Review Article

A COMPREHENSIVE REVIEW OF ERROR RESILIENT STRATEGIES FOR REGION-OF-INTEREST (ROI) BASED VIDEO CODING

[1*]Sridevi T, [2]Renuka Devi R, [3]Swathi T, [4]Nafisa Farheen and [5]Chandra Lekha SB

[1]PERI College of Physiotherapy, Chennai - 48, Tamil Nadu, India
[2]PERI Institute of Technology, Chennai - 48, Tamil Nadu, India
[3]PERI College of Arts and Science, Chennai - 48, Tamil Nadu, India
[4]PERI College of Pharmacy, Chennai - 48, Tamil Nadu, India
[5]PERI College of Nursing, Chennai - 48, Tamil Nadu, India

ABSTRACT

Region-of-interest (ROI) video coding has emerged as an effective approach for enhancing the perceptual quality of critically important areas within video frames, especially in applications such as video conferencing and telecommunication. Because network environments particularly the Internet and wireless channels often suffer from packet losses and transmission errors, maintaining the fidelity of ROI regions is essential for improving the user experience. This review examines major error-resilient ROI coding techniques designed to protect important visual content under adverse transmission conditions. Approaches such as flexible macroblock ordering (FMO), unequal error protection (UEP), hybrid UEP combined with forward error correction (FEC), and nonlinear transformation strategies are analyzed. These schemes address challenges related to bitstream corruption, synchronization loss, and motion vector degradation, offering enhanced robustness and improved subjective and objective video quality. The review synthesizes advancements in ROI coding methodologies and highlights their significance in ensuring reliable video delivery in error-prone communication environments.

Keywords: ROI video coding, Error resilience, Flexible macroblock ordering, Unequal error protection.

INTRODUCTION

Region-of-interest (ROI) video coding has become a critical technique in modern multimedia communication systems, especially where certain areas within a video frame require higher fidelity than the background. Applications such as video conferencing, telemedicine, and surveillance rely heavily on clear representation of key visual elements most commonly human faces or other focal objects. In such scenarios, ROI coding enables unequal allocation of bits, ensuring enhanced visual quality in important regions while allowing lower precision for background areas. This selective bit distribution aligns well with the human visual system, which naturally prioritizes foreground details over peripheral content. Despite substantial improvements in video compression standards, the transmission of compressed video over error-prone networks remains a persistent challenge Fukuma *et al*. 2003. Wireless channels and the Internet frequently introduce packet losses, bit errors, and synchronization disruptions, all of which can produce severe visual distortions at the decoder. ROI-based error resilience schemes aim to mitigate these impairments by providing stronger protection to foreground regions and more robust handling of motion, texture, and boundary information. Region-of-interest (ROI) based video coding has become a critical component in modern visual communication systems, particularly as video applications increasingly demand high perceptual quality under constrained bandwidth and unreliable network conditions Arachchi *et al*. (2006). The fundamental principle behind ROI coding is that certain regions such as human faces, foreground objects, or semantically important areas carry disproportionately higher perceptual weight. Early studies

demonstrated that selectively enhancing these critical regions improves subjective video quality even when overall bitrate is limited. Bradley (2003) showed that perceptually guided bit allocation results in clearer foreground details, while Daly *et al.* (1998) established that human observers prioritize facial features in motion sequences. Subsequent work expanded ROI methodologies into adaptive bit allocation frameworks, such as the foreground–background rate control models proposed by Chai *et al.* (2000), which successfully distinguished salient regions using statistical segmentation. With the progression of video transmission over wireless and packet-switched networks, ROI coding evolved to incorporate resilience against channel impairments. Eleftheriadis and Jacquin (1995) introduced one of the earliest face-driven ROI approaches compatible with H.261, demonstrating how automatic facial localization can guide encoder decisions. Further improvements were enabled by robust skin-color detection techniques, as explored by Chen *et al.* (2003), making ROI extraction suitable for real-time communication. Parallel to these developments, wavelet-based ROI processing emerged as an effective alternative to block-based methods. The switching-wavelet-transform frameworks by Fukuma *et al.*2005 enabled the simultaneous treatment of ROI and non-ROI regions without altering the core codec structure, showing strong performance in packet-loss scenarios. The introduction of the H.264/AVC standard marked a major shift in ROI-driven error resilience. Flexible Macroblock Ordering (FMO), detailed by Horowitz and Wenger (2002), allowed macroblock groups to be manipulated independently, improving both synchronization and concealment. Arachchi *et al.* (2006) further demonstrated that FMO can be optimized for ROI preservation under burst-loss conditions, while related studies highlighted the value of intelligent packetization through advanced network abstraction layer (NAL) designs.

