A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Zhi Liu*, Houyu Qian and Mengmeng Zhang

North China University of Technology, BeiJing, China

lzliu@ncut.edu.cn,597205668@qq.com,muchmeng@126.com

Abstract

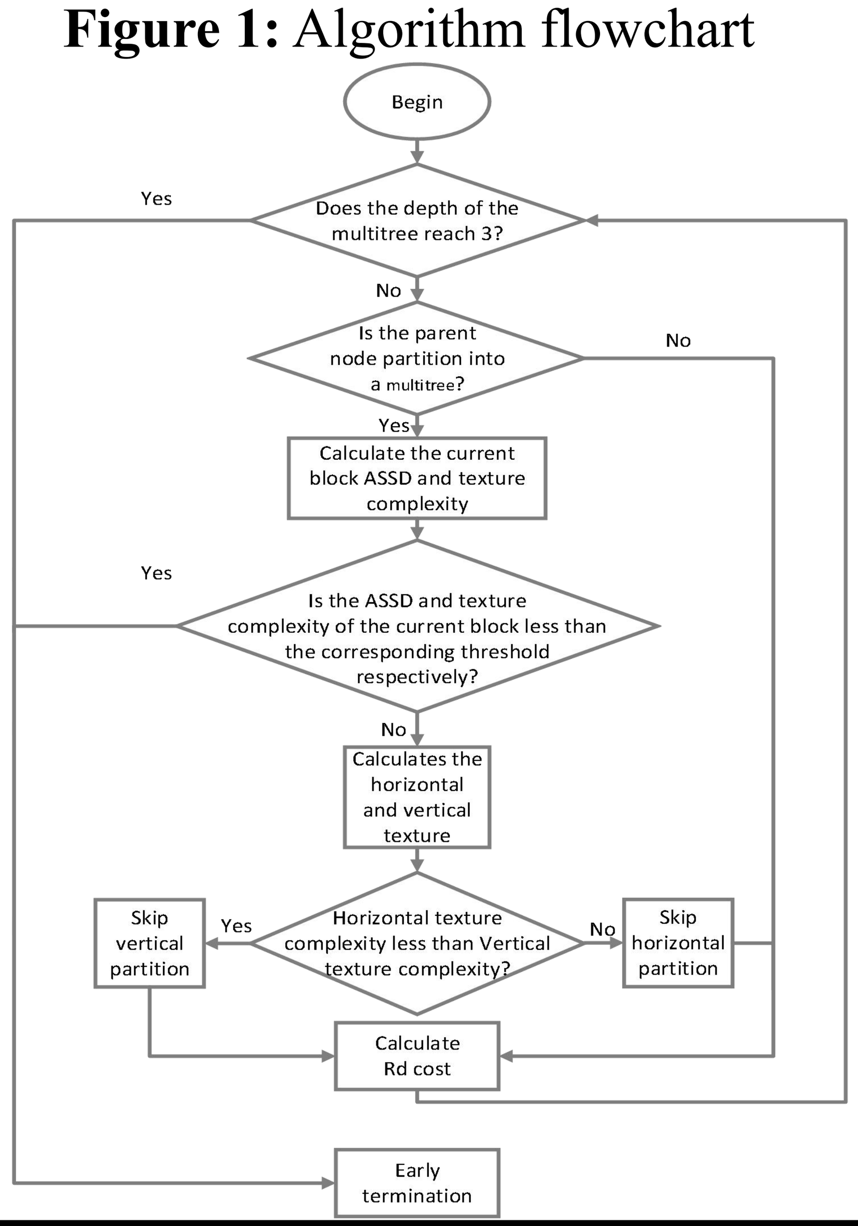
The quadtree nested multi-type trees (MTT) partitioning scheme is adopted by the latest generation of video coding standard H.266/VVC. While the performance of the encoder is improved, the complexity has increased by 2.2-5.6 times. It has been found that in the inter-coding mode, blocks with intense motion or complex texture tend to be further divided by the encoder. In this paper, a fast multi-tree partition algorithm is proposed based on both the spatial and temporal information. The motion characteristics and texture complexity of the current coding block are utilized simultaneously to determine whether the multi-tree partition can be terminated early. For blocks that need to be further divided, the partition direction is determined in advance.

