SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper
Open Access

A Wavelet-Guided Content-Adaptive Framework For Fast Intra-Mode Prediction In H.264/AVC

Umesh W. Kaware[1] , Sanjay M. Gulhane[2]

[1]Jawaharlal Darda Institute of Engineering and Technology Yavatmal, Maharashtra, India-445001
uwkaware@gmail.com Orcid iD-0009-0001-7947-9418
[2]Pravara Rural Engineering College Loni, Ahmednagar, Maharashtra, India-413736 sanjay.gulhane@pravara.in Orcid iD-0000-0003-2783-5940

Abstract
Intra-mode decision is an expensive step in H.264/AVC encoding, as prediction modes need to be evaluated to find the mode with the least rate-distortion cost (RDC). In this paper, we propose a Wavelet-Guided Content-Adaptive Intra-Mode Prediction (WG-CAMP) approach to reduce such intra-mode evaluation in the coding process while still offering competitive coding performance. The proposed approach employs a two-level Daubechies-2 (db2) wavelet decomposition on each 16 × 16 luminance block and incorporates horizontal, vertical, and diagonal wavelet information in the space using a structure tensor to estimate the orientation of the dominant block. The estimated orientation is mapped to the three nearest directional H.264/AVC intra modes, while vertical, horizontal, and DC modes are kept as safety candidates and a small set of content-adaptive candidate information is used for rate-distortion optimisation (RDO). We compare our proposed framework with Full RD Search and SATD-based Fast Search algorithms on Coastguard, Container, Foreman, and News CIF sequences with QP of 20, 24, 28, and 32. Experiments showed that WG-CAMP reduces the average encoding time from 72.07 s for Full RD Search to 25.87 s, bringing 63.73% time savings and a 2.79× speedup and improves the average PSNR from 39.443 dB to 39.491 dB and SSIM from 0.95214 to 0.95273. In particular, the proposed method introduces only a 0.78% BDBR increase compared to Full RD Search (2.11% for SATD Fast Search) while keeping the Full-RD optimal mode in approximately 81.03% of blocks with an average candidate set of about 5.04 modes. These results prove that wavelet-domain orientation provides a good low-complexity descriptor of content-adaptive intra-mode prediction, and the coding efficiency and computational complexity trade-off can be made with WG-CAMP. Our WG-CAMP framework is a good candidate for low-complexity H.264/AVC encoding in applications where less encoding time is needed, and the RD performance is not significantly affected.

Keywords: H.264/AVC, Intra-Mode Prediction, Wavelet Transform, Content-Adaptive Coding, Fast Mode Decision, RDO, and Computational Complexity.

1. Introduction

Video compression is an important technology for multimedia communication, streaming, surveillance, video conferencing, and storage applications, where continuously increasing video resolution and quality must be efficiently managed to overcome bandwidth and computational limitations. H.264/AVC is a popular video coding standard that is efficient in terms of compression, quality, memory, and implementation. In particular, H.264/AVC can utilise the spatial redundancy within a frame by predicting the current block from the previous reconstruction. To achieve high coding efficiency using rate-distortion optimisation (RDO) the encoder has to search multiple prediction modes and choose the least-cost RD mode with the best prediction of that. This extensive search significantly increases the coding complexity of H.264/AVC and makes it difficult for the conventional H.264/AVC coding to be used in real-time applications. Recent developments in video coding are very much about reducing the encoder complexity while maintaining coding efficiency, with video applications moving to high-resolution and low-latency processing.

The intra-mode decision process is therefore an important target for optimisation. In H.264/AVC, the possibility of different directional and non-directional prediction modes allows the encoder to effectively represent different spatial structures. However, even if the data is predicted, the full RDO process requires many more predictions and transform and rate-cost calculations. Wang and Cheng [1] solved this problem by using an early termination approach based on the most probable mode (MPM) and the sum of absolute transformed difference (SATD). In this approach, the spatially similar neighbouring blocks are likely to select similar modes, and a good prediction will generally result in less residual. The proposed method evaluates the MPM and a small number of SATD-selected candidates to reduce the number of mode evaluations while maintaining almost the same PSNR and bitrate as the exhaustive search. In the same way, Mathew et al. [2] presented a computationally efficient intra- and inter-mode decision framework for H.264/AVC, for which unlikely prediction modes are eliminated prior to RDO using prediction direction and SATD information. Their experiment achieved a 50.73% reduction in coding time and only 0.13 dB PSNR degradation and 1.13% bitrate increase, indicating that we can achieve significant complexity reduction with appropriate candidate-mode pruning.

The need for efficient mode selection has also motivated research in scalable and adaptive video coding. Balaji et al. [3] have proposed an optimal mode selection algorithm for H.264 scalable video coding based on pixel orientation. It is based on balancing the encoding time, PSNR and bitrate instead of just optimising the computational complexity. Compared to the H.264/SVC reference software, the algorithm achieved 57.44% encoding time reduction with 0.43 dB PSNR improvement and 0.23% compression improvement. These findings demonstrate that directional properties and spatial orientation may provide useful information to reduce the mode-search space, but recent developments in video prediction have also explored more flexible spatial and spatio-temporal representation in recent years. Seiler et al. [4] proposed a two-stage spatio-temporal prediction technique using temporal information first and then refining the prediction with spatially adjacent reconstructed samples. Using H.264/AVC, they concluded that their method can reduce data rates by up to nearly 13% for the evaluated sequences and iteratively refine the prediction significantly as compared to the previous approach.

More recently, research in video coding has moved toward adaptive and data-driven prediction mechanisms. Lu et al. [5] proposed an adaptive online-learning framework for intra-prediction based on separable path graph transforms. The approach learned transforms from previously reconstructed blocks with similar properties and applied them to improve residual decorrelation for complex texture regions. They compared their approach to H.264/AVC intra-coding tools and found transform efficiency to be higher for intra-predicted texture data than DCT. More generally, learned approaches have also shown that machine learning can capture and predict the local directional texture characteristics and therefore coding decisions. Chen et al. [6] developed a learned fast intra-coding framework for HEVC where it leveraged a shallow convolutional neural network to extract the directional texture features and predict a smaller number of RDO candidates. They found up to 75.2% reduction in the intra-mode encoding complexity with a low RD loss, demonstrating the potential of predictive rather than full mode evaluation. Learning-based approaches are largely directed towards newer coding standards, but they do bring a new challenge for researchers to explore lightweight content-adaptive prediction approaches for H.264/AVC when computational resources and encoding latency are limited.

The main contribution of this work is the development of a lightweight, content-adaptive intra-mode prediction framework for H.264/AVC that exploits wavelet-domain directional information to reduce the number of modes subjected to full RD evaluation. The proposed method extracts multi-level directional information from the discrete wavelet transform (DWT) coefficients using a structure-tensor formulation and maps the estimated orientation to the most relevant directional H.264/AVC intra modes. Three directional candidates are combined with vertical, horizontal, and DC safety modes, thereby retaining potentially optimal modes while substantially reducing the search space. Two orientation strategies are investigated, namely direct edge orientation and its perpendicular (+90°) direction, and their effectiveness is evaluated against exhaustive Full-RD and conventional SATD-based fast search. In addition to PSNR, SSIM, bitrate, and encoding time, the framework provides a detailed analysis of full-RD optimal-mode retention, mode-wise hit rate, candidate count, mode-selection confusion, time saving, speedup, and Bjøntegaard Delta Bitrate (BDBR), enabling a comprehensive assessment of the trade-off between coding efficiency and computational complexity.

The significant contributions can be summarised as:

1. A lightweight DWT-based directional feature extraction mechanism is introduced to estimate the dominant spatial orientation of each block and adaptively identify the most promising H.264/AVC intra-prediction modes.
2. The predicted directional modes are combined with three essential safety modes (vertical, horizontal, and DC), substantially restricting the candidate set before full RD evaluation. Both direct edge orientation and perpendicular orientation are investigated to determine the more effective mode-mapping strategy.

3. The proposed framework is systematically compared with exhaustive Full-RD and SATD-based fast search using PSNR, SSIM, bitrate, BDBR, encoding time, time saving, speedup, optimal-mode retention, mode-wise hit rate, candidate count, and mode-confusion analysis across multiple QP values.

2. Related Work

Intra-mode prediction is a key aspect of the encoding process of H.264/AVC due to how much the mode selection of the prediction can impact the performance and complexity of the encoder. Although the full RDO can lead to efficient mode selection, its high cost greatly limits the practicality of H.264/AVC in real-time and resource-constrained applications, and therefore, researchers have been experimenting with various approaches such as edge and gradient prediction, statistical and feature-based mode selection, RDC optimisation, and hardware-efficient architectures in recent years. We categorise the existing studies based on their approaches and discuss their major contributions and improvements in performance and limitations.

2.1 Edge-, Gradient, and Direction-Based Intra-Mode Decision

Intra prediction in H.264/AVC utilises the proximity of neighbouring pixels to reduce the prediction residual and improve coding efficiency. However, considering all available intra-prediction modes using RDO significantly increases the complexity of the encoder. Accordingly, several authors have studied the edge, gradient, and directional information to find good prediction modes early. Hamdy et al. [7] proposed the Best Prediction Matrix Mode algorithm for H.264 intra prediction to improve the mode selection based on the spatial characteristics of the coding block. Zheng et al. [8] proposed a gradient-based interpolation for angular intra prediction and showed the effectiveness of gradient information for directional prediction, but their method was focused on HEVC. Rath et al. [9] explored directional intra prediction in different angular directions and showed the importance of local texture orientation. For H.264/AVC, the work in [10], Efficient mode decision scheme based on edge detection with Gaussian pulse for Intra-prediction in H.264/AVC, used Prewitt-gradient edge detection and block homogeneity to find the appropriate intra modes. The method delivered an average 1.14 dB BD-PSNR improvement and 11.34% BDBR reduction for QCIF sequences. Still, edge- and gradient-based methods are based on handcrafted features and empirically selected thresholds, which may make them ineffective for highly textured, weak-edge, or non-directional blocks.

