International Journal of Technology
http://ijtech.eng.ui.ac.id

Research Article

Efficient Block Partitioning Method for Spatial Scalable Encoding in Versatile Video Coding

Yo Koura[1], **Naoya Niwa** [1*], **Hiroe Iwasaki**[1,2]

[1]Tokyo University of Agriculture and Technology, 2-24-16 Nakacho, Koganei-shi, Tokyo-to 184-8588 Japan

[2]Tohoku University, 6-6 Aoba, Aramaki, Aoba-ku, Sendai-shi, Miyagi-ken 980-0845 Japan

[*]Corresponding author: naoya@go.tuat.ac.jp; Tel.: +81423887696; Fax: +81423887696

Abstract: The rapid increase in video consumption, further accelerated by the coronavirus disease 2019 (COVID-19) pandemic, has driven a significant rise in demand for live streaming services and video delivery across various devices, such as smartphones and tablets. Screen size and display resolution of these devices vary widely, creating the need for flexible and efficient video transmission methods. Spatial scalable coding addresses this challenge by encoding multiple resolutions within a single bitstream, allowing devices to decode the appropriate resolution without requiring separate streams. This approach reduces redundancy and improves transmission efficiency. Versatile Video Coding (VVC), the latest international video compression standard, supports spatial scalability through its multilayer profile. VVC enhances compression performance by reusing information from lower-resolution layers; however, this added complexity increases computational overhead, particularly during encoding. In this paper, we propose an efficient block partitioning method specifically designed for scalable VVC-based coding. The method exploits structural similarities between low-resolution and high-resolution encoded data to guide partitioning decisions in the high-resolution layer, thereby reducing unnecessary computations. The experimental results demonstrate that the proposed method reduces encoding time by approximately 55%, with a BD-rate increase below 3.45%. These results validate the effectiveness of the approach in accelerating scalable video encoding without compromising visual quality, making it suitable for real-time applications and environments with limited computational resources.

Keywords: Block partitioning algorithm; Spatial scalable coding; Versatile Video Coding (VVC); Video coding

1. Introduction

Since the COVID-19 pandemic, there has been a continuous increase in overall network traffic due to the growing consumption of video content (Peroni and Gorinsky, 2025; Timmerer et al., 2025; Business Research Insights, 2024; Jhang-Li and Liou, 2023; Sandvine, 2023; Feldmann et al., 2020). The demand for live streaming has grown rapidly due to the increase in online events, remote communication, and real-time broadcasting services among various types of video services (Cong et al., 2021). Simultaneously, more users are watching videos on portable devices such as smartphones and tablets (Zheng, 2025; Marketing Scoop, 2023). Multiple bitstreams are often required to support various screen sizes and resolutions, which increases overall transmission bandwidth requirements.

Spatial scalable coding enables efficient multi-resolution video encoding using a layered bitstream with multiple resolutions (Sun et al., 2007). It is supported as a standard feature in the multilayer profile of Versatile Video Coding (VVC), the latest international video coding standard that began with MPEG-2, followed by Advanced Video Coding (AVC/H.264) and High Efficiency Video Coding (HEVC/H.265) (Bross et al., 2021b; Bross et al., 2021a; Fischer, 2020; Zhang and Mao, 2019; Sullivan et al., 2012; Wiegand et al., 2003). More block shapes are adopted

https://doi.org/10.14716/ijtech.v17i1.8218
Received July 2025; Revised November 2025; Accepted November 2025

in VVC than in previous standardization specifications, and determining the appropriate block shape consumes computational resources. Furthermore, the increased computational complexity of the basic encoding process significantly increases the search complexity between videos at different resolutions.

We propose a new method for accelerating spatial scalable coding using similarities in the block partitioning structure between videos at different resolutions. The proposed method achieves an average reduction of 55% in encoding time, with a maximum BD-rate increase of less than 3.45%, which is generally acceptable in terms of perceived visual quality. The remainder of this paper is organized as follows. Section 2 provides an overview of spatial scalable coding in the VVC framework. Section 3 presents the details of the proposed method. Section 4 presents the experimental results and discussion. Finally, Section 5 concludes the paper.