The Proposed Strategy

This paper finds that, different from intra-coding mode, in inter-coding mode, areas with intense motion or complex texture will be divided more finely by the encoder, and background regions with slight motion and simple textures will be represented by large blocks. Therefore, the two indicators are used to measure the motion characteristics and texture complexity of the coding block. The Average Sum of Square Difference (ASSD) of the luma pixel values of the current block and the reference block at the same position in the co-located reference frame is used in this paper to describe the motion characteristics of the current block. The calculation formula of ASSD is shown in Equation 1.Since the Scharr operator has a larger weight of adjacent pixels and higher accuracy, it can obtain sharper edge features, so the Scharr operator is used in this paper to calculate the texture complexity of CU. The calculation formula of texture complexity is shown in Equation 2, 3 and 4. The corresponding threshold are set by the corresponding statistical mean. Tables 1 and Figure 2 show the statistical results of ASSD and texture complexity respectively. Before dividing the current block, the ASSD values and texture complexity values are calculated separately, if both are less than the threshold, Cu partition will be terminated early. Otherwise, the partition direction is further determined by the texture direction.Finally, the proportion of different partition depth is calculated, as shown in Figure 3. Because the proportion of continuous use of multi-tree partition for 4 times or more is very small and the complexity is very high, this paper proposes to terminate the partition early when the multi-tree partition depth reaches 3.

The Flowchart of Proposed Strategy

Figure 1: Algorithm flowchart



$$ASSD = \frac{1}{2N \times 2N}\sum_{i=0}^{2N-1}\sum_{j=0}^{2N-1}(CU_{Cur(i,j)} - CU_{Col(i,j)})^2 \qquad (1)$$

$$S_x = \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix}, S_y = \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix} \qquad (2)$$

$$G_x = \sum_{i=1}^{W}\sum_{j=1}^{H} A \times S_x \ , G_y = \sum_{i=1}^{W}\sum_{j=1}^{H} A \times S_y \qquad (3)$$

$$G_{scharr} = |G_x| + |G_y| \qquad (4)$$

Figure 2: ASSD average value

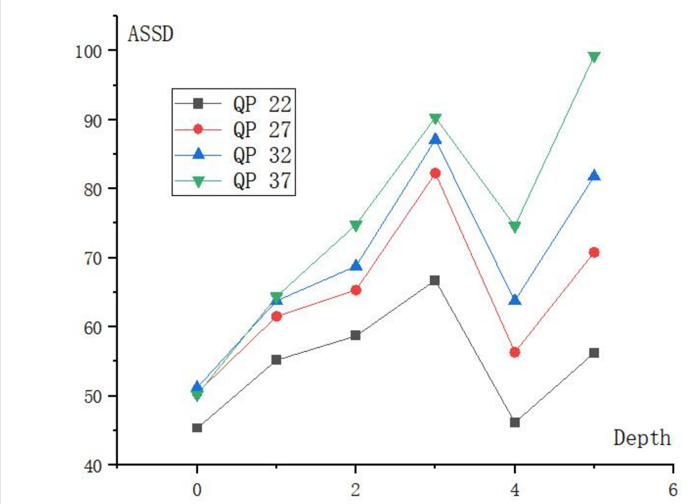
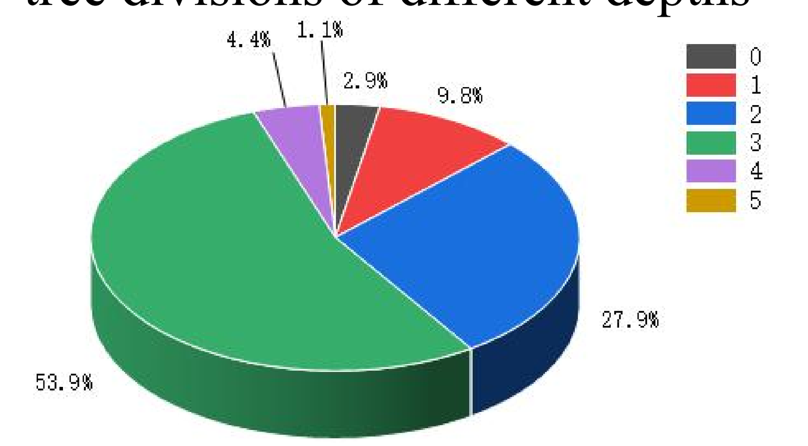


