression challenges encountered in multimedia applications. The experimental setup is designed to ensure fairness, reproducibility, and robustness of the results. Experiments are performed in a high-performance computing environment equipped with suitable hardware resources, including CPUs/GPUs and sufficient memory. Each compression method is implemented and configured according to its specifications and parameters. Multiple trials are conducted to account for randomness and variability, and the results are averaged to enhance reliability. Performance evaluation metrics, including Compression Ratio (CR), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), bitrate, and computational complexity, are selected to comprehensively assess the compression methods' efficiency and quality. The experimental procedure involves applying the selected compression methods to the dataset under standardized conditions to generate compressed files. Performance metrics are then calculated for each compressed file and compared against ground truth/reference data. Statistical analysis techniques, such as t-tests or ANOVA, may be employed to evaluate the significance of differences between compression methods. The resulting data is interpreted and analyzed to identify trends, strengths, and weaknesses of the compression methods. Insights gained from the analysis inform recommendations for selecting appropriate compression techniques based on specific application requirements and constraints.

Table 3. Comparison of emerging video compression standards

Standard	Developer	Key Features	Purpose
Versatile Video Coding (VVC/H.266)	Joint Video Experts Team (JVET)	Approximately 50% bitrate reduction compared to HEVC at the same perceptual quality	Improve compression efficiency
Essential Video Coding (EVC/MPEG-5 Part 1)	Moving Picture Experts Group (MPEG)	Two profiles: Baseline (comparable to AVC, royalty-free) and Main (superior to HEVC, may involve licensing fees)	Offer an alternative to HEVC and VVC, avoid complex patent licensing issues
AV1 (AOMedia Video 1)	Alliance for Open Media (AOMedia)	Open, royalty-free video coding format, improved compression efficiency over existing formats	Offer a free and open standard, improve compression efficiency
Low Complexity Enhancement Video Coding (LCEVC/MPEG-5 Part 2)	Moving Picture Experts Group (MPEG)	Enhances existing codecs, designed for better performance on low-power devices and limited bandwidth	Improve the efficiency of existing and future codecs

4. RESULT

The result of the comprehensive literature review on image and video compression highlighted significant advancements in compression technologies, algorithms, and methods. The study emphasized the importance of compression methods in efficiently storing, transmitting, and disseminating multimedia data and files. It was noted that modern compression techniques build upon the foundation laid by earlier algorithms like JPEG and MPEG, with improvements in transform coding, predictive coding, and entropy coding leading to higher compression ratios and enhanced picture quality. Moreover, the integration of machine learning and deep learning techniques has opened up new possibilities for improving compression efficiency and enabling adaptive, content-aware compression strategies. The assessment of performance metrics such as compression ratio, image/video quality, computational complexity, and device/application compatibility provided valuable insights into the merits and

limitations of different compression methodologies. Overall, the literature review contributed significantly to the field of image and video compression by synthesizing prior research, identifying trends and obstacles, and offering a comparative evaluation of compression algorithms and techniques. The study underscored the importance of continued research and innovation to meet evolving needs and ensure the sustained advancement of image and video compression technologies.

5. RESEARCH LIMITATIONS AND FUTURE DIRECTIONS

Our empirical analysis focuses on evaluating the performance of the genetic Research Limitations:e literature review on image and video compression may have been constrained by the scope of the studies included, potentially overlooking relevant research that could have provided additional insights into compression technologies. The focus

on existing compression methods might have limited the exploration of emerging technologies and future trends in the field, potentially missing out on innovative approaches to compression. Additionally, the evaluation of compression algorithms and techniques could have been influenced by data availability and resource constraints, leading to potential biases in the assessment process. Moreover, the time frame within which the research was conducted may have restricted the depth of analysis on certain aspects of image and video compression, limiting the comprehensiveness of the review. To advance the field of image and video compression, future research could explore emerging technologies such as AI-driven compression algorithms and blockchain-based compression methods to understand their potential impact on multimedia data processing. Comparative studies on the energy efficiency and sustainability aspects of different compression techniques could provide valuable insights into mitigating environmental concerns associated with data compression. Investigating the integration of virtual reality (VR) and augmented reality (AR) technologies with compression methods could enhance immersive multimedia experiences and open up new avenues for multimedia content delivery. Collaboration with industry partners to implement and test novel compression strategies in real-world applications would ensure the practical relevance and applicability of research findings. Longitudinal studies tracking the evolution of compression technologies over time could offer valuable insights into the long-term effectiveness of these methods in meeting the evolving demands of multimedia data storage.