2. VVC-Based Spatial Scalable Coding

2.1 Overview of the VVC

Versatile Video Coding (VVC) is the most recent video compression standard in the lineage of international standards. Developed to meet the growing demand for high-resolution and diverse video content, VVC supports a wide range of video formats, including 4K and 8K ultra-high-definition video, high dynamic range (HDR), and 360° immersive video. To accommodate such formats, VVC introduces a highly flexible and efficient coding framework equipped with numerous new tools and advanced features and achieves approximately 40%–50% reduction in bit rate compared to HEVC (Amestoy et al., 2023; Bonnineau et al., 2022; Grois et al., 2021). However, this improvement in compression efficiency comes at the cost of significantly increasing the encoding time.

Among the many factors that contribute to this complexity, block partitioning is particularly critical. Block partitioning refers to the process of dividing a frame into CUs of various sizes and shapes. Each CU is a fundamental unit for prediction, transformation, and quantization, as shown in Figure 1. As illustrated in Figure 2, VVC extends the block partitioning structure of HEVC by adding BT and TT splits to the conventional QT structure (Huang et al., 2021). However, the increased number of partitioning possibilities significantly increases computational complexity (Mercat et al., 2021; Pakdaman et al., 2020). Block partitioning consumes more than 90% of the total VVC encoding time (Tissier et al., 2019).

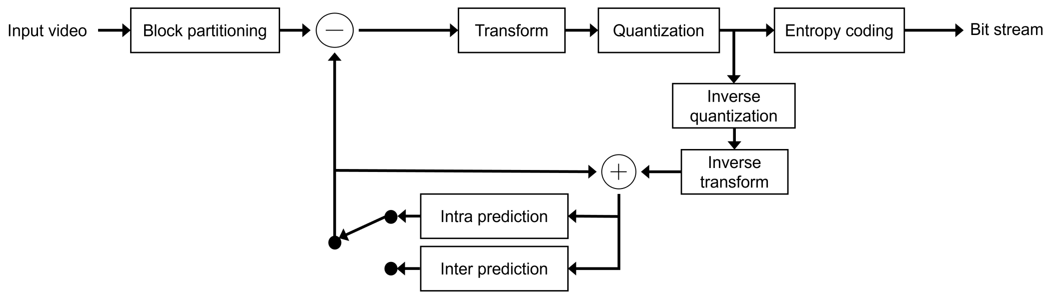

Figure 1 Overview of the VVC coding process

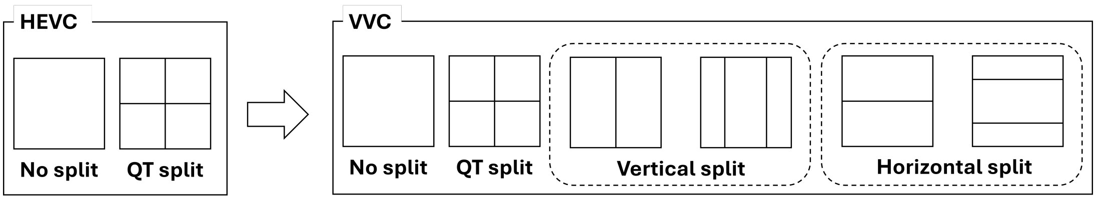

Figure 2 The block partition patterns specified in VVC

2.2 Spatial scalable coding

As shown in Figure 3, spatial scalable coding enables the efficient transfer of videos at different resolutions within a single coding framework (Shahid et al., 2017; Unanue et al., 2011; Schwarz et al., 2007). This technique has evolved through earlier scalable extensions such as SVC for H.264/AVC and SHVC for H.265/HEVC (Boyce et al., 2015; Fraunhofer Heinrich-Hertz-Institut, 2014; Sullivan et al., 2013; Hayase et al., 2012; Ohm, 2005). VVC provides spatial scalability by hierarchically encoding multiple layers—namely, a base layer (BL) and an enhancement layer (EL). The standard VVC encoding tool is used to encode the BL, and its decoded frames are used as prediction references in the EL. The residual between the up-sampled BL frame and the original high-resolution input is calculated. In addition, the EL supports the reuse of coding information from the BL, such as motion vectors. This reduces redundancy and improves compression efficiency.