Table 1. The average value of texture complexity

Depth \ QP	22	27	32	37
0	53180	62871	75609	87442
1	36343	44287	53252	63494
2	13262	20486	39929	46495
3	9191	10969	13399	21759
4	28360	31301	34304	36689
5	34998	40600	46873	54070

Figure 3: The proportion of multi-tree divisions of different depths



Experimental Results

In order to verify the performance of the proposed algorithm, this algorithm is implemented in the VVC reference software VTM-11.2, each sequence are tested in different QPs (22, 27, 32, 37) and the encoding configuration is Randomaccess. BD-rate and TS are respectively used as the evaluation indexes of algorithm rate distortion and complexity. The definition of TS is:

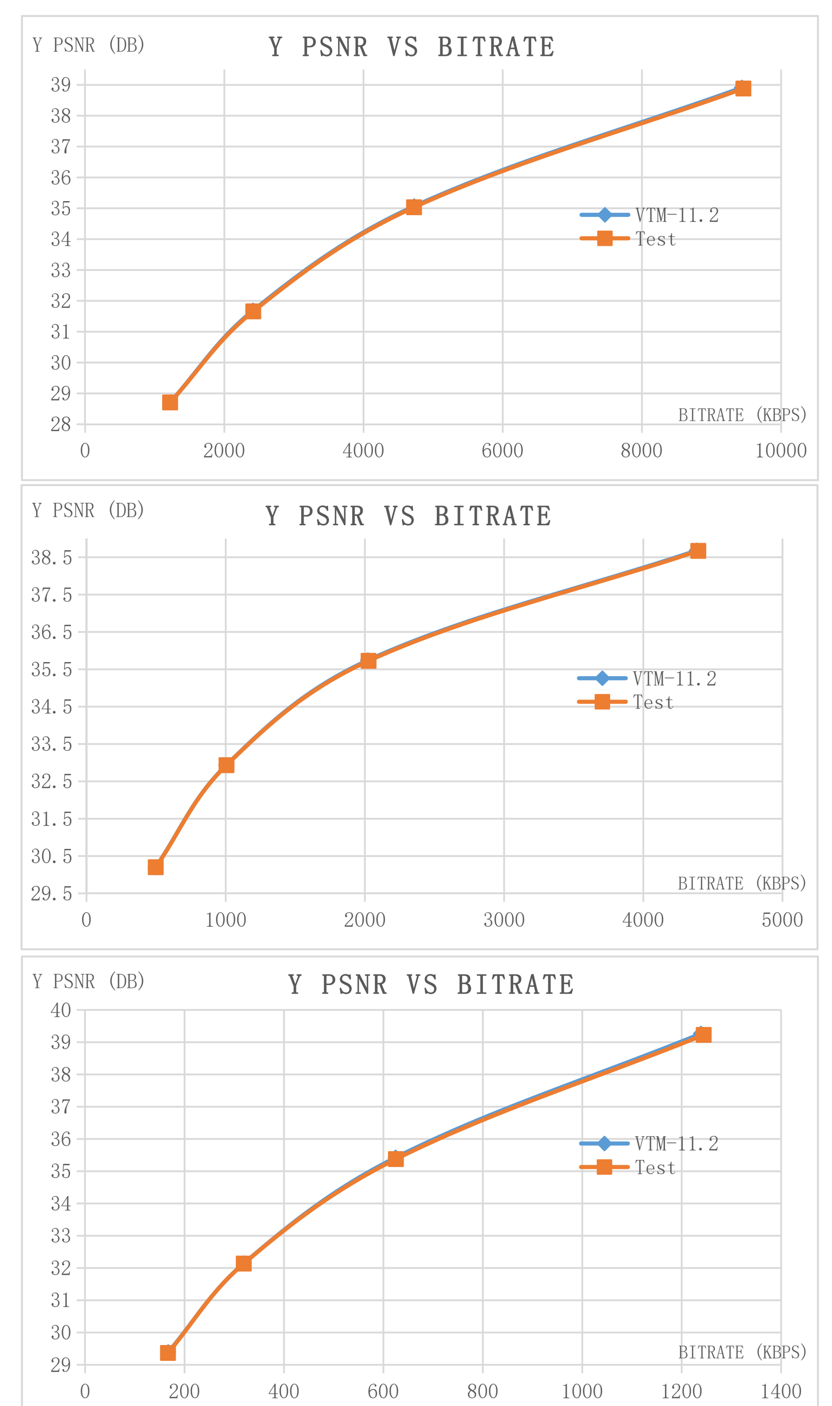
$$TS = \frac{1}{4}\sum_{QP \in \{22,27,32,37\}} \frac{Time_{VTM}(QP_i) - Time_{Pro}(QP_i)}{Time_{VTM}(QP_i)} \qquad (5)$$