2.2. Fast mode decision based on features and statistical analysis

Another direction in research is to use statistical and low-complexity features to eliminate unlikely modes before complete RDO. Zhang et al. [11] proposed statistical early-termination and early-skip models for HEVC intra coding and showed that block-level statistical features can reduce the number of unnecessary mode evaluations. Sun et al. [12] used a deep convolutional network for perceptual-based intra-coding optimisation, while Jamali and Coulombe [13] proposed RDO-cost prediction for fast intra-mode decision. Chen et al. [14] and Yang and Wei [15] also studied fast intra-mode decision for VVC and HEVC, respectively. Vaidya and Metkar [16] used gradient-based features for early termination and fast mode selection. While these studies mainly focus on HEVC/VVC and not H.264/AVC, they indicate a key development towards replacing exhaustive mode testing with feature-guided candidate selection. They show that inexpensive spatial, gradient, and statistical descriptors can be used to predict promising modes. However, the direct application of these approaches to H.264/AVC is relatively limited. This provides an opportunity for lightweight feature-based mode prediction for the H.264/AVC mode structure.

2.3. RDC-based mode optimisation

RDC modelling is another way to improve intra-mode decision without modifying the H.264/AVC prediction structure. Jamali and Coulombe [13] investigated RDO-cost prediction to reduce the computational burden of HEVC intra-mode evaluation. More recently, Anitha Kumari and Udupa [17] presented an optimal cost-modelling scheme for efficient intra-prediction mode selection in video compression. Their approach alleviated the limitations of the conventional Lagrangian cost function and proposed a new cost formulation to improve the trade-off between reconstructed quality, bitrate, and encoding time. The work from the authors shows that a better decision criterion can reduce unnecessary computational operations while maintaining coding efficiency. Even so, cost-based approaches may still need prediction and cost calculations for multiple candidate modes and thus may not eliminate the computational overhead associated with intra-mode evaluation. A more efficient approach would be to combine cost modelling with an initial feature-based mode-pruning approach so that expensive RDO is only performed for highly probable candidates.

2.4. Hardware-oriented and low-power Intra-Prediction.

Hardware-efficient implementation has also attracted much attention in real-time and resource-constrained video applications. Pathak et al. [18] proposed a 16-intra-prediction architecture for error concealment in H.264. Bhavikatti and Banakar [19] investigated a low-power intra-prediction design for the H.264 baseline encoder. The authors noted the importance of reducing hardware complexity, power consumption, and processing requirements. Tayyebi et al. [20] developed a parametrised low complexity H.264-based video encoder architecture for FPGA implementation using parallel and pipelined processing. The architecture was able to achieve a throughput of up to 2.3 Gbit/s and is suitable for high-throughput H.264 encoding. While these hardware-based architectures are able to offer much more throughput and performance in the implementation, they are not so much about the optimal intra mode prediction. Therefore, a computationally cheap and content-adaptive mode-decision algorithm for such hardware architectures is a major area of research.

2.5. Variance-, Similarity and Adaptive Mode-Decision Approach

Recently, H.264/AVC researchers have explored adaptive mode choice with a block variance and similarity measure. Sildas et al. [21] proposed variance-based reconfigurable modules for H.264 intra-prediction mode selection with different similarity measures, such as cosine similarity, SAD, SSD, Hamming distance, and Euclidean distance. They showed that a variance-based Hamming distance could reduce computational complexity by 44% while still maintaining low hardware resource usage. This work is important because it moves from fixed mode evaluation to content-dependent and reconfigurable mode selection. However, the decision process still depends on predefined similarity measures and variance thresholds. Such handcrafted decision rules may not work well for blocks with multiple directional structures, weak edges, or complex textures. Moreover, the method does not fully exploit the combined discriminative information of multiple spatial and statistical features.

Table 1 summarises the latest H.264/AVC intra-prediction studies based on their main methodological contributions. The studies are divided into edge/gradient-based, cost-model optimisation, hardware acceleration, and similarity/reconfigurable decision approaches. The representative work, the underlying principle, the major improvements reported, and the main limitation are listed for each category. This comparison illustrates the progress made in overcoming intra-mode decision complexity and the challenges to develop more adaptive and efficient mode prediction methods.

Table 1 - Comparative summary of recent H.264/AVC intra-prediction and fast intra-mode decision approaches.

Category	Representative work	Main principle	Significant reported result	Main limitation
Edge/gradient-based	[10]	Prewitt edge direction + homogeneity	1.14 dB BD-PSNR gain; 11.34% BDBR reduction for QCIF	Fixed thresholds; mainly luminance/texture dependent
Cost-model optimization	[17]	Alternative RDC function	Better balance of PSNR, bitrate and encoding time	Does not fully eliminate mode-evaluation complexity
Hardware acceleration	[20]	Parallel/pipelined FPGA architecture	Up to 2.3 Gbit/s throughput	Focuses on implementation rather than adaptive coding efficiency
Similarity/reconfigurable decision	[21]	Variance-based similarity measures	Up to 44% computational-complexity reduction	Relies on predefined similarity measures and thresholds

Reviewed studies demonstrate that the transition from full RDO to edge-based, statistical, cost-driven, adaptive and hardware-efficient intra-mode decision is gradual. H.264/AVC approaches, however, still struggle with the very basic trade-off between coding efficiency and computational complexity. Most approaches depend on handcrafted features, fixed thresholds or predefined similarity measures, which may not generalise to different video contents, resolutions and quantisation parameters. Edge- and gradient-based approaches are good for strongly directional blocks but less reliable for complex or non-directional textures, while cost modelling and hardware solutions can still have a significant computational overhead or rely on algorithms and predetermined decision-making. Moreover, compared to HEVC and VVC, little work has been done on lightweight, adaptive or data-driven intra-mode prediction for H.264/AVC. These limitations motivate the development of a content-adaptive fast intra-mode prediction framework that exploits cheap block-level features to identify the most useful prediction modes before full RDO, thereby eliminating unnecessary mode evaluations while maintaining competitive RD performance and finding a good balance between encoding time, computational complexity, bitrate and reconstructed video quality.

3. Proposed Methodology

The work proposes a framework shown in Figure 1 for reducing the burden of performing full intra-mode RDO in H.264/AVC by predicting a small number of promising modes from the spatial direction of each 16×16 luminance block. Instead of performing transform, quantisation, reconstruction, rate estimation and RDC analysis of all nine intra-prediction modes, the block in the wavelet domain is estimated to determine its dominant orientation. This information is mapped to a smaller number of directional candidates that can be evaluated using the same RD procedure. The overall processing flow consists of block extraction, wavelet decomposition, directional feature extraction, orientation interpretation, mode mapping, candidate selection, intra prediction, transform/quantization, RDO and reconstruction, which are completed in a short time.

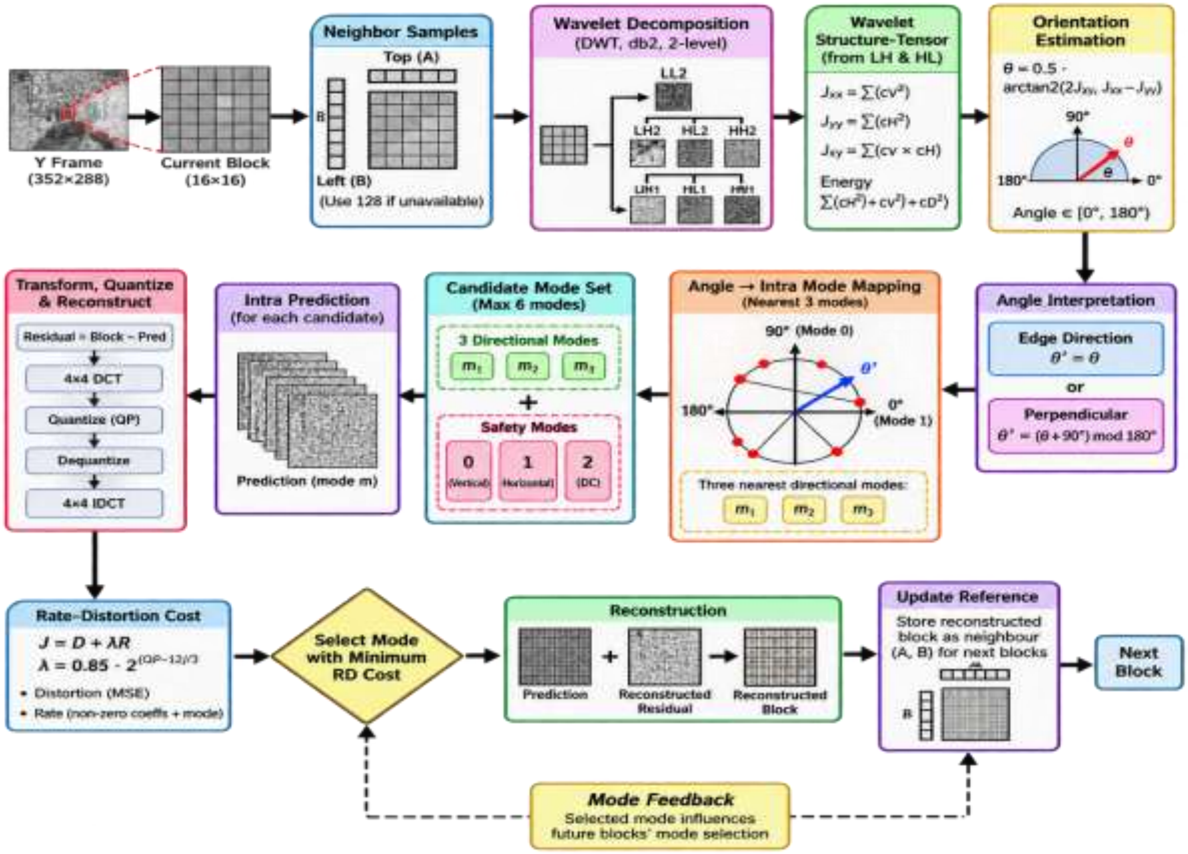


Figure 1 - Proposed wavelet-guided content-adaptive intra-mode prediction framework for H.264/AVC, showing DWT-based directional feature extraction, structure-tensor orientation estimation, candidate-mode selection, reduced RD evaluation, and reconstructed-block feedback.