With this layered structure, decoding only the BL provides a low-resolution version, whereas decoding both BL and EL reconstructs the high-resolution video. Compared with encoding each resolution independently, spatial scalable coding achieves bitrate savings by avoiding duplicate encoding.

However, this approach increases the complexity of encoding (Marquant et al., 2022; Wang et al., 2010; Takahashi and Yamada, 2008). In single-layer coding, prediction is solely based on spatially or temporally adjacent blocks. In contrast, spatial scalable coding must also evaluate the effectiveness of reusing information from the base layer, which introduces additional computational overhead. Figure 4 illustrates the extra processing steps required in the scalable coding process.

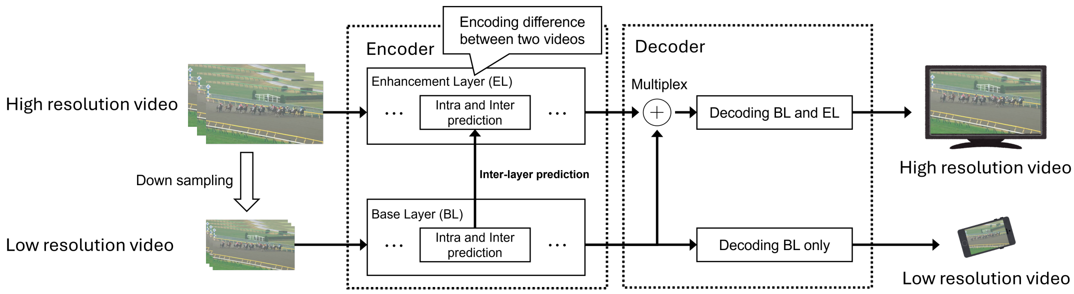

Figure 3 Structure of prediction in spatial scalable coding

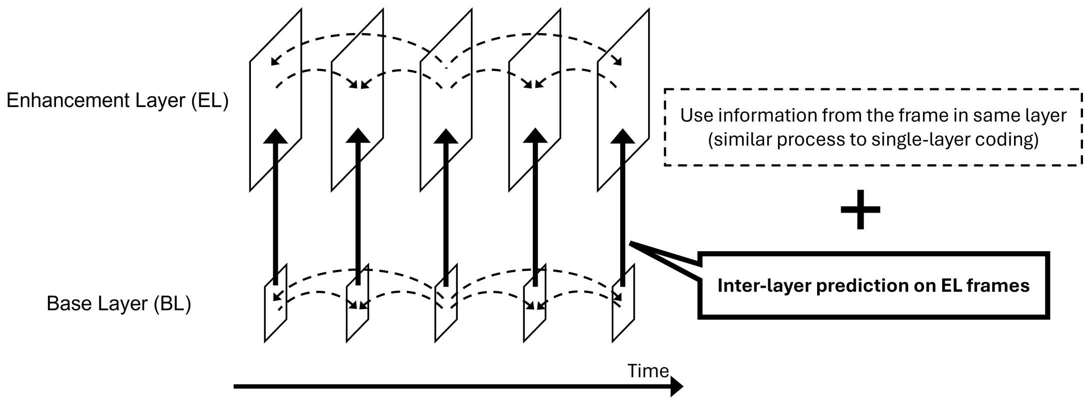

Figure 4 Additional processing for inter-layer prediction in spatial scalable coding

3. Proposed Method

3.1 Overview of the proposed approach

In spatial scalable coding, BL and EL represent the same visual content at different resolutions. Because of this similarity, the block partitioning in the BL is often similar to that required in the EL. By referencing the corresponding region in the BL, the proposed method accelerates block partitioning in the EL. Specifically, it uses the block boundaries and shapes in the BL to reduce the number of partitioning candidates in the EL. This approach effectively reduces computational complexity.