The experiment chose the standard test sequence of HEVC. Different scenes and different resolutions are included in these sequences, especially those with simple backgrounds with intense motion and sequences with slight motions but complex backgrounds. Table 2 shows the performance comparison between the proposed VVC inter frame multi tree partition fast algorithm based on spatio-temporal correlation and VVC standard test model .The experimental results show that, compared with VTM-11.2, the coding time of the algorithm proposed in this paper can be saved by 24.42%, and the BD-rate only increases by 1.34%. Figure 4 shows the comparison of RD performance curves of different sequences.

Table 2. Experimental results of the proposed algorithm

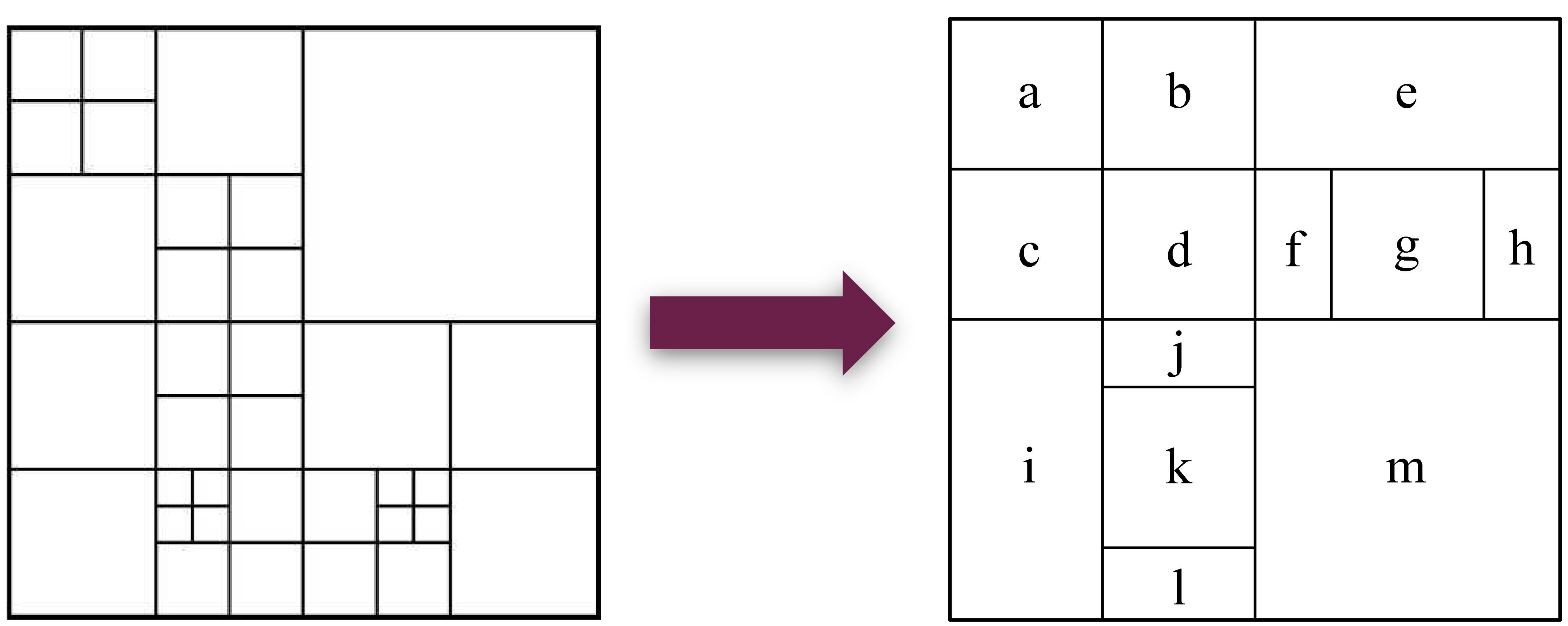
	Sequence	Random Access			TS	△PSNR (dB)
		Y	U	V		
Class B 1080p	Kimono	1.54%	-0.15%	1.60%	19.81%	-0.02