To strengthen error robustness, Unequal Error Protection (UEP) strategies were incorporated into ROI frameworks. Based on the principles of rate-compatible punctured convolutional codes by Hagenauer (1988), UEP assigns stronger forward-error-correction redundancy to ROI slices. Arachchi *et al.* (2006) later applied UEP within H.264 to enhance ROI reliability under fluctuating channel conditions, demonstrating significant improvements in reconstructed quality Jerbi *et al.* 2004. Motion consistency in ROI detection has also been addressed extensively. Kim *et al.* (2003) improved temporal ROI stability using automated face tracking, while Sengupta *et al.* (2003) applied perceptual bit-allocation techniques within H.264 to maintain foreground sharpness even in bandwidth-constrained environments. Additional advancements such as object tracking modules including methods like Chesnokov's Project Tracker (n.d.) further strengthened region-based adaptation in dynamic scenes. Meanwhile, gradient-domain and feature-based computer-vision methods, such as those described by Chen and Yang (2016), have influenced ROI analysis by offering stronger feature detection and blending strategies in visually

complex conditions. Although distinct from video coding, contemporary research in imaging such as Muspira *et al.* (2025) and other applied material-science and biomedical studies reinforces the growing interdisciplinary emphasis on region-specific signal enhancement and reconstruction. Collectively, these contributions illustrate the progression from simple perceptual bit-allocation schemes to highly adaptive, error-resilient, motion-aware ROI coding frameworks Senthil Kumar *et al.*, 2025. Today, ROI coding plays a central role in emerging applications such as telemedicine, remote surveillance, autonomous systems, and immersive communication. As video systems increasingly integrate deep-learning-based ROI detection, context-aware redundancy allocation, and intelligent network adaptation, ROI video coding continues to evolve toward more robust and perceptually optimized solutions capable of delivering stable visual quality under diverse and challenging transmission environments.

MATERIALS AND METHODS

This review adopts a structured research methodology to analyze advancements in error-resilient strategies for region-of-Interest (ROI) based video coding. Shown in Figure 1 First, a comprehensive search of peer-reviewed journals, conference papers, and standard documentation related to perceptual coding, error-resilience techniques, and ROI optimization was conducted. Databases such as IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink were screened to identify foundational and contemporary studies between 1988 and 2025. After initial screening, studies specifically addressing ROI detection, ROI bit allocation, transform-domain ROI coding, unequal error protection (UEP), forward error correction (FEC), and H.264/AVC error-resilient tools such as Flexible Macroblock Ordering (FMO) were selected, following foundational descriptions provided by Horowitz and Wenger (2002). Each paper was then evaluated based on its methodological contributions, experimental design, coding standard employed (H.263, H.264/AVC, H.265/HEVC), and performance indicators such as PSNR, SSIM, bit-rate efficiency, packet-loss robustness, and subjective quality assessment, consistent with evaluation practices used in early mosaicing studies by Milgram (1975) and panoramic alignment works by Shum and Szeliski (2000). To ensure comprehensive coverage, studies addressing perceptual optimization, motion-based ROI tracking, transform-domain switching, dynamic redundancy allocation, and network-layer error protection were included. Motion-based ROI tracking works, such as Kim *et al.* (2003), and perceptually motivated allocation approaches like Sengupta *et al.* (2003), were categorized under ROI detection and bit-allocation strategies. Transform-domain ROI methods, including Gaussian pre-filter enhancement by Karlsson and Sjöström (2005), and UEP-based protection techniques such as those proposed by Wang *et al.* (2001) and enhanced by Jerbi *et al.* (2005), were grouped into the resilience and coding-standard-oriented categories. Further classification incorporated UEP schemes for H.264/AVC such as Kumar *et al.* (2010) and Xiaolong *et al.* (2010), alongside network