3.2 Block Partitioning Algorithm for the Enhancement Layer Based on the Base Layer

The proposed algorithm efficiently performs block partitioning in the enhancement layer by utilizing both boundary and area information from the base layer. The algorithm's overall structure is detailed in the following subsections.

3.2.1 Overall Algorithm Flow

A flowchart of the proposed algorithm is shown in Figure 5. The algorithm first checks whether the current CU's width or height is 32 pixels or more. If this condition is satisfied, the candidate partitioning patterns are refined using the block boundary information from the base layer. If the condition is not met or boundary-based refinement is not applicable, the algorithm uses the CU shapes in the corresponding region of the base layer for refinement.

3.2.2 Block Partitioning Decision Based on Base Layer Boundaries

The proposed method determines the direction of partitioning in the EL by referencing block boundaries from the BL. Specifically, it detects straight split lines within the corresponding region of the BL, either vertical (dividing the region into left and right) or horizontal (dividing the region into top and bottom), as illustrated in Figure 6.

In the VVC Test Model (VTM), coding units are recursively partitioned, and the boundary line that divides a region into two parts typically corresponds to the initial split applied. The proposed method selects the longest vertical or horizontal split line within the corresponding BL region and uses it to restrict the candidate partitioning directions in the EL. This approach eliminates unnecessary block partitioning pattern evaluation and reduces computational complexity. If neither vertical nor horizontal split lines are detected, the region is considered to have a flat texture. In such cases, no further partitioning is performed in the EL.

3.2.3 Block Partitioning Decision Based on the Base Layer Area

The partitioning direction in the EL is determined based on the total area of vertical and horizontal CUs in the corresponding BL region. In the corresponding BL region, the total area of vertically oriented CUs (Sv) and horizontally oriented CUs (SH) is calculated. Since larger CUs have a greater influence on the overall structure, using area as a metric allows for a more reliable estimation of the dominant partitioning trend. By comparing SV and SH, the method selects the more appropriate split direction. As a result, the algorithm can effectively reduce the number of partitioning patterns of candidates, even in regions with weak directional boundaries or small block sizes.

4. Results and Discussion

The proposed method is evaluated by comparing it with the VTM's spatial scalable coding functionality.