3.1 Input Frame and Block Extraction

The input video is initially decoded as the luminance component of a YUV 4:2:0 video. For each frame, the luminance image is partitioned into non-overlapping 16×16 blocks. The block size is selected to correspond to the principal macroblock-level structure commonly associated with H.264/AVC intra coding.

$$Input\ Y\ frame\ \rightarrow 16\ x\ 16\ block$$

Processing individual blocks enables the proposed method to determine the local texture and directional characteristics rather than applying a single prediction strategy to the entire frame. This content-adaptive processing is important because smooth, edge-dominated, and highly textured regions generally exhibit different optimal intra-prediction modes.

3.2 Neighbour-Sample Acquisition

For each current block, the reconstructed top and left boundary samples are obtained from previously reconstructed blocks. These samples are required to generate H.264/AVC-like intra predictions.

$$Top\ samples + Left\ samples \ \rightarrow Intra\ Predictions$$

When a boundary is unavailable, the implementation initialises the corresponding samples to 128. This provides a valid reference for blocks located at the frame boundaries and avoids additional boundary-processing complexity. The reconstructed-neighbour mechanism also preserves the sequential dependency of intra coding: once the optimal mode and reconstructed block are obtained, that block develops into a reference for subsequent blocks.

3.3 Two-Level Wavelet Decomposition

The central feature-extraction stage applies a two-level discrete wavelet transform (DWT) using the Daubechies-2 (db2) wavelet:

$$B \xrightarrow{DWT_{db2}} \{LL_2, LH_2, HL_2, HH_2, LH_1, HL_1, HH_1\}$$

The DWT separates the block into low-frequency approximation and high-frequency directional components. In particular, the horizontal and vertical detail sub-bands provide information about local directional structures, while the diagonal component contributes information about texture complexity. Using two decomposition levels allows the method to capture directional characteristics of more than one spatial scale. The main advantage of this stage is that directional information can be extracted before performing expensive RD calculations. Thus, the wavelet transforms act as a low-complexity screening mechanism for the nine possible intra modes.

3.4 Wavelet-Domain Structure-Tensor Analysis

The horizontal and vertical wavelet detail coefficients are then used to construct directional statistics. The implementation accumulates the following structure-tensor components given by expressions from (1) to (4):

$$J_{xx} = \sum (cV)^2 \qquad (1)$$

$$J_{yy} = \sum (cH)^2 \qquad (2)$$

$$J_{xy} = \sum (cV * cH) \qquad (3)$$

The dominant orientation is subsequently estimated as:

$$\theta = \frac{1}{2}\, tan^{-1}\left(\frac{2*J_{xy}}{J_{xx}-J_{yy}}\right) \qquad (4)$$

The resulting angle is converted to the interval $\theta \ \in [0^0, 180^0]$. Conceptually, this stage converts the wavelet coefficients into a single directional descriptor representing the dominant orientation of the current block.

$$Wavelet\ Coefficients \ \rightarrow \ J_{xx}, J_{yy}, J_{xy} \ \rightarrow Dominant\ Angle\ \theta$$

This is the key mechanism that is responsible for making the proposed method content adaptive. Instead of testing every prediction mode equally, the algorithm uses the measured block orientation to determine which directional modes are most likely to produce a low prediction error.

3.5 Edge and Perpendicular Orientation Interpretation

Two alternative interpretations of the estimated wavelet orientation are investigated. In the first strategy, the estimated orientation itself is used, which is given by expression (5):

$$\theta' = \theta \qquad (5)$$

In the second strategy, the perpendicular direction is considered using expression (6):

$$\theta' = (\theta + 90^0)\ mod\ 180^0 \qquad (6)$$

Thus, the proposed framework evaluates both direct edge-orientation and perpendicular orientation.

$$\theta = \begin{cases} \theta, & Edge\ strategy \\ (\theta + 90^0)\ mod\ 180^0, & Perpendicular\ strategy \end{cases} \qquad (7)$$

This comparison is important because the orientation of an image edge and the direction along which prediction samples should be propagated are not necessarily identical. Evaluating both interpretations allows the experimental framework to determine which relationship provides better correspondence with the Full-RD optimal mode.

3.6 Orientation-to-Intra-Mode Mapping

The estimated angle is compared with the predefined directional angles given in Table 1: $\{0^0, 22.5^0, 45^0, 67.5^0, 90^0, 112.5^0, 135^0, 157.5^0\}$. Circular angular distance is used so that orientations near and are treated consistently:

$$d(\theta, \theta_m) = min(|\theta - \theta_m|, 180^0 - |\theta - \theta_m|) \qquad (8)$$

The three directional modes having the smallest angular distances are selected.

$$\theta \ \rightarrow Nearest\ 3\ directional\ modes$$

This mapping is computationally inexpensive because it requires only angle-distance comparisons rather than performing transform and RD calculations for all modes.

Table 1 - Orientation-to-intra-mode mapping used for directional candidate selection.

Wavelet Orientation Angle	Directional Interpretation	H.264/AVC Intra Mode	Mode Description
0°	Vertical direction	Mode 1	Vertical
22.5°	Near-diagonal direction	Mode 6	Horizontal Down
45°	Diagonal direction	Mode 4	Diagonal Down Right
67.5°	Near-diagonal direction	Mode 7	Vertical Left
90°	Horizontal direction	Mode 0	Vertical*
112.5°	Near-diagonal direction	Mode 5	Vertical Right
135°	Diagonal direction	Mode 3	Diagonal Down Left
157.5°	Near-diagonal direction	Mode 8	Horizontal Up

3.7 Safety-Mode Inclusion and Reduced Candidate Set

Although directional information is highly useful for edge-dominated blocks, relying exclusively on directional modes could cause the optimal mode to be missed for smooth or non-directional regions. To improve robustness, three additional safety modes described in Table 2 are included: Mode 0: Vertical, Mode 1: Horizontal, and Mode 2: DC. Consequently, the candidate set contains the three wavelet-selected directional modes together with these three safety modes given by expression (9):

$$C = \{m_1, m_2, m_3\} \cup \{0, 1, 2\} \quad (9)$$

The maximum number of candidates is therefore six rather than nine.

$$3\ predicted\ modes + 3\ safety\ modes\ \leq 6\ candidates$$

This is an important design feature because it provides a compromise between aggressive mode pruning and optimal-mode retention. The directional modes exploit texture orientation, while the safety modes provide additional protection against incorrect directional prediction.

Table 2 - Safety modes retained to improve robustness of the proposed intra-mode prediction framework.

Safety Mode	Mode ID	Purpose
Vertical	**0**	Provides a basic spatial prediction candidate
Horizontal	**1**	Provides an alternative directional candidate
DC	**2**	Provides a non-directional prediction for smooth/weakly textured blocks

3.8 Intra-Prediction Generation

For every candidate mode, an intra-predicted block is generated using the reconstructed top and left samples. The implementation supports all nine H.264/AVC intra-mode identifiers used in the experimental framework. For a candidate mode,

$$P_m = f(T, L, m) \quad (10)$$

where T and L denote the top and left reference samples. The prediction residual is then obtained using expression (11):

$$R_m = B - P_m \quad (11)$$

This stage is performed only for the reduced candidate set, rather than all nine modes. Therefore, the number of subsequent transform, quantisation, reconstruction, and RD operations is directly reduced.

3.9 Transform, Quantisation and Reconstruction

For each candidate mode, the residual undergoes block-wise DCT processing given by (12):

$$R_m \xrightarrow{DCT} C_m \xrightarrow{Q(QP)} \widehat{C_m} \xrightarrow{Q^{-1}} \tilde{C}_m \xrightarrow{IDCT} \tilde{R}_m \quad (12)$$

The reconstructed candidate block is then expressed as:

$$\widehat{B_m} = clip(P_m + \tilde{R}_m, 0, 255) \quad (13)$$

The QP-dependent quantisation step is calculated by expression (14):

$$Q_{step} = 2^{(QP-4)/6} \quad (14)$$

This stage allows the candidate modes to be compared under the same quantisation conditions. Since the proposed method reduces the number of candidates entering this computationally expensive stage, the overall encoding complexity is reduced.

3.10 RDC Calculation

Each candidate is finally evaluated using the implemented RDC given by expression (15):

$$J_m = D_m + \lambda * R_m \quad (15)$$

Where D_m represents the mean-squared reconstruction error, R_m is the estimated rate, and

$$\lambda = 0.85 * 2^{(QP-12)/3} \quad (16)$$

The rate estimate is based on the number of non-zero (N_{NZ}) quantised coefficients together with a mode signalling contribution:

$$R_m = 1.5 * N_{NZ} + 3 \tag{17}$$

The candidate with minimum RDC is selected using equation (18):

$$m^* = arg \min_{m \in C} J_m \tag{18}$$

Thus, the wavelet stage does not directly determine the final mode. Instead, it narrows the search space, after which RDO determines the final coding decision. This separation is important because it allows the proposed framework to retain the coding-efficiency advantage of RDO while avoiding unnecessary evaluation of unlikely modes.