	ParkScene	1.12%	0.90%	1.22%	28.04%	-0.02
	Cactus	1.45%	1.81%	1.49%	28.04%	-0.02
	BasketballDrive	1.33%	1.58%	0.09%	24.46%	-0.02
	BQTerrace	1.49%	1.54%	0.04%	28.58%	-0.02
Class C WVGA	BasketballDrill	1.34%	1.07%	0.51%	30.46%	-0.04
	BQMall	1.52%	0.30%	1.28%	28.02%	-0.04
	PartyScene	0.57%	0.54%	0.25%	15.94%	-0.05
	RaceHorses	0.40%	1.30%	0.17%	8.61%	-0.02
Class D WQVGA	BasketballPass	2.20%	1.75%	2.93%	25.27%	0
	BQSquare	1.09%	1.55%	0.62%	18.40%	-0.05
	BlowingBubbles	1.51%	1.31%	1.06%	25.61%	-0.04
	RaceHorses	0.65%	-0.40%	0.51%	11.14%	-0.04
ClassE 720p	FourPeople	1.54%	0.70%	0.39%	32.17%	-0.03
	Johnny	2.44%	1.31%	1.54%	33.81%	-0.03
	KristenAndSara	1.99%	1.04%	1.27%	35.98%	-0.04
Class F	BasketballDrillText	1.22%	1.41%	1.33%	32.47%	-0.04
	ChinaSpeed	2.60%	2.42%	2.04%	37.32%	-0.1
	SlideEditing	0.18%	0.06%	-0.01%	12.72%	-0.01
	SlideShow	0.70%	4.88%	6.80%	11.52%	-0.01
	Average	1.34%	1.25%	1.26%	24.42%	-0.03