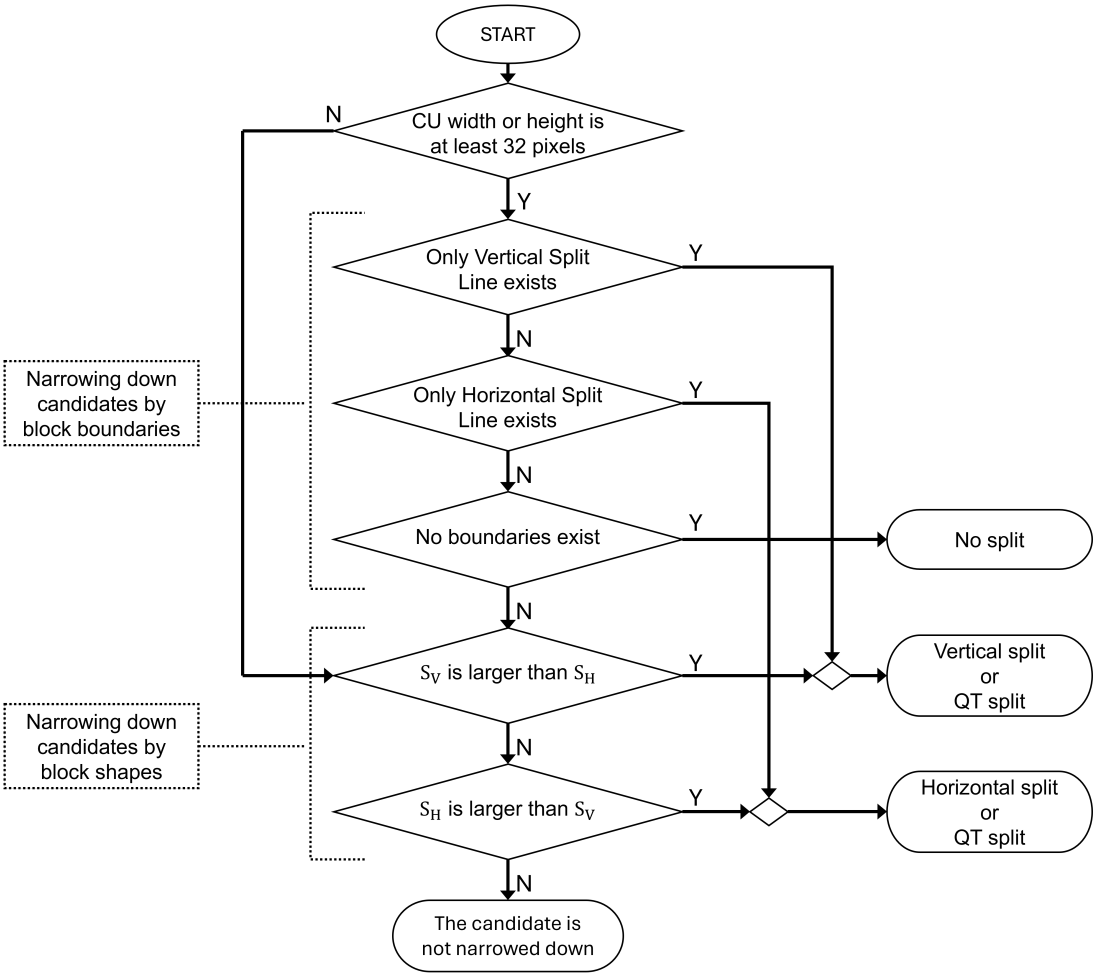

Figure 5 The flowchart of the proposed block partition methods

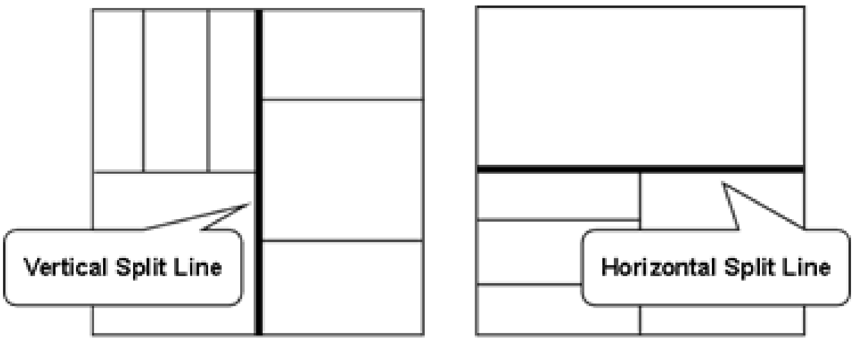

Figure 6 The example of Vertical Split Line and Horizontal Split Line

4.1 Evaluation environment

The experiments were performed using VTM-23.6 reference software on a PC equipped with an AMD Ryzen 9 5950X processor. Encoding was conducted under the random access configuration with quantization parameters (QPs) of 22, 27, 32, and 37. The first 33 frames of several test sequences with different characteristics were used in the experiments(ITE and ARIB, 2017; ITE and ARIB, 2016). The video resolution was set to 2 K (1920×1080) for the BL and 4K (3840×2160) for the EL.

4.2 Measurement Criteria

In this experiment, the evaluation is based on coding efficiency and the rate of reduction of encoding time in EL.The Bjøntegaard Delta (BD)-rate, which represents the average increase in bitrate at the same quality, is used as an indicator of coding efficiency (Bjontegaard, 2001). The BD-rate in this experiment was calculated from the Y-PSNR in EL and the sum of the bitrates of BL and EL.

The encoding time reduction rate in EL, TS_{EL} is calculated as follows:

$$TS_{EL} = \frac{Time_{\text{default}} - Time_{\text{proposed}}}{Time_{\text{default}} - Time_{\text{BL}}} \tag{1}$$

where $Time_{\text{proposed}}$ is the coding time when using the proposed method, $Time_{\text{default}}$ is the encoding time when using the spatial scalable encoding of the original setting, and $Time_{\text{BL}}$ is the coding time of BL. This formula derives the coding time reduction rate for EL only, independent of the coding time of BL.