3.11 Optimal Mode Selection and Block Reconstruction

The candidate producing the minimum RDC is selected as the final intra mode. Its reconstructed block is inserted into the reconstructed frame as:

$$arg \min_{m \in C} J_m \rightarrow Selected\ Mode \rightarrow Reconstructed\ block$$

The reconstructed block is then stored as a reference for subsequent blocks. Consequently, the framework maintains the spatial dependency required for sequential intra prediction.

3.12 Reference Update and Processing of the Next Block

After reconstruction, the selected block becomes part of the reference image. The next block can therefore use its reconstructed samples as top or left neighbours.

$$Selected\ Mode \rightarrow Reconstruction \rightarrow Reference\ Update \rightarrow Next\ block$$

This feedback mechanism ensures that the proposed method does not operate independently on each block; rather, it follows the reconstructed-frame dependency of the encoding process.

3.13 Overall Computational Advantage

The principal computational advantage of the proposed framework comes from moving directional analysis before full RDO. In exhaustive Full-RD search, all nine modes undergo prediction, followed by other stages, and RD-cost computation. In the proposed approach, only the wavelet-selected candidates and safety modes proceed to the expensive RD stage. The theoretical candidate reduction is therefore

$1 - \frac{6}{9} = 33.33\%$.

In the worst case, the actual computational saving depends on the measured candidate distribution and implementation overhead of the wavelet analysis. Importantly, the framework does not simply remove modes arbitrarily; it uses block-specific directional information to determine which modes are most promising and retains vertical, horizontal, and DC modes as safety candidates.

The proposed framework evaluates three intra-mode search strategies under a common H.264/AVC-like coding environment: Full RD Search, SATD-based Fast Search, and the proposed Wavelet-Guided Fast Search (Edge and Perpendicular). Full RD Search provides the reference for assessing coding efficiency, while the two fast strategies reduce the number of computationally expensive RD evaluations. The overall algorithm is summarised in Algorithm 1.

Algorithm 1: Intra-Mode Prediction Framework

Input	Video sequence V, Block size B = 16 × 16, QP set Q = {20, 24, 28, 32}, Number of intra modes M = 9
Output	PSNR, SSIM, bitrate, encoding time, time saving, speedup, BDBR, optimal-mode retention and candidate statistics

Read the luminance component of the input YUV video sequence.
For each quantization parameter QP ∈ Q do
    Divide every frame into non-overlapping 16 × 16 blocks.
    For each current block X do
        Obtain the reconstructed top and left neighbouring samples from previously processed blocks.
        Generate the nine intra-predicted blocks corresponding to all available intra prediction modes.

        Full RD search

1. Evaluate all nine modes using transform, quantisation, reconstruction, rate estimation, and RDC.
2. Select the minimum-RD-cost mode.

SATD-based fast search

1. Rank the nine modes using SATD.
2. Select the three lowest-SATD candidates.
3. Perform RDO on the selected candidates.

WAVELET-guided fast search (Edge and Perpendicular)

1. Perform two-level db2 wavelet decomposition.
2. Estimate block orientation using wavelet-domain structure-tensor analysis.
3. Map the orientation to the three nearest directional modes.
4. Add vertical, horizontal, and DC safety modes.
5. Perform RDO on the resulting candidate set.

Performance and Mode analysis

Calculate PSNR, SSIM, bitrate, and encoding time.

Calculate time saving, speedup, and BDBR relative to Full RD Search.

Evaluate wavelet optimal-mode retention, mode-wise hit rate, average candidate count, candidate frequency, and confusion matrix.

Report coding-efficiency and computational-complexity results.

4. Results and Discussion

We evaluate the performance of the proposed content-adaptive wavelet-guided intra-mode prediction method against Full RD Search and SATD-based Fast Search using four CIF video sequences: Coastguard, Container, Foreman, and News. The experiments were performed at QP values of 20, 24, 28, and 32 with 16 × 16 blocks. Full RD Search is the reference method because it evaluates all nine H.264/AVC intra modes using the complete RD process. The SATD-based method is based on a typical fast search, whereas the proposed method uses two-level db2 wavelet decomposition and wavelet-domain orientation estimation to select a smaller number of directional candidates along with vertical, horizontal, and DC safety modes. We evaluate reconstructed quality, bitrate, encoding time, time saving, speedup, BDBR, and optimal mode retention.

4.1 Overall Coding Performance

Table 3 summarises the major average results reported for the evaluated sequences. The proposed Wavelet Edge method always provides a better computational efficiency/quality trade-off than the SATD-based Fast method. For the reported results, the Wavelet Edge method has an average PSNR of 39.491 dB and 39.443 dB for Full RD Search and 39.403 dB for SATD Fast Search, respectively, and the average SSIM for both cases increases from 0.95214 for Full RD Search to 0.95273 for the proposed method. Thus, the reduction in the number of RD evaluations does not result in a degradation of reconstructed quality, but a small improvement when followed by the proposed coding model.

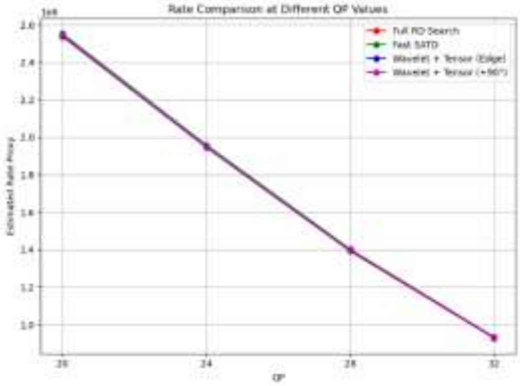
Table 3 - Overall comparison of Full RD Search, SATD-based Fast Search, and the proposed Wavelet-guided Edge Search in terms of PSNR, SSIM, bitrate, and encoding time.

Method	Avg. PSNR (dB)	Avg. SSIM	Avg. Rate	Avg. Encoding Time (s)
Full RD Search	39.443	0.95214	1,070,354.63	72.07
SATD Fast Search	39.403	0.95177	1,087,574.63	60.09
Proposed Wavelet Edge	**39.491**	**0.95273**	1,082,682.75	**25.87**

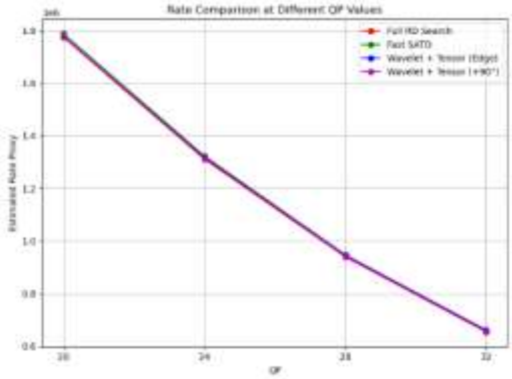
The main benefit of our proposed method is not just mode pruning but also being able to tune the mode content. SATD Fast Search has only three modes with the lowest SATD, but this strategy would require first evaluation of all nine modes. The proposed method estimates the most dominant block orientation in the wavelet domain and directly builds a small candidate set. This is the reason for the much lower encoding time of the proposed method. The average encoding time for Full RD Search is reduced from 72.07 s to 25.87 s. This is an average time savings of 63.73% and a speedup of 2.79×. By comparison, SATD Fast Search offers only 16.12% time saving and about 1.20× speedup.

4.2 RD Performance

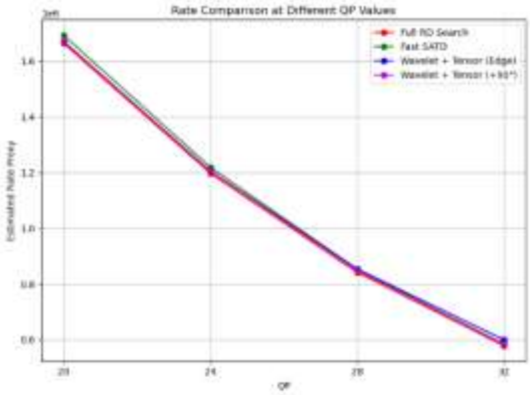
The RD results in Figure 2 (hereafter, the four figures will indicate results on coastguard, container, foreman, and news test sequences) also show that the proposed method is still competitive in coding even though it reduces the computational complexity very significantly. Full RD Search has the lowest average rate in our experiment, but SATD Fast and Wavelet Edge have slightly higher rates. We find that the difference is rather small compared to the large reduction in encoding time. The BDBR of SATD Fast is 2.11% compared with Full RD Search, while this difference is only 0.78% for Wavelet Edge. Therefore, our proposed method offers a much better RD compromise than the usual SATD pruning method.



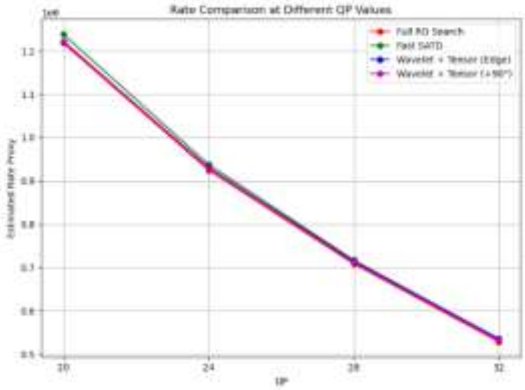
(a)



(b)



(c)



(d)

Figure 2 - Estimated rate proxy versus QP for the Full RD Search, SATD-based Fast Search, and proposed Wavelet-guided Edge Search for the (a) Coastguard, (b) Container, (c) Foreman and News CIF sequence.