Figure 4:RD performance curves of different sequences



Partition mode of quadtree nested multitree in VVC

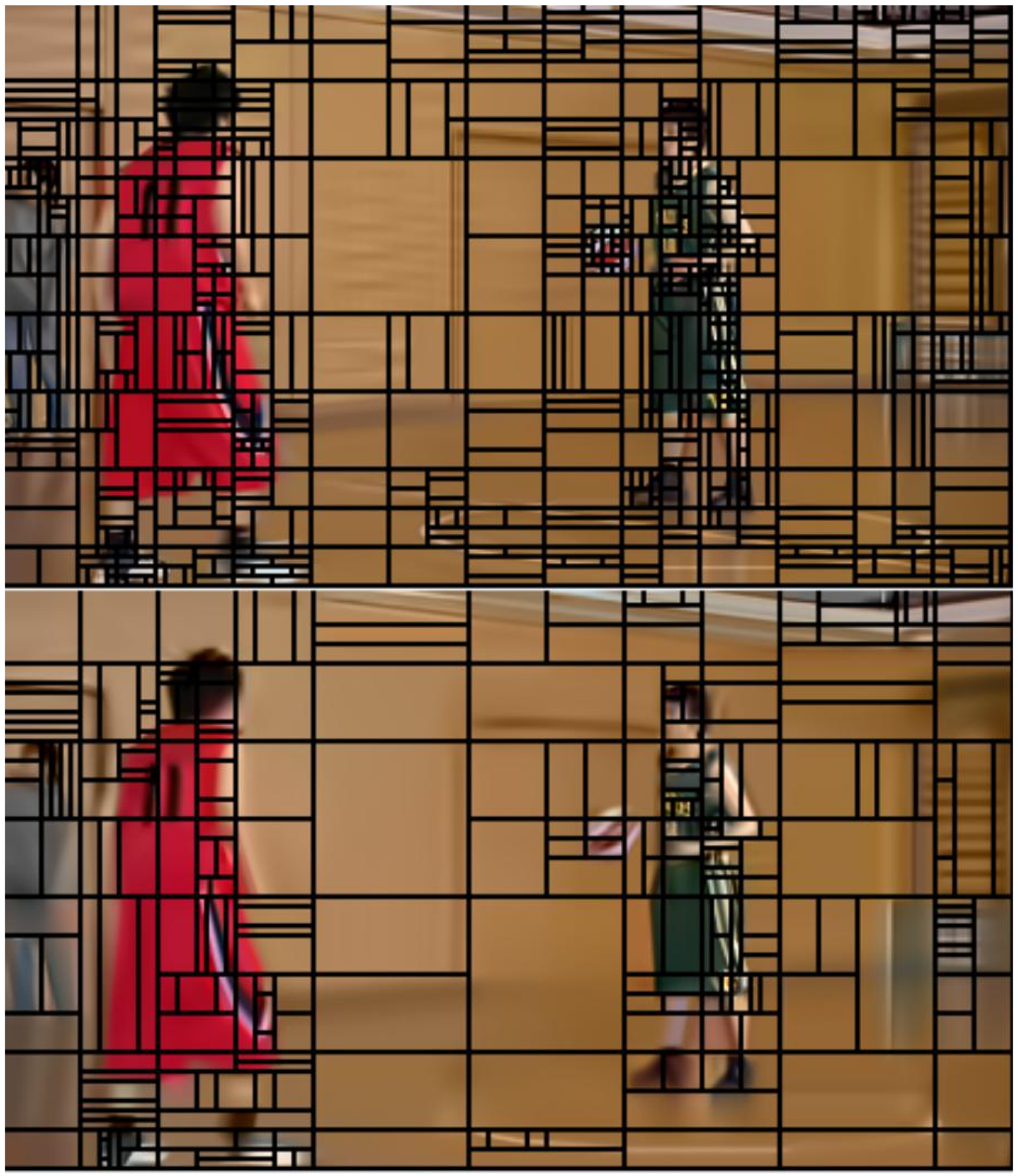
The QTBT partition structure proposed by MediaTek was adopted by JVET, replacing the quadtree structure of the previous generation video coding standard HEVC



Advantages: more flexible partition, improved encoder performance

Disadvantages: the complexity is increased by 2.2-5.6 times

Comparison of the same frame division under the AI and RA configurations



All Intra (above)