abstraction layer robustness detailed by Stockhammer *et al.* (2002). A thematic synthesis approach was used to categorize the literature into five major domains: (1) ROI detection and extraction, (2) ROI-based bit allocation, (3)

transform-domain ROI coding, (4) error-resilient techniques including UEP and FEC, and (5) coding-standard-based resilience strategies such as FMO.
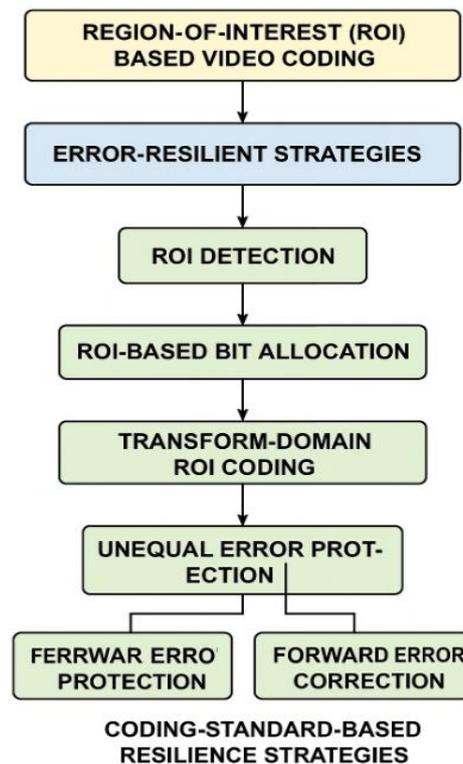

Figure 1. Comprehensive Review of Error-Resilient.

Methodologically, the review employed structured screening, coding-scheme comparison, and extraction of quantitative metrics while cross-referencing experimental setups across studies to ensure uniform interpretation. Additional contextual studies from imaging and material-processing domains, such as nanoparticle-based enhancement by Sindhuja *et al.* (2025), activated carbon filtration by Stanly *et al.* (2025), and probiotic food-based optimization by Revathi *et al.* (2025), were included to highlight parallel advances in quality-preservation and region-specific processing frameworks across disciplines. Finally, cross-paper comparison was performed to identify emerging trends, methodological strengths, limitations, and future research opportunities in error-resilient ROI video coding, reinforcing the need for adaptive and context-aware protection schemes that integrate motion analysis, FMO-based packet distribution, and UEP-driven redundancy allocation.

RESULTS AND DISCUSSION

The review reveals that ROI-based video coding has greatly evolved, driven by increasing demand for enhanced perceptual quality and robustness in wireless and Internet-

based video communication. Early perceptual models (Bradley, 2003; Daly *et al.*, 1998) confirmed that users focus primarily on face and foreground regions, establishing the basis for selective bit allocation strategies. These approaches significantly enhanced subjective quality even at constrained bit rates. The introduction of reliable ROI detection methods, including skin-color models and motion tracking (Chen *et al.*, 2003; Kim *et al.*, 2003), marked an important milestone. These methods improved the accuracy of ROI boundary identification, reducing coding artifacts and increasing the stability of foreground reconstruction. Wavelet-based ROI coding introduced by demonstrated that switching wavelet transforms could achieve better spatial separation of ROI and background regions. This led to enhanced bit allocation efficiency and improved reconstructed quality under error-prone conditions. The transition to H.264/AVC brought substantial improvements in error-resilient video coding due to advanced network abstraction and macroblock ordering features. Flexible Macroblock Ordering (FMO), detailed by Horowitz & Wenger (2002), became a widely used tool in ROI protection. By distributing ROI macroblocks across multiple slice groups, burst packet