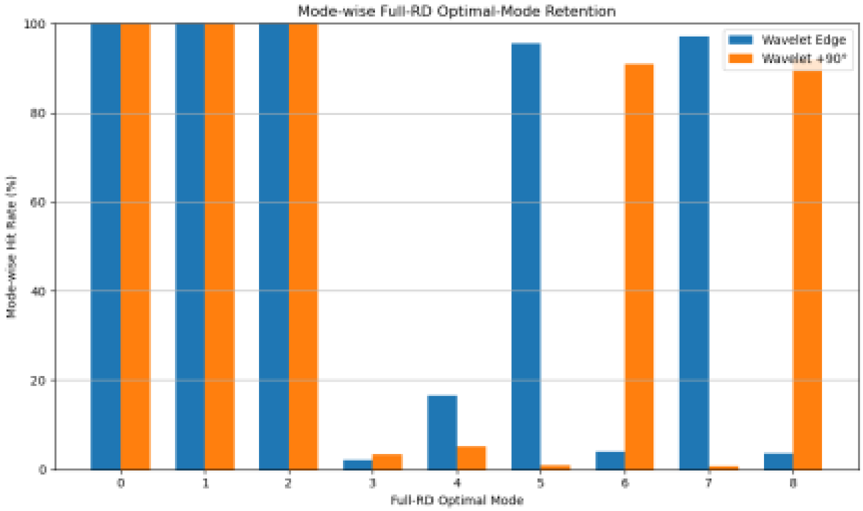
The difference is most evident at higher QP values, where we still have competitive PSNR while we save significant computational energy. For example, at QP = 32, the PSNR is 35.082 dB for Full RD Search, 35.049 dB for SATD Fast Search and 35.189 dB for Wavelet Edge. At the same QP, the encoding time is reduced from 51.03 s for Full RD Search to 43.49 s for SATD Fast and only 19.34 s for Wavelet Edge. In short, our proposed method achieves approximately 62.10% time saving and 2.64× speedup, but slightly improves PSNR compared to the reference search.

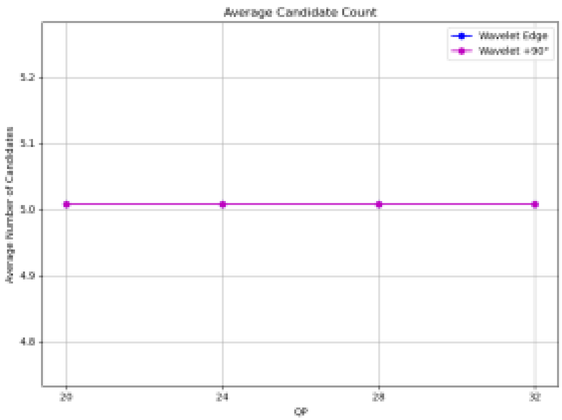
4.3 Sequence-Level Behaviour

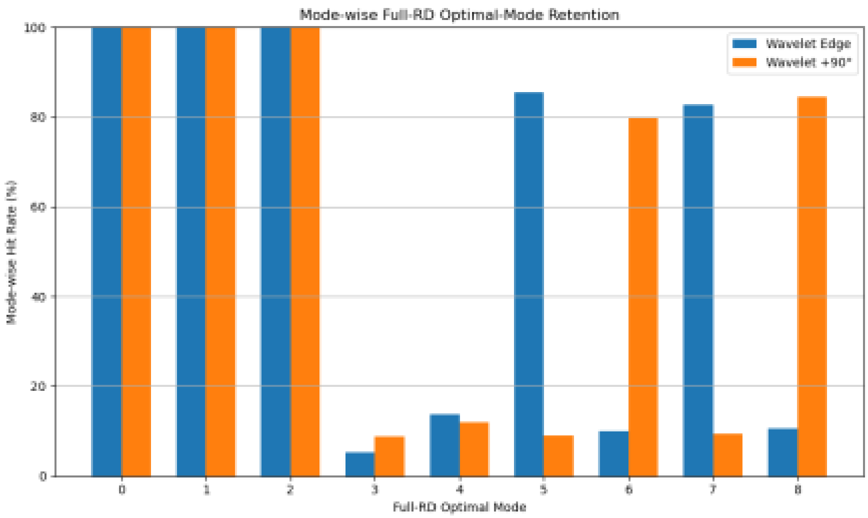
The sequence-level analysis in Figure 3 shows that the effectiveness of our method depends on the spatial properties of the video content. The optimal mode retention analysis shows that this is the case. For Coastguard, the Wavelet Edge candidate sets have a hit rate of 91.14%, 88.22%, 87.48%, and 87.45% for QP values of 20, 24, 28 and 32, respectively. The average number of candidate sets is 5.04 modes.

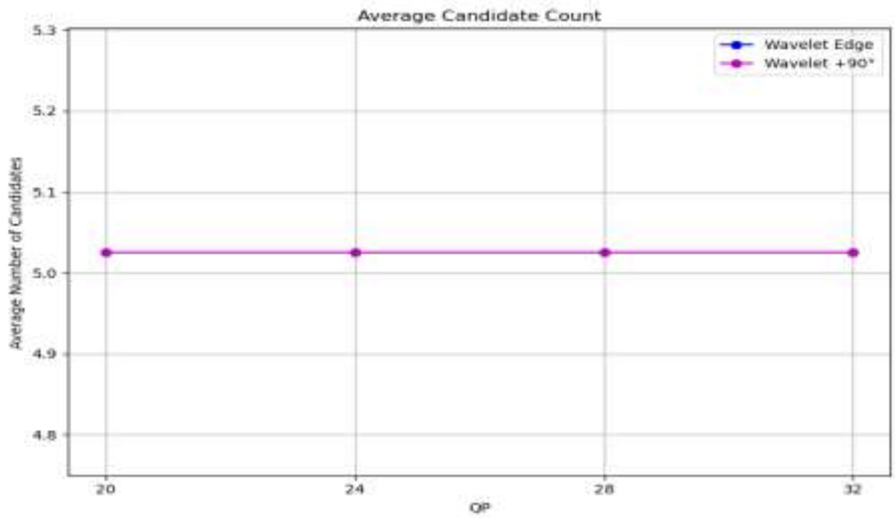
For Container, the reported Wavelet Edge hit rates are lower, ranging from 86.75% at QP 20 to 75.27% at QP 32. This means that blocks with more complex or less clearly defined directional structures are more difficult to characterise with one dominant wavelet orientation. However, the potential selection set is still limited to around 5.06 modes, and still the method provides substantial pruning.

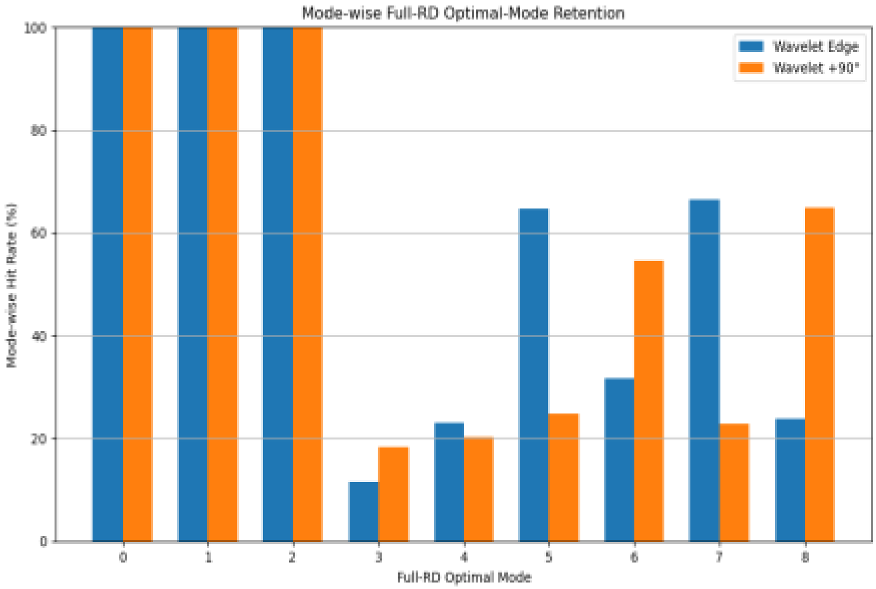
The Foreman results are similar, with the reported Wavelet Edge retention decreasing from 86.75% at QP 20 to 75.27% at QP 32. On the other hand, the News sequence shows a very good mode retention, with the reported Wavelet Edge hit rate reaching 91.14% at QP 20. These results suggest that the proposed method is most effective if we have strong spatial structures in the block and can represent the spatial structure by dominant directional information.

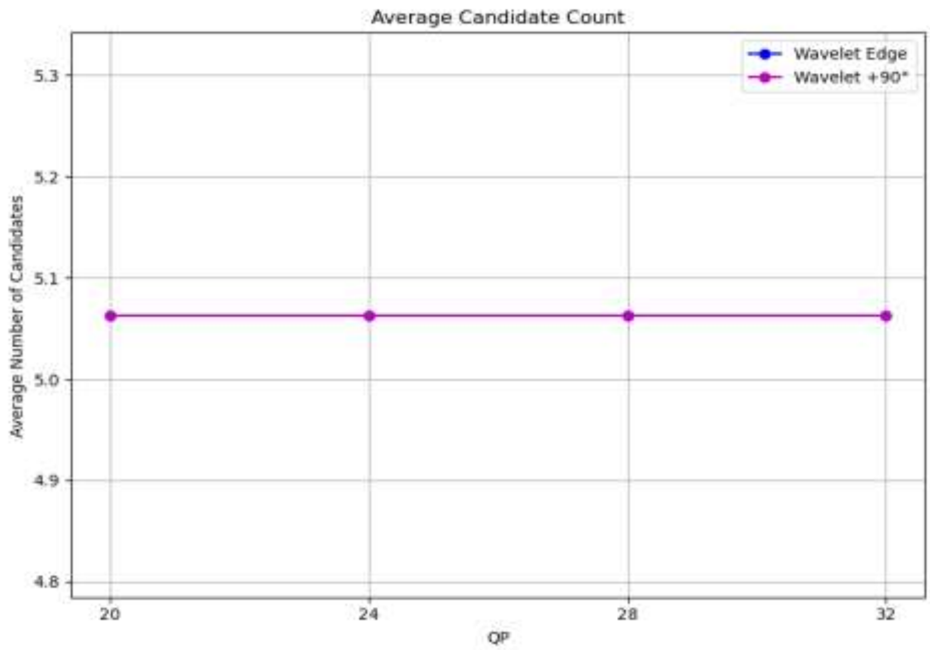







Average Candidate Count
Wavelet Edge
Wavelet +90°
Average Number of Candidates
5.3
5.2
5.1
5.0
4.9
4.8
20
24
28
32
QP