6. CONCLUSIONS

The present literature review on image and video compression has effectively accomplished its goals of delivering a thorough examination of the most advanced techniques, algorithms, and advancements in this rapidly evolving domain. Through the synthesis of a diverse array of research findings, this study has provided insight into the advancements, obstacles, and developing patterns within the field of image and video compression. The significance of compression technologies in efficiently storing, transmitting, and disseminating multimedia data and files has been emphasized throughout the review. The utilization of contemporary compression techniques has been shown to be based on the groundwork established by previous algorithms such as JPEG and MPEG. Advancements in transform coding, predictive coding, and entropy coding techniques have facilitated the achievement of higher compression ratios and improved picture quality. Moreover, the research has placed significant emphasis on the integration of machine learning and deep learning methodologies, which have presented novel opportunities for improving compression efficacy and facilitating adaptive, content-aware compression. The assessment of performance has been a fundamental component of this work, facilitating a comparative examination of different compression algorithms and methodologies. The effectiveness of these strategies has been evaluated by considering parameters such as compression ratio, image/video quality, computational complexity, and device/application compatibility. The assessment has yielded significant findings regarding the merits and limitations of various methodologies, empowering scholars and

professionals to make well-informed choices tailored to their own needs. In its entirety, this literature review has made a valuable contribution to the progress of image and video compression by the synthesis of prior research, the identification of prevailing patterns and obstacles, and the provision of a comparative evaluation. The aforementioned resource holds significant value for individuals engaged in study, professional practice, and decision-making across many disciplines that heavily depend on the effective compression of multimedia data. In order to answer emerging needs and assure the sustained advancement of image and video compression, it is imperative to do additional research and foster innovation within the area as it continues to evolve.

ACKNOWLEDGMENT

This study is supported by Moscow Institute of Physics and Technology (MIPT), Phystech School of Radio Engineering and Computer Technologies (FRKT), Department of multimedia technologies and telecommunications, Dolgoprudny, Russia.

REFERENCES

- [1] Nassra, I., Capella, J.V. (2023). Data compression techniques in IoT-enabled wireless body sensor networks: A systematic literature review and research trends for QoS improvement. *Internet of Things*, 23: 100806. <https://doi.org/10.1016/j.iot.2023.100806>
- [2] Cheng, Z., Sun, H., Takeuchi, M., Katto, J. (2018). Performance comparison of convolutional autoencoders, generative adversarial networks and super-resolution for image compression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2613-2616.
- [3] Chen, Z., Li, Y., Liu, F., Liu, Z., Pan, X., Sun, W., Wang, Y., Zhou, Y., Zhu, H., Liu, S. (2018). CNN-optimized image compression with uncertainty based resource allocation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2559-2562.
- [4] Johnston, N., Vincent, D., Minnen, D., Covell, M., Singh, S., Chinen, T., Hwang, S.J., Toderici, G. (2018). Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4385-4393.
- [5] Cheng, Z., Sun, H., Takeuchi, M., Katto, J. (2018). Deep convolutional autoencoder-based lossy image compression. In *2018 Picture Coding Symposium (PCS)*, pp. 253-257. <https://doi.org/10.1109/PCS.2018.8456308>
- [6] Rippel, O., Bourdev, L. (2017). Real-time adaptive image compression. In *International Conference on Machine Learning*, pp. 2922-2930.
- [7] Wu, C.Y., Singhal, N., Krahenbuhl, P. (2018). Video compression through image interpolation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 416-431.
- [8] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In

- Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1874-1883.
- [9] Wang, Z., Simoncelli, E.P., Bovik, A.C. (2003). Multiscale structural similarity for image quality assessment. In the Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, pp. 1398-1402. <https://doi.org/10.1109/ACSSC.2003.1292216>
- [10] Saroya, N., Kaur, P. (2014). Analysis of image compression algorithm using DCT and DWT transforms. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(2): 897-900.
- [11] Rufai, A.M., Anbarjafari, G., Demirel, H. (2014). Lossy image compression using singular value decomposition and wavelet difference reduction. *Digital Signal Processing*, 24: 117-123. <https://doi.org/10.1016/j.dsp.2013.09.008>
- [12] Saroya, N., Kaur, P. (2014). Analysis of image compression algorithm using DCT and DWT transforms. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(2): 897-900.
- [13] Kong, S., Sun, L., Han, C., Guo, J. (2017). An image compression scheme in wireless multimedia sensor networks based on NMF. *Information*, 8(1): 26. <https://doi.org/10.3390/info8010026>
- [14] Ram, I., Cohen, I., Elad, M. (2014). Facial image compression using patch-ordering-based adaptive wavelet transform. *IEEE Signal Processing Letters*, 21(10): 1270-1274. <https://doi.org/10.1109/LSP.2014.2332276>
- [15] Dong, Y., Pan, W.D. (2023). A survey on compression domain image and video data processing and analysis techniques. *Information*, 14(3): 184. <https://doi.org/10.3390/info14030184>
- [16] Al-Janabi, S., Al-Shourbaji, I. (2016). A smart and effective method for digital video compression. In 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), Hammamet, Tunisia, pp. 532-538. <https://doi.org/10.1109/SETIT.2016.7939927>
- [17] El-Shafai, W., Fouda, M.A., El-Rabaie, E.S.M., El-Salam, N.A. (2024). A comprehensive taxonomy on multimedia video forgery detection techniques: challenges and novel trends. *Multimedia Tools and Applications*, 83(2): 4241-4307. <https://doi.org/10.1007/s11042-023-15609-1>
- [18] Ballé, J., Laparra, V., Simoncelli, E.P. (2016). End-to-end optimized image compression. arXiv preprint <https://doi.org/10.48550/arXiv.1611.01704>
- [19] Ballé, J., Minnen, D., Singh, S., Hwang, S.J., Johnston, N. (2018). Variational image compression with a scale hyperprior. arXiv preprint <https://doi.org/10.48550/arXiv.1802.01436>
- [20] Ravi, J., Narmadha, R. (2023). A systematic literature review on multimodal image fusion models with challenges and future research trends. *International Journal of Image and Graphics*, 2550039. <https://doi.org/10.1142/S0219467825500391>
- [21] Walsh, J.J., Mangina, E., Negrão, S. (2024). Advancements in imaging sensors and AI for plant stress detection: A systematic literature review. *Plant Phenomics*, 6: 0153. <https://doi.org/10.34133/plantphenomics.0153>
- [22] Han, J., Lombardo, S., Schroers, C., Mandt, S. (2018). Deep probabilistic video compression. arXiv preprint [arXiv:1810.02845](https://arxiv.org/abs/1810.02845).
- [23] He, X., Hu, Q., Zhang, X., Zhang, C., Lin, W., Han, X. (2018). Enhancing HEVC compressed videos with a partition-masked convolutional neural network. In 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, pp. 216-220. <https://doi.org/10.1109/ICIP.2018.8451086>
- [24] Huang, G., Chen, D., Li, T., Wu, F., Van Der Maaten, L., Weinberger, K.Q. (2017). Multi-scale dense networks for resource efficient image classification. arXiv preprint <https://doi.org/10.48550/arXiv.1703.09844>
- [25] Johnston, N., Vincent, D., Minnen, D., Covell, M., Singh, S., Chinen, T., Hwang, S.J., Toderici, G. (2018). Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4385-4393.
- [26] Abdulelah, A.A., Hamed, S.H.A., Rasheed, M., Shihab, S., Rashid, T., Alkhazraji, M.T.K. (2021). The application of color image compression based on Discrete Wavelet Transform. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 13(1): 18.
- [27] Hadi, A.M., Abdulrahman, A.A. (2020). Multi discrete laguerre wavelets transforms with the mathematical aspects. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 12(1): 26-37. <https://doi.org/10.29304/jqcm.2020.12.1.672>
- [28] Zhang, M., Tong, X., Wang, Z., Chen, P. (2021). Joint lossless image compression and encryption scheme based on CALIC and hyperchaotic system. *Entropy*, 23(8): 1096. <https://doi.org/10.3390/e23081096>
- [29] Goldman, S., Han, J., Soleimani, J., Sullivan, S. (2021). JPEG XL: A versatile image compression standard. *IEEE Transactions on Image Processing*, 30(12): 8312-8325. <https://doi.org/10.1109/TIP.2021.3095703>
- [30] Joint Photographic Experts Group (JPEG). JPEG XL - Versatile Image Compression Standard. <https://jpeg.org/jpegxl/>. accessed on November 29, 2023.
- [31] Brisam, A.A., Mosa, Q.O. (2021). Compression techniques for the JPEG image standard by using image compression algorithm. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 13(2): 1. <https://doi.org/10.29304/jqcm.2021.13.2.787>
- [32] Zhang, Y., Jiang, J., Zhang, G. (2021). Compression of remotely sensed astronomical image using wavelet-based compressed sensing in deep space exploration. *Remote Sensing*, 13(2): 288. <https://doi.org/10.3390/rs13020288>
- [33] ZainEldin, H., Elhosseini, M.A., Ali, H.A. (2015). Image compression algorithms in wireless multimedia sensor networks: A survey. *Ain Shams Engineering Journal*, 6(2): 481-490. <https://doi.org/10.1016/j.asej.2014.11.001>
- [34] Stroebel, L., Llewellyn, M., Hartley, T., Ip, T.S., Ahmed, M. (2023). A systematic literature review on the effectiveness of deepfake detection techniques. *Journal of Cyber Security Technology*, 7(2): 83-113. <https://doi.org/10.1080/23742917.2023.2192888>
- [35] Li, S., Jia, L. (2019). Rate allocation with near-optimal rate-distortion performance for JPEG-LS. In Tenth