4.3 Experimental results

TS_{EL} and BD-rate are shown in Table 1. TS_{EL} is the average of the results in the four QPs (22, 27, 32, and 37). It can be seen from Figure 6 that the proposed method achieves an average speed-up of approximately 55% and a minimum of 40%. Table 2 also shows that the BD rate is at most 3.45%, and in particular, paddock, water polo (scrolling text), drama (apple), and drama (bouquet) have a BD rate of less than 3%, which is a slight degradation.

Table 1 Time Saving and BD-rate in Enhancement Layer

Sequence	Time Saving in Enhancement Layer (%)	BD-rate (%)
Japanese Maple	66.64	3.45
Layered Kimono	52.07	1.24
Horse race (dirt)	41.31	3.25
Horse race (finish)	62.57	3.07
Paddock	53.17	1.69
Marathon (start)	52.65	3.34
Water polo (scrolling text)	55.96	2.32
Drama (apple)	52.35	2.67
Drama (bouquet)	57.79	2.66

The proposed method achieved approximately a 55% reduction in encoding time across all test sequences. For sequences such as Paddock, Water polo (scrolling text), Drama (apple), and Drama (bouquet), the increase in BD-rate remained below 3%, and no significant issues were observed during subjective visual verification. In contrast, sequences including Japanese maple, horse race (dirt), horse race (finish), and marathon (start) showed increases in the BD rate exceeding 3%. For instance, Japanese maple contains numerous thin branches and highly detailed leaf structures, which generate fine, high-frequency patterns that normally require detailed block partitioning. Similarly, the horse race (dirt) sequence features irregular granular textures of the dirt field, resulting in rapidly changing local variations that demand flexible and fine-grained splits. Because the proposed method restricts the candidate partitions based on BL-derived information, these sequences suffer from reduced ability to adaptively refine the partition structure, resulting in greater coding efficiency degradation. As a result, the proposed method shows a slight degradation in the BD rate, but it remains effective owing to the significant time saving contribution.

5. Conclusions

In this study, a speed-up method for spatial scalable coding in VVC is proposed, utilizing the block boundary and shape information of the BL. The experimental results show that the proposed method can achieve an average speed-up of approximately 55%, and the coding efficiency reduction is reduced to a maximum of 3.45%. This study shows the feasibility of a more efficient live video streaming method. In the future, we would like to develop even more efficient speed-up methods that consider texture fineness and motion information.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Numbers JP23K24827, JP22H03571, JP24K20751, Support Center for Advanced Telecommunications Technology Research, Hoso Bunka Foundation, and Telecommunication Advancement Foundation, ROIS NII Open Collaborative Research 2023-(23S0202), 2024-(24S0103, 24S0108).

Author Contributions

Yo Koura : Conceptualization, Methodology, and Original Draft Preparation. Naoya Niwa : Survey, Review, and Editing. Hiroe Iwasaki : Contributed to the enrichment of concepts and discussion.

Conflict of Interest

The authors declare no conflicts of interest.

References

Amestoy, T., Sidaty, N., Hamidouche, W., Philippe, P., & Menard, D. (2023). Video quality assessment and coding complexity of the versatile video coding standard. *arXiv preprint*.

Bjontegaard, G. (2001). *Calculation of average psnr differences between rd-curves* (tech. rep. No. VCEG-M33). ITU-T. https://www.itu.int/md/T01-SG16-010402-TD-0033

Bonnineau, C., Hamidouche, W., Fournier, J., Sidaty, N., Travers, J. F., & Deforges, O. (2022). Perceptual quality assessment of hevc and vvc standards for 8k video. *IEEE Transactions on Broadcasting*, *68*(1), 246–253. https://doi.org/10.1109/TBC.2022.3140710

Boyce, J. M., Ye, Y., Chen, J., & Ramasubramonian, A. K. (2015). Overview of shvc: Scalable extensions of the high efficiency video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *26*(1), 20–34. https://doi.org/10.1109/TCSVT.2015.2461951