Mode-wise Full-RD Optimal-Mode Retention
Wavelet Edge
Wavelet +90°
Mode-wise Hit Rate (%)
100
80
60
40
20
0
0
1
2
3
4
5
6
7
8
Full-RD Optimal Mode


Average Candidate Count
Wavelet Edge
Wavelet +90°
Average Number of Candidates
5.3
5.2
5.1
5.0
4.9
4.8
20
24
28
32
QP

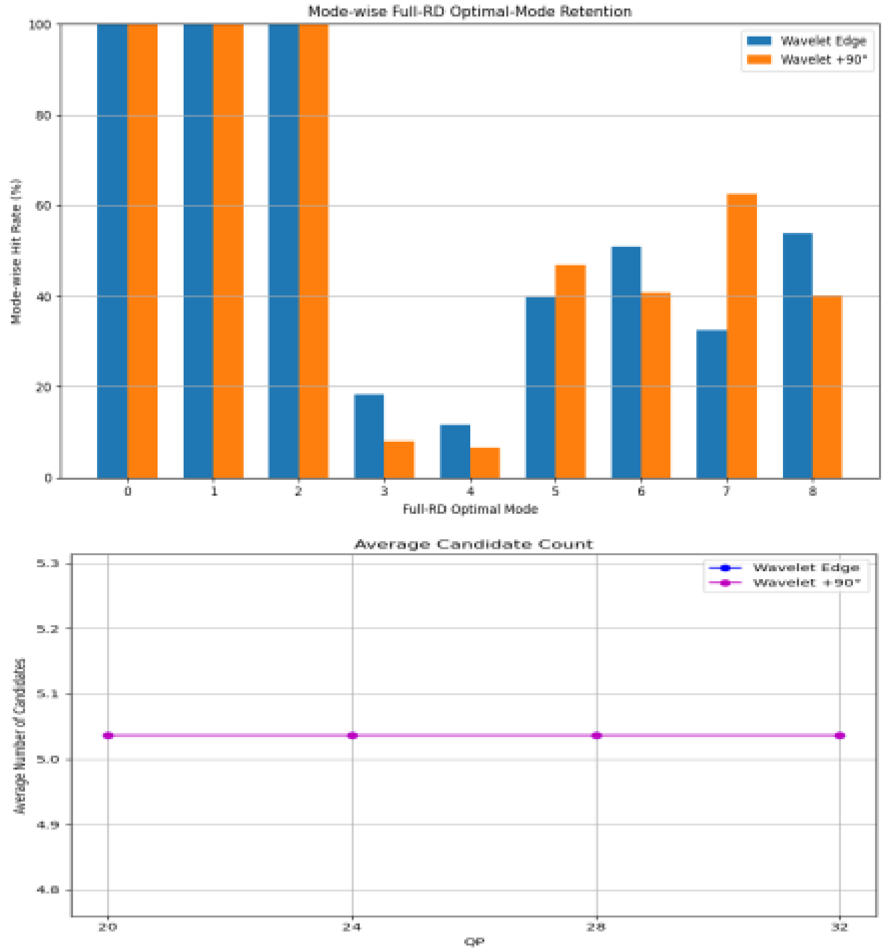
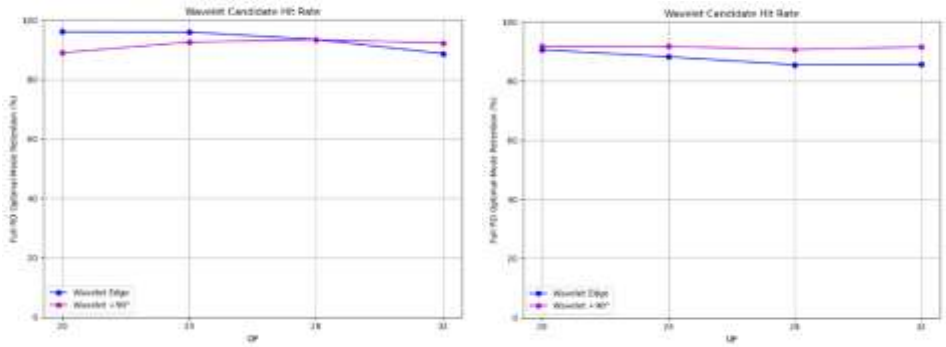


Figure 3 - Mode-wise Full-RD optimal-mode retention and average number of candidate intra modes selected of the proposed Wavelet-guided Edge Search across the Coastguard, Container, Foreman, and News CIF sequences.

4.4 Effectiveness of Wavelet-Based Orientation Selection

One of the key components of the proposed framework is the use of wavelet-domain orientation information to generate candidate modes. We extract the horizontal, vertical and diagonal wavelet information at different scales and integrate this information into a structure tensor. The resulting orientation is mapped to the three nearest directional H.264/AVC intra modes. Vertical, horizontal and DC modes are then kept as safety candidates. This means that the maximum number of candidate modes is 6 and the average number of candidates is around 5.0.

The high optimal mode retention in Figure 4 indicates that this compact candidate set captures a substantial proportion of the modes selected by Full RD Search. For example, the average Wavelet Edge hit rates are 93.52% for one of the sequence sets and 88.57% for another, while the candidate count is still close to five modes. This confirms that wavelet orientation is useful low-complexity information for predicting the suitability of intra prediction modes before complete RDO.



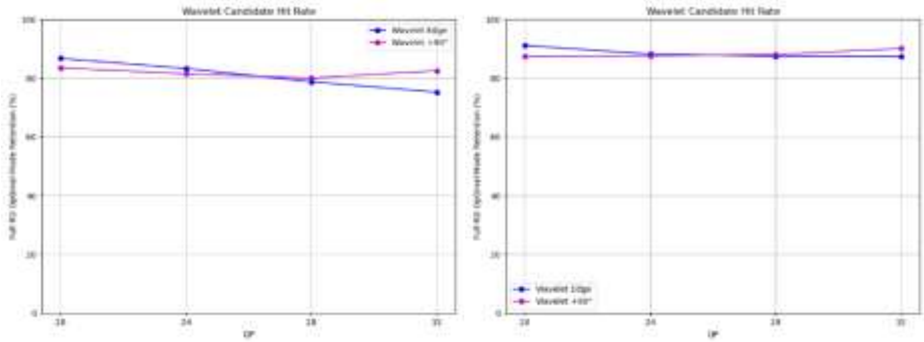


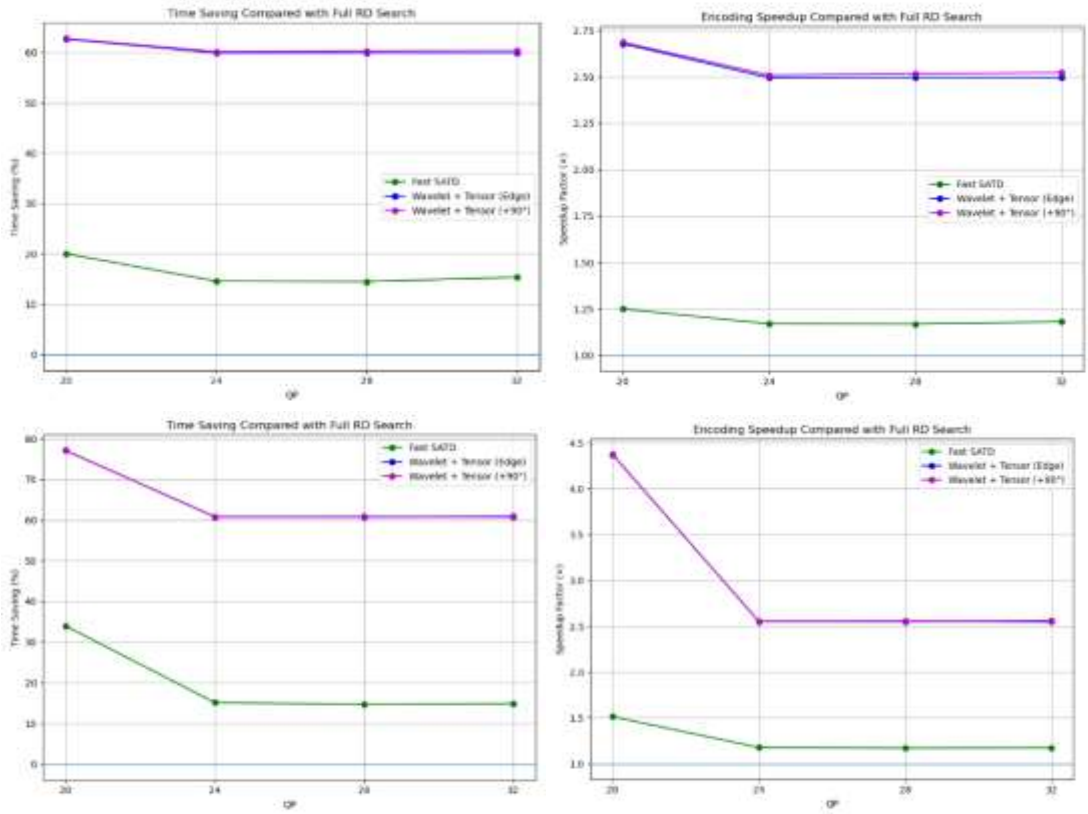
Figure 4 - Full-RD optimal-mode retention versus QP for the proposed Wavelet-guided Edge Search across the Coastguard, Container, Foreman, and News CIF sequences.