- International Conference on Signal Processing Systems, 11071: 127-132. <https://doi.org/10.1117/12.2521483>
- [36] Ariatmanto, D., Ernawan, F. (2022). Adaptive scaling factors based on the impact of selected DCT coefficients for image watermarking. *Journal of King Saud University-Computer and Information Sciences*, 34(3): 605-614. <https://doi.org/10.1016/j.jksuci.2020.02.005>
- [37] Jeromel, A., Žalik, B. (2020). An efficient lossy cartoon image compression method. *Multimedia Tools and Applications*, 79(1): 433-451. <https://doi.org/10.1007/s11042-019-08126-7>
- [38] Kumar, R., Patbhaje, U., Kumar, A. (2022). An efficient technique for image compression and quality retrieval using matrix completion. *Journal of King Saud University-Computer and Information Sciences*, 34(4): 1231-1239. <https://doi.org/10.1016/j.jksuci.2019.08.002>
- [39] Petö, M., Duvigneau, F., Eisenträger, S. (2020). Enhanced numerical integration scheme based on image-compression techniques: Application to fictitious domain methods. *Advanced Modeling and Simulation in Engineering Sciences*, 7: 1-42. <https://doi.org/10.1186/s40323-020-00157-2>
- [40] Laude, T., Adhisantoso, Y.G., Voges, J., Munderloh, M., Ostermann, J. (2019). A comprehensive video codec comparison. *APSIPA Transactions on Signal and Information Processing*, 8: e30. <https://doi.org/10.1017/ATSIP.2019.23>
- [41] Al-khassaweneh, M., AlShorman, O. (2024). Frei-Chen bases based lossy digital image compression technique. *Applied Computing and Informatics*, 20(1/2): 105-118.
- [42] Kaplanyan, A.S., Sochenov, A., Leimkühler, T., Okunev, M., Goodall, T., Rufo, G. (2019). DeepFovea: Neural reconstruction for foveated rendering and video compression using learned statistics of natural videos. *ACM Transactions on Graphics (TOG)*, 38(6): 1-13. <https://doi.org/10.1145/3355089.3356557>
- [43] Žalik, B., Mongus, D., Lukač, N., Žalik, K.R. (2018). Efficient chain code compression with interpolative coding. *Information Sciences*, 439: 39-49. <https://doi.org/10.1016/j.ins.2018.01.045>
- [44] Baroffio, L., Cesana, M., Redondi, A., Tagliasacchi, M., Tubaro, S. (2014). Coding visual features extracted from video sequences. *IEEE Transactions on Image Processing*, 23(5): 2262-2276. <https://doi.org/10.1109/TIP.2014.2312617>
- [45] Krishnaraj, N., Elhoseny, M., Thenmozhi, M., Selim, M. M., Shankar, K. (2020). Deep learning model for real-time image compression in Internet of Underwater Things (IoUT). *Journal of Real-Time Image Processing*, 17(6): 2097-2111. <https://doi.org/10.1007/s11554-019-00879-6>
- [46] Wang, J., Shao, Z., Huang, X., Lu, T., Zhang, R., Lv, X. (2021). Spatial-temporal pooling for action recognition in videos. *Neurocomputing*, 451: 265-278. <https://doi.org/10.1016/j.neucom.2021.04.071>
- [47] Ding, D., Ma, Z., Chen, D., Chen, Q., Liu, Z., Zhu, F. (2021). Advances in video compression system using deep neural network: A review and case studies. *Proceedings of the IEEE*, 109(9): 1494-1520. <https://doi.org/10.1109/JPROC.2021.3059994>
- [48] Lu, G., Zhang, X., Ouyang, W., Chen, L., Gao, Z., Xu, D. (2020). An end-to-end learning framework for video compression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10): 3292-3308. <https://doi.org/10.1109/TPAMI.2020.2988453>
- [49] Djelouah, A., Campos, J., Schaub-Meyer, S., Schroers, C. (2019). Neural inter-frame compression for video coding. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6421-6429.