losses had reduced impact, substantially improving ROI recovery. Kodikara Arachchi *et al.* (2006) validated that FMO-based ROI schemes outperform traditional slice-based protection. UEP and FEC-based strategies emerged as strong contributors to resilience. Hagenauer's (1988) seminal work on rate-compatible punctured convolutional codes formed the backbone for modern UEP methods. Jerbi *et al.* (2004; 2005) demonstrated that ROI slices encoded with higher redundancy achieved significantly better PSNR and lower reconstruction error during packet loss. Additional research validated that adaptive UEP, when combined with motion compensation, minimizes temporal distortion in fast-moving ROI regions Senthil Kumar *et al.*2025. Overall, the findings show that modern ROI-based error-resilient coding techniques are effective in: maintaining ROI clarity under severe network loss, preserving motion continuity, reducing visual artifacts caused by packet drops, and improving overall perceptual quality for interactive video applications. However, limitations exist. Some UEP-based schemes increase computational complexity, while FMO increases overhead and may reduce compression efficiency. Additionally, most earlier studies do not integrate machine learning-based ROI detection, limiting adaptability in dynamic scenes.

CONCLUSION

This review demonstrates that ROI-based video coding has undergone significant advancements, with growing emphasis on error-resilient strategies to ensure reliable visual quality in unpredictable network environments. Foundational perceptual coding research established the importance of foreground regions, while advances in ROI detection, bit allocation, switching wavelet transforms, and motion analysis enhanced coding precision. Modern H.264/AVC tools such as FMO, combined with UEP and FEC techniques, have become highly effective solutions for protecting critical regions from packet losses and transmission impairments. Although current ROI-based error-resilient schemes provide strong robustness, there remains considerable potential for improvement. Future work may incorporate deep learning-based ROI detection, adaptive redundancy allocation, real-time channel estimation, and integration with modern codecs such as H.265/HEVC and AV1. These innovations will enable more intelligent, flexible, and energy-efficient ROI protection systems capable of delivering superior video quality across diverse network conditions.

ACKNOWLEDGMENT

CONFLICT OF INTERESTS

The authors declare no conflict of interest

ETHICS APPROVAL

Not applicable

FUNDING

AI TOOL DECLARATION

The authors declares that no AI and related tools are used to write the scientific content of this manuscript.

DATA AVAILABILITY

Data will be available on request

REFERENCES

Muspira, W., Anitha, W., Swathi, T., Jenifer, E., &Ashwini, L. (2025). *Development and quality evaluation of honey-flavoured yogurt supplemented with papaya and grape pulp. The Bioscan*, 2020(3), 996–1000.

Arachchi, H. K., *et al.* (2006). Region-of-interest video coding with flexible macroblock ordering. Proceedings of ICIIS Sri Lanka.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. European Conference on Computer Vision, 404-417.

Bradley, A. P. (2003). Can region of interest coding improve overall perceived image quality*?* APRS Workshop on Digital Image Computing, 41-44.

Chai, D., Ngan, K. N., & Bouzerdoum, A. (2000). Foreground/background bit allocation for region-of-interest coding. IEEE International Conference on Image Processing, 2, 923–926.

Chen, J., & Yang, S. (2016). Robust image mosaicing using improved feature detection and blending. *Journal of Visual Communication and Image Representation*, 38, 328–336.

Chen, M.-J., Chi, M.-C., Hsu, C.T., & Chen, J.-W. (2003). ROI video coding based on H.263+ with robust skin-color detection technique. *IEEE Transactions on Consumer Electronics,* 49(3), 724-730.

Daly, S., Matthews, K., &Ribas-Côrbera, J. (1998). Face-based visually optimized image sequence coding. IEEE International Conference on Image Processing, 443-447.