The mode-wise results also reveal an important feature of the proposed strategy. Vertical, horizontal and DC modes are 100% retained in all the experiments reported here, because these three modes are chosen to be safety candidates. Directional modes have a lower individual retention rate, especially for a mode with complex diagonal structures. This is not surprising since a single orientation estimate cannot perfectly represent blocks with multiple edges or weak directional structure. The safety mode mechanism therefore plays an important role in avoiding excessive mode pruning.

4.5 Comparison with SATD-Based Fast Search

SATD Fast Search is a good baseline as it reduces the number of complete RD evaluations from nine to three. However, the candidate generation process requires SATD evaluation for all nine modes before choosing the best three of them. Therefore, the computational reduction is not that great. The average time saving is only 16.12%, with a speedup of 1.19× and BDBR of 2.11%.

The proposed Wavelet Edge method follows a different approach. Instead of evaluating all modes to determine which candidates are promising, it uses inexpensive block-level structural information to predict promising directional modes before RD evaluation. This significantly reduces the encoding time. From Figure 5, the average time saving of 63.73% and speedup of 2.76× are nearly four times greater than the time saving of SATD Fast Search, and the BDBR penalty is reduced from 2.11% to 0.78%. Thus, the proposed method provides a more suitable balance between computational complexity and RD performance.



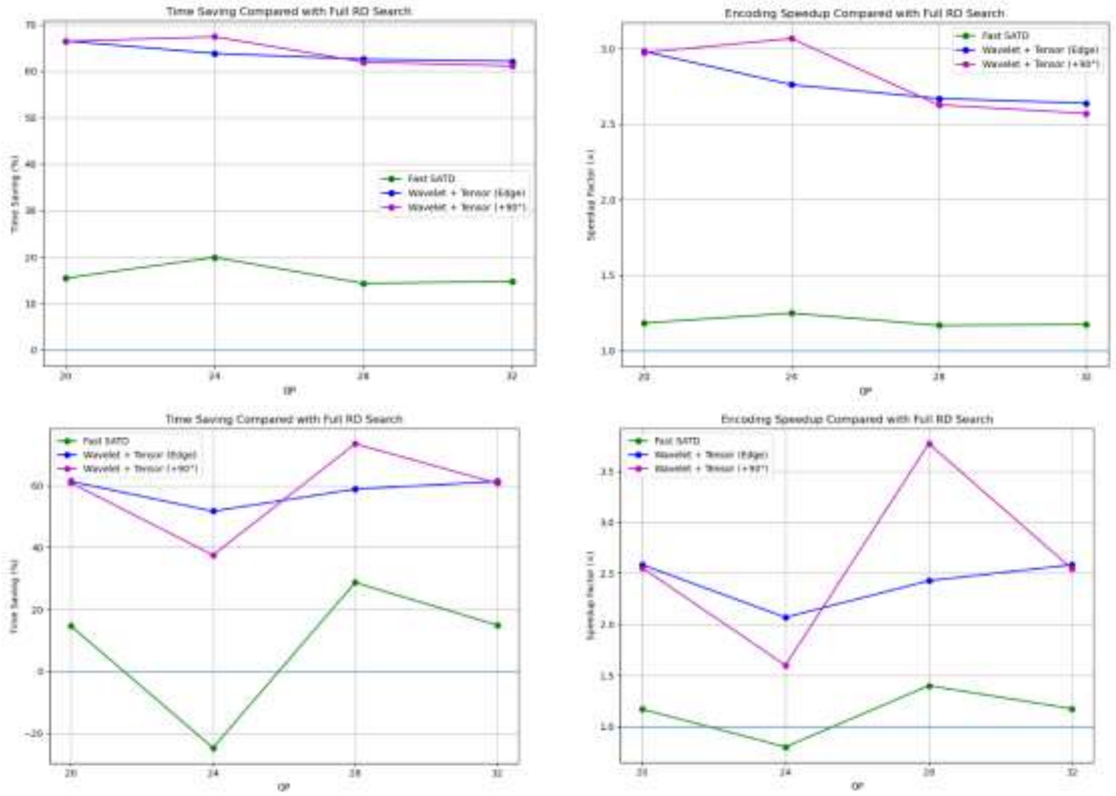
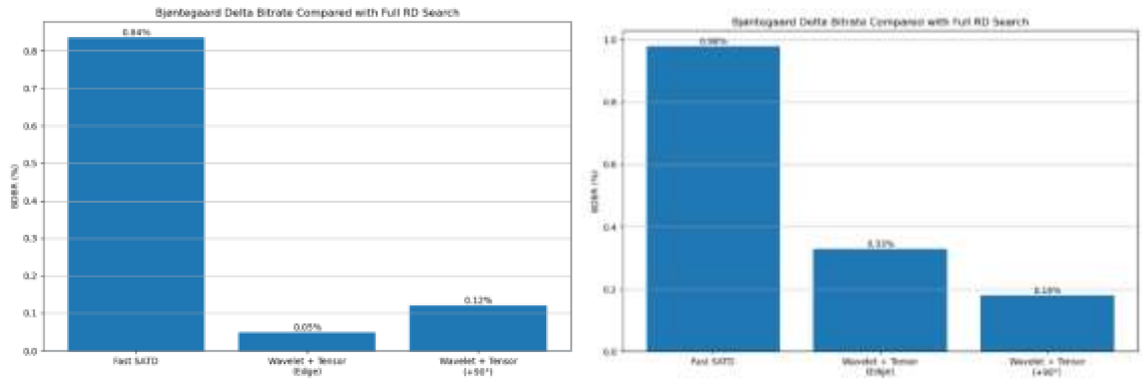


Figure 5 - QP-wise time saving and speedup factor achieved by the proposed Wavelet-guided Edge Search compared with Full RD Search for the Coastguard, Container, Foreman, and News CIF sequences.

4.6 Effect of the Perpendicular Orientation Strategy

The second wavelet configuration uses perpendicular orientation (θ + 90°) to see whether the prediction direction is better represented by the edge direction itself or by the direction normal to the detected edge. The results in Figure 6 show that neither interpretation is always better. The Wavelet Edge configuration achieves 81.03% average mode retention while the perpendicular configuration achieves 81.91%. The second also gives a slightly higher speedup of 2.81× as compared to 2.76× for the edge interpretation. However, the edge configuration has a lower BDBR of 0.78% compared to 0.53% for the perpendicular configuration; both representations can be useful depending on the content.

Therefore, the comparison indicates that orientation information should be the main decision variable and that the relationship between detected edge orientation and the optimal prediction direction is not always direct. This is important in future development of the proposed framework because adaptive selection between (θ) and (θ + 90°) may improve mode retention even if the interpretation is fixed.



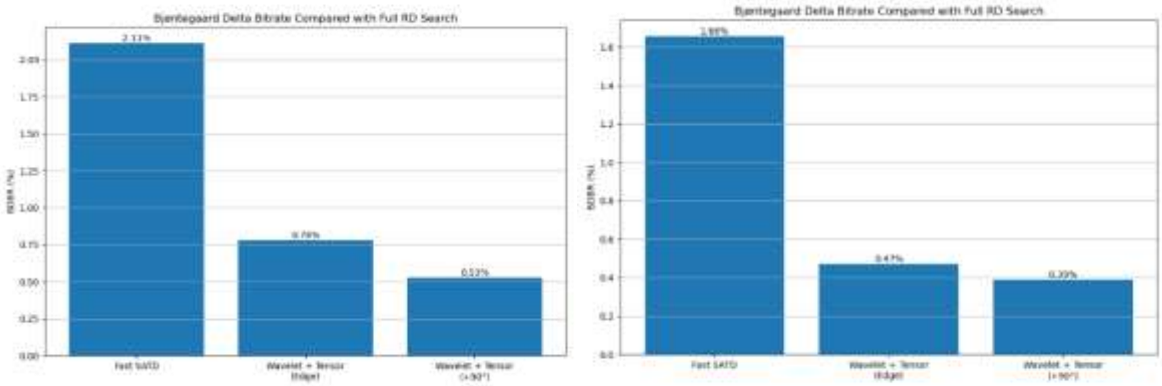


Figure 6 - BDBR comparison of the SATD-based Fast Search and proposed Wavelet-guided Edge Search relative to Full RD Search for the Coastguard, Container, Foreman, and News CIF sequences.

4.7 Overall Discussion

Overall, our experimental results in Table 4 illustrate that our Wavelet-guided intra-mode prediction framework successfully addresses the central trade-off between coding efficiency and computational complexity. Full RD Search is the best reference to optimise all modes but requires considerably more time. SATD Fast Search offers some relief to some extent, but the requirement to evaluate all the modes in the data before ranking them limits the speedup. The proposed method makes use of multiscale wavelet-domain structure to find the most suitable directional modes before the RD evaluation and thus reduces the need to evaluate all the modes.

The most important result is the simultaneous achievement of competitive or slightly better reconstruction quality and significantly lower encoding complexity. Across the reported average results, Wavelet Edge improves PSNR from 39.443 dB to 39.491 dB and SSIM from 0.95214 to 0.95273 with reduced encoding time from 72.07 s to 25.87 s. The resulting 63.73% time saving and 2.79× speedup are achieved with a relatively small BDBR increase of 0.78%.

As a result of this, our proposed approach makes use of low-complexity, content-adaptive orientation information in candidate mode prediction instead of only ranking the cost for the post-evaluation assessment. The almost five-mode average candidate set is a good compromise between aggressive pruning and preserving the Full-RD optimal mode. The mode-retention analysis also indicates that non-directional and complex directional blocks are still the main source of prediction misses. Hence, future improvements might include adaptive confidence estimation, block-texture classification or dynamic expansion of the candidate set for blocks with ambiguous wavelet orientation.