Bross, B., Chen, J., Ohm, J. R., Sullivan, G. J., & Wang, Y. K. (2021a). Developments in international video coding standardization after avc, with an overview of versatile video coding (vvc). *Proceedings of the IEEE*, *109*(9), 1463–1493. https://doi.org/10.1109/JPROC.2020.3043399

Bross, B., Chen, J., Ohm, J. R., Sullivan, G. J., & Wang, Y. K. (2021b). Overview of the versatile video coding (vvc) standard and its applications. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*(10), 3736–3764. https://doi.org/10.1109/TCSVT.2021.3101953

Business Research Insights. (2024). Live streaming market size, share, and growth report 2024–2033 [Accessed 4 July 2025]. https://www.businessresearchinsights.com/market-reports/live-streaming-market-111696

Cong, Z., Liu, J., & Manchanda, P. (2021). The role of "live" in livestreaming markets: Evidence using orthogonal random forest. *arXiv*.

Feldmann, A., Gasser, O., Lichtblau, F., Pujol, E., Poese, I., Dietzel, C., Wagner, D., Wichtlhuber, M., Tapiador, J., Vallina-Rodriguez, N., Hohlfeld, O., & Smaragdakis, G. (2020). The lockdown effect: Implications of the covid-19 pandemic on internet traffic. *Proceedings of the ACM Internet Measurement Conference (IMC '20)*, 1–18. https://doi.org/10.1145/3419394.3423658

Fischer, W. (2020). Video coding (mpeg-2, mpeg-4/avc, hevc). In *Digital video and audio broadcasting technology* (pp. 125–175). Springer. https://doi.org/10.1007/978-3-030-32185-7_7

Fraunhofer Heinrich-Hertz-Institut. (2014). Hevc scalability extension (shvc) [Accessed 4 July 2025]. https://hevc.hhi.fraunhofer.de/shvc

Grois, D., Giladi, A., Choi, K., & Park, M. W. (2021). Performance comparison of emerging evc and vvc video coding standards with hevc and av1. *SMPTE Motion Imaging Journal*, *130*(4), 1–12. https://doi.org/10.5594/M001916

Hayase, K., Fujii, H., Bandoh, Y., & Jozawa, H. (2012). Recent advances on scalable video coding. *IEICE Transactions on Fundamentals*, *E95.A*(8), 1230–1239. https://doi.org/10.1587/transfun.E95.A.1230

Huang, Y. W., An, J., Huang, H., Li, X., Hsiang, S. T., Zhang, K., Gao, H., Ma, J., & Chubach, O. (2021). Block partitioning structure in the vvc standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *31*(10), 3818–3833. https://doi.org/10.1109/TCSVT.2021.3088134

ITE and ARIB. (2016). *Ultra-high definition/wide-color-gamut standard test sequences – series a manual* (tech. rep.). The Institute of Image Information, Television Engineers (ITE), Association of Radio Industries, and Businesses (ARIB).

ITE and ARIB. (2017). *Ultra-high definition/wide-color-gamut standard test sequences – series b manual* (tech. rep.). The Institute of Image Information, Television Engineers (ITE), Association of Radio Industries, and Businesses (ARIB).

Jhang-Li, J. H., & Liou, J. H. (2023). An analysis of operating strategy for a video live streaming platform: Advertisement, advertorial, and donation. *Information Technology & Management*, 1–18. https://doi.org/10.1007/s10799-023-00387-x

Marketing Scoop. (2023). The latest online video consumption statistics in 2024: An in-depth analysis [Accessed 4 July 2025]. https://www.marketingscoop.com/small-business/online-video-consumption-statistics

Marquant, G., Salmon-Legagneur, C., Urban, F., & de Lagrange, P. (2022). Spatial scalability with vvc: Coding performance and complexity. *Proceedings of SPIE*. https://doi.org/10.1117/12.2632821

Mercat, A., Makinen, A., Sainio, J., Lemmetti, A., Viitanen, M., & Vanne, J. (2021). Comparative rate-distortion-complexity analysis of vvc and hevc video codecs. *IEEE Access*, *9*, 67813–67828. https://doi.org/10.1109/ACCESS.2021.3077116