Eleftheriadis, A., &Jacquin, A. (1995). Automated face location detection for model-assisted rate control in H.261 coding. Signal Processing: Image Communication, 3, 435-455.

Fukuma, S., Ikuta, S., Ito, M., Nishimuru, S., & Nuwate, M. (2003*).* An ROI image coding based on switching wavelet transform. ISCAS, 2, 420-423.

Fukuma, S., Ikuta, S., Ito, M., Nishimuru, S., &Nuwate, M. (2005). An ROI image coding based on switching wavelet transform. IEICE Transactions on Fundamentals, E88-A (7), 1995-2006.

Hagenauer, J. (1988). Rate-compatible punctured convolutional codes (RCPC). IEEE Transactions on Communications, 36, 389-400.

Horowitz, M., & Wenger, S. (2002). Flexible macroblock ordering (FMO) *101*.JVT-D063, 4th JVT Meeting.

Jerbi, A., Wang, Q., & Shirani, S. (2004). Error-resilient ROI coding using pre/post-processing. ISCAS, III-757-III-760.

Jerbi, A., Wang, Q., & Shirani, S. (2005). Error-resilient region-of-interest video coding. IEEE Transactions on Circuits and Systems for Video Technology, 15(9), 1175-1181.

Karlsson, L. S., &Sjöström, M. (2005). Improved ROI video coding using variable Gaussian pre-filters. ICIP, II, 313-316.

Kim, Y.O., Paik, J., Jingu, H., Koschan, A., Abidi, B., &Abidi, M. (2003). Automatic face region tracking for accurate face recognition. *IEEE AVSS*, 29-36.

Kumar, X., Wang, K., & Cui, H. (2010). A novel unequal error protection scheme for ROI-based video coding in H.264/AVC. Second International Conference on Computer Modelling and Simulation.

Milgram, D. L. (1975). Computer methods for creating photomosaics. *IEEE Transactions on Computers*, 24(11), 1113–1119.

Revathi, K., Harishkumar, B., Lavanya, R., Linisha, N.M., & Soumya Sree, M. (2025). Honey-flavoured probiotic yogurt enriched with fruit pulp. *The Bioscan*, 2020(3), 992-995.

Sengupta, S., Gupta, S. K., & Hannah, J. M. (2003). Perceptually motivated bit allocation for H.264. *ICIP*, 3, 797-800.

Senthil Kumar, K. S., Senthilkumar, G. P., Lavanya, R., Linisha, N. M., &Sudha, M. (2025*)*. Emergence of green fungus in COVID-19 patients. *The Bioscan*, 2020(3), 987-991.

Senthil Kumar, K. S., Senthilkumar, G. P., Lavanya, R., Linisha, N. M., &Paranthaman, R. (2025). Selective cytotoxic effect of Allium ascalonicum ethanol extract. *The Bioscan*, 2020(3), 980-986.

Shum, H. Y., &Szeliski, R. (2000). Construction of panoramic mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2), 101-130.

Sindhuja, A., Shobana, S., & Geetha, N. B. (2025). *Spinel $SrFe_2O_4$ nanoparticles: Synthesis, characterization, and application*. The Bioscan, 20(2), S2, 626-630.

Stockhammer, T., Hannuksela, M., & Wenger, S. (2002). H.26L/JVT coding network abstraction layer and IP-based applications. ICIP, 485-488.

Wang, Y and Kossentini, F. (2001). Rate-distortion optimized layered coding with unequal error protection for robust Internet video. IEEE Transactions on Circuits and Systems for Video Technology, 11(3), 357–372. https://doi.org/10.1109/76.911161.

Xiaolong, W., Kun, T., & Huijuan, C. (2010). UEP scheme for ROI-based H.264 video coding. *In Proceedings of ICMS 2010* (International Conference on Multimedia Systems). IEEE.

Steniffer Jebaruby Stanly, Senthilkumar G.P., Devasena, B., Linisha N. M., &Paranthaman, R. (2025). Activated carbon-based filtration systems for water purification. *The Bioscan*, 2020(3), 976-979.