Table 4 - Summary of the principal advantages of the proposed Wavelet-guided method

Metric	Full RD Search	SATD Fast Search	Proposed Wavelet Edge	Interpretation
PSNR (dB)	39.443	39.403	**39.491**	Proposed method maintains competitive reconstruction quality
SSIM	0.95214	0.95177	**0.95273**	Slight improvement in structural quality
Encoding time (s)	72.07	60.09	**25.87**	Significant reduction in computational time
Time saving (%)	—	16.12	**63.73**	Proposed method provides substantially higher time saving
Speedup (×)	1.00	1.19	**2.79**	Proposed method achieves greater encoding acceleration
BDBR (%)	0.00	2.11	**0.78**	Lower bitrate penalty than SATD Fast Search
Average candidates	9	3	**≈5.06**	Adaptive candidate reduction

Optimal-mode retention (%)	100.00	—	**81.03**	Most Full-RD optimal modes are retained

The results therefore establish that the proposed framework offers a strong compromise between RD efficiency, reconstructed quality, and computational complexity, making it a promising lightweight intra-mode prediction strategy for H.264/AVC encoding.

5. Conclusion

In this paper, we proposed WG-CAMP to reduce the computational complexity of H.264/AVC intra coding. The proposed method uses multi-dimensional db2 wavelet decomposition and wavelet domain structure-tensor analysis to estimate the dominant spatial orientation of a 16 × 16 block. This orientation is then mapped to the three nearest directional intra modes, vertical, horizontal and DC, and these three intra modes are then used as safety candidates. With the content-adaptive candidate selection, RDO can be performed on a much smaller number of modes than in the traditional Full RD Search.

The comparison of the Coastguard, Container, Foreman and News CIF sequences over QP values of 20, 24, 28 and 32 shows that WG-CAMP is more efficient. The average encoding time for WG-CAMP is reduced by 63.73% to 25.87 s, i.e. speedup is achieved by 2.79 times. Even though the cost is much lower, the reconstruction quality of the method is still very good. The average PSNR is increased from 39.443 dB to 39.491 dB, and the SSIM is improved from 0.95214 to 0.95273. The BDBR is reduced to 0.78%, and the Full-RD best mode is still used in about 81.03% of blocks with only about 5.04 candidate modes on average.

Despite the promising results, there are still a few limitations. Firstly, the experiments are based on four CIF test sequences. At this stage, we are not able to generalise WG-CAMP to different spatial resolutions, video contents, and coding conditions. Secondly, the orientation-to-mode mapping uses specific angular relationships and a set of candidate selection strategies. Blocks containing weak, multiple, or ambiguous directional structure may not be well represented with an overall orientation. The mode-retention studies confirmed that complex and non-directional blocks are the main source of prediction misses. Second, the use of fixed safety modes guarantees robustness but limits the level of complexity reduction that is possible. Finally, the current implementation uses a fixed wavelet configuration and does not dynamically determine whether the detected edge orientation or its perpendicular direction is more suitable for one block or not. Between the two configurations, we could say that neither is always optimal for all video content.

These limitations provide many directions for future study. An adaptive confidence mechanism can be developed to estimate the reliability of the wavelet orientation and dynamically increase the candidate set for blocks with uncertain structural information. Block-level texture classification using wavelet energy, variance, gradient strength or directional coherence could further distinguish strongly directional blocks from complex and non-directional regions. The choice between the detected orientation θ and its perpendicular direction (θ + 90°) can also be made content-adaptive rather than fixed, potentially improving mode retention and RD performance. Moreover, lightweight machine-learning models could be investigated to predict the most promising intra modes using wavelet and spatial features while maintaining low computational overhead. Future experiments should also incorporate larger H.264/AVC data sets, more resolutions, different video contents, and hardware-based implementations, as well as test the robustness and real-time capability of WG-CAMP. Further developments in this area could improve the mode-pruning accuracy of the system, the encoding complexity, bitrate and reconstructed video quality.

References

[1] P. Wang and H. Cheng, "Early termination of intra mode decision based on most probable mode and SATD for H.264/AVC encoding," Journal of Signal Processing Systems, vol. 92, no. 7, pp. 173–186, 2020, doi: 10.1007/s11265-019-01455-8.

[2] D. Mathew, B. M. C., and B. A. Jose, "Computationally efficient intra and inter mode decision in H.264/AVC," Procedia Computer Science, vol. 171, pp. 360–368, 2020, doi: 10.1016/j.procs.2020.04.037.

[3] L. Balaji, K. K. Thyagharajan, C. Raja, and A. Dhanalakshmi, "An optimal mode selection algorithm for scalable video coding," International Journal of Computational Vision and Robotics, vol. 10, no. 2, pp. 143–155, 2020, doi: 10.1504/IJCVR.2020.105685.

[4] J. Seiler, H. Lakshman, and A. Kaup, "Spatio-temporal prediction in video coding by best approximation," in Proc. IEEE International Conference on Image Processing (ICIP), 2022. [Online]. Available: arXiv:2207.09727.

[5] W.-Y. Lu, E. Pavez, A. Ortega, X. Zhao, and S. Liu, "Adaptive online learning of separable path graph transforms for intra-prediction," in Proc. Picture Coding Symposium (PCS), 2024, doi: 10.1109/PCS60826.2024.10566291.

[6] Z. Chen, J. Shi, and W. Li, “Learned fast HEVC intra coding,” IEEE Transactions on Image Processing, 2020. [Online]. Available: arXiv:1907.02287.
[7] S. Hamdy, A. Zekry, and W. A. Mohamed, “Development of the best prediction matrix mode algorithm for intra prediction in H.264 encoder,” International Journal of Engineering & Technology, vol. 7, no. 4, pp. 3595–3601, 2018, doi: 10.14419/ijet.v7i4.19497.
[8] Y. Zheng, P. Xu, J. Sun, and Z. Guo, “Gradient based interpolation for intra angular prediction in HEVC,” in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 2018, pp. 1–5, doi: 10.1109/ISCAS.2018.8350914.
[9] G. Rath, F. Racapé, F. Urban, F. Le Léannec, F. Galpin, and K. Naser, “A general framework for directional intra prediction with varying angle for video coding,” in Proceedings of the 28th European Signal Processing Conference (EUSIPCO), 2020, pp. 1–4, doi: 10.23919/Eusipco47968.2020.9287723.
[10] M. A. El-Mowafy, A. A. Abd El-Haleem, and M. M. Ibrahim, “Efficient mode decision scheme based on edge detection with Gaussian pulse for intra-prediction in H.264/AVC,” Alexandria Engineering Journal, vol. 61, no. 4, pp. 2709–2722, 2022, doi: 10.1016/j.aej.2021.07.044.
[11] Y. Zhang, N. Li, S. Kwong, G. Jiang, and H. Zeng, “Statistical early termination and early skip models for fast mode decision in HEVC INTRA coding,” ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 15, no. 3, Art. no. 70, 2019, doi: 10.1145/3321510.
[12] X. Sun, M. Han, W. Zuo, and M. Liu, “Perceptual-based HEVC intra coding optimization using deep convolution networks,” IEEE Access, vol. 7, pp. 59872–59883, 2019, doi: 10.1109/ACCESS.2019.2910245.
[13] M. Jamali and S. Coulombe, “Fast HEVC intra mode decision based on RDO cost prediction,” IEEE Transactions on Broadcasting, vol. 65, no. 1, pp. 109–122, 2019, doi: 10.1109/TBC.2018.2847464.
[14] Y. Chen, L. Yu, H. Wang, T. Li, and S. Wang, “A novel fast intra mode decision for versatile video coding,” Journal of Visual Communication and Image Representation, vol. 71, Art. no. 102849, 2020, doi: 10.1016/j.jvcir.2020.102849.
[15] J. Yang and A. Wei, “Fast mode decision algorithm for intra prediction in HEVC,” in Proceedings of the IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2020, pp. 1018–1022.
[16] Y. M. Vaidya and S. P. Metkar, “Gradient-based feature extraction for early termination and fast intra prediction mode decision in HEVC,” in Intelligent Data Engineering and Analytics, Advances in Intelligent Systems and Computing, vol. 1177, Springer, Singapore, 2021, pp. 221–227, doi: 10.1007/978-981-15-5679-1_21.
[17] R. D. Anitha Kumari and A. N. Udupa, “Optimal cost modeling scheme for efficient intra-prediction mode in video compression,” SN Computer Science, vol. 4, Art. no. 352, 2023, doi: 10.1007/s42979-023-01737-w.
[18] K. C. Pathak, A. D. Darji, and J. N. Sarvaiya, “Novel low-complex 4 × 4 and 16 × 16 intra-prediction architecture for error concealment for H.264,” in Proceedings of the Fourth International Conference on Microelectronics, Computing and Communication Systems (MCCS 2019), Lecture Notes in Electrical Engineering, vol. 673, Springer, Singapore, 2021, pp. 599–615, doi: 10.1007/978-981-15-5546-6_51.
[19] S. Bhavikatti and R. M. Banakar, “Low power design of intra prediction of H.264 baseline encoder,” SSRN Electronic Journal, 2021, doi: 10.2139/ssrn.3769771.
[20] A. Tayyebi, D. M. Hanna, and B. Jones, “Parametrized low-complexity hardware architecture of an H.264-based video encoder for FPGAs,” Microprocessors and Microsystems, vol. 105, Art. no. 105017, 2024, doi: 10.1016/j.micpro.2024.105017.
[21] V. Sildas, P. V. Chandramani, R. Srinivasan, and R. Ananthanarayanan, “Variance-based reconfigurable modules for mode decision in intra prediction algorithm,” Ain Shams Engineering Journal, vol. 16, no. 7, Art. no. 103403, 2025, doi: 10.1016/j.asej.2025.103403.