Ohm, J. R. (2005). Advances in scalable video coding. *Proceedings of the IEEE*, *93*(1), 42–56. https://doi.org/10.1109/JPROC.2004.839611

Pakdaman, F., Adelimanesh, M. A., Gabbouj, M., & Hashemi, M. R. (2020). Complexity analysis of next-generation vvc encoding and decoding. *2020 IEEE International Conference on Image Processing (ICIP)*, 3134–3138. https://doi.org/10.1109/ICIP40778.2020.9190983

Peroni, L., & Gorinsky, S. (2025). An end-to-end pipeline perspective on video streaming in best-effort networks: A survey and tutorial. *ACM Computing Surveys*, *57*(12), 1–47. https://doi.org/10.1145/3742472

Sandvine. (2023). 2023 global internet phenomena report [Accessed 4 July 2025]. https://www.advanced-television.com/2023/01/19/report-23-jump-in-online-traffic-in-2023/

Schwarz, H., Marpe, D., & Wiegand, T. (2007). Overview of the scalable video coding extension of the h.264/avc standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *17*(9), 1103–1120. https://doi.org/10.1109/TCSVT.2007.905532

Shahid, Z., Chaumont, M., & Puech, W. (2017). Scalable video coding [Accessed 4 July 2025]. https://web.archive.org/web/20170829172743/http://www.lirmm.fr/~chaumont/Publications/Shahid2017_ScalableVideoCoding.pdf

Sullivan, G. J., Boyce, J. M., Chen, Y., Ohm, J. R., Segall, C. A., & Vetro, A. (2013). Standardized extensions of high efficiency video coding. *IEEE Journal of Selected Topics in Signal Processing*, *7*(6), 1001–1016. https://doi.org/10.1109/JSTSP.2013.2283657

Sullivan, G. J., Ohm, J. R., Han, W. J., & Wiegand, T. (2012). Overview of the high efficiency video coding (hevc) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *22*(12), 1649–1668. https://doi.org/10.1109/TCSVT.2012.2221191

Sun, H., Vetro, A., & Xin, J. (2007). An overview of scalable video streaming. *Wireless Communications and Mobile Computing*, *7*(2), 159–172. https://doi.org/10.1002/wcm.471

Takahashi, T., & Yamada, K. (2008). *Information processing device and method, and program* (JP2008516556A) [Accessed 4 July 2025]. https://patents.google.com/patent/JP2008516556A/en

Timmerer, C., Amirpour, H., Tashtarian, F., Afzal, S., Rizk, A., Zink, M., & Hellwagner, H. (2025). Http adaptive streaming: A review on current advances and future challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *21*(7), 1–27. https://doi.org/10.1145/3736306

Tissier, A., Mercat, A., Amestoy, T., Hamidouche, W., Vanne, J., & Menard, D. (2019). Complexity reduction opportunities in the future vvc intra encoder. *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, 1–6. https://doi.org/10.1109/MMSP.2019.8901754

Unanue, I., Urteaga, I., Husemann, R., J, D., Roesler, V., Rodriguez, A., & Sanchez, P. (2011). A tutorial on h.264/svc scalable video coding and its tradeoff between quality, coding efficiency and performance. In *Recent advances on video coding*. InTech. https://doi.org/10.5772/19227

Wang, Z., Zhang, J., & Li, H. (2010). Spatially scalable video coding with an efficient two-layered architecture. *Multimedia Tools and Applications*, *48*, 247–265. https://doi.org/10.1007/s11042-009-0327-3

Wiegand, T., Sullivan, G. J., Bjontegaard, G., & Luthra, A. (2003). Overview of the h.264/avc video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(7), 560–576. https://doi.org/10.1109/TCSVT.2003.815165

Zhang, T., & Mao, S. (2019). An overview of emerging video coding standards. *GetMobile: Mobile Computing and Communications*, *22*(4), 13–20. https://doi.org/10.1145/3325867.3325873

Zheng, J. (2025). Study on the future development of streaming media and traditional media. *Communications in Humanities Research*, *58*, 19–24. https://doi.org/10.54254/2753-7064/2025.22075