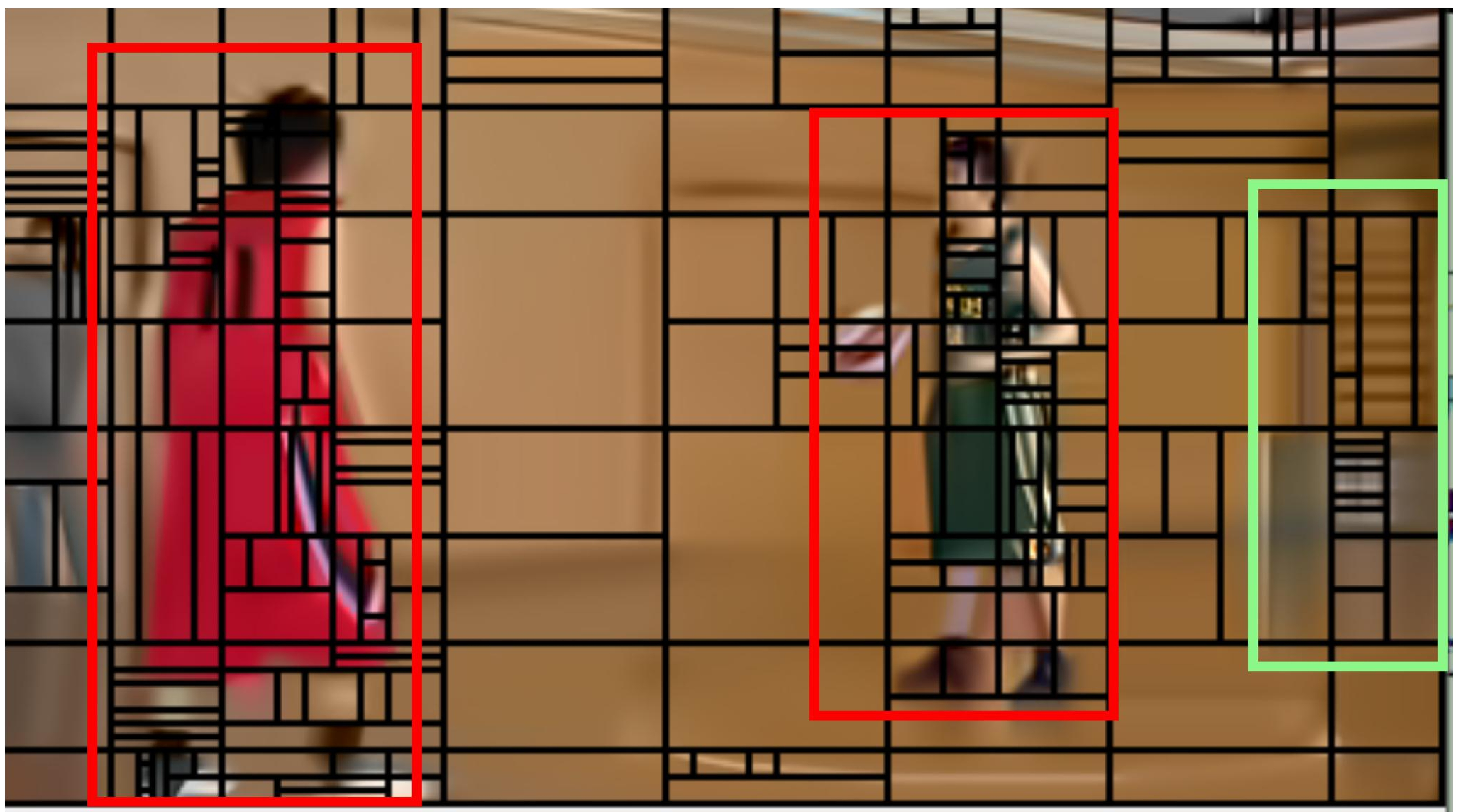
Randomaccess (below)
(High compression rate)

Existing problems:
The traditional intra block partition decision can not be well applied to inter coding mode.

It is necessary to study a method that comprehensively considers motion features and image textures to make decisions in advance for the multi-tree partition mode to reduce the complexity of the encoder

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

The characteristics of quadtree nested multi-type tree division in inter coding mode



Areas with intense motion will be divided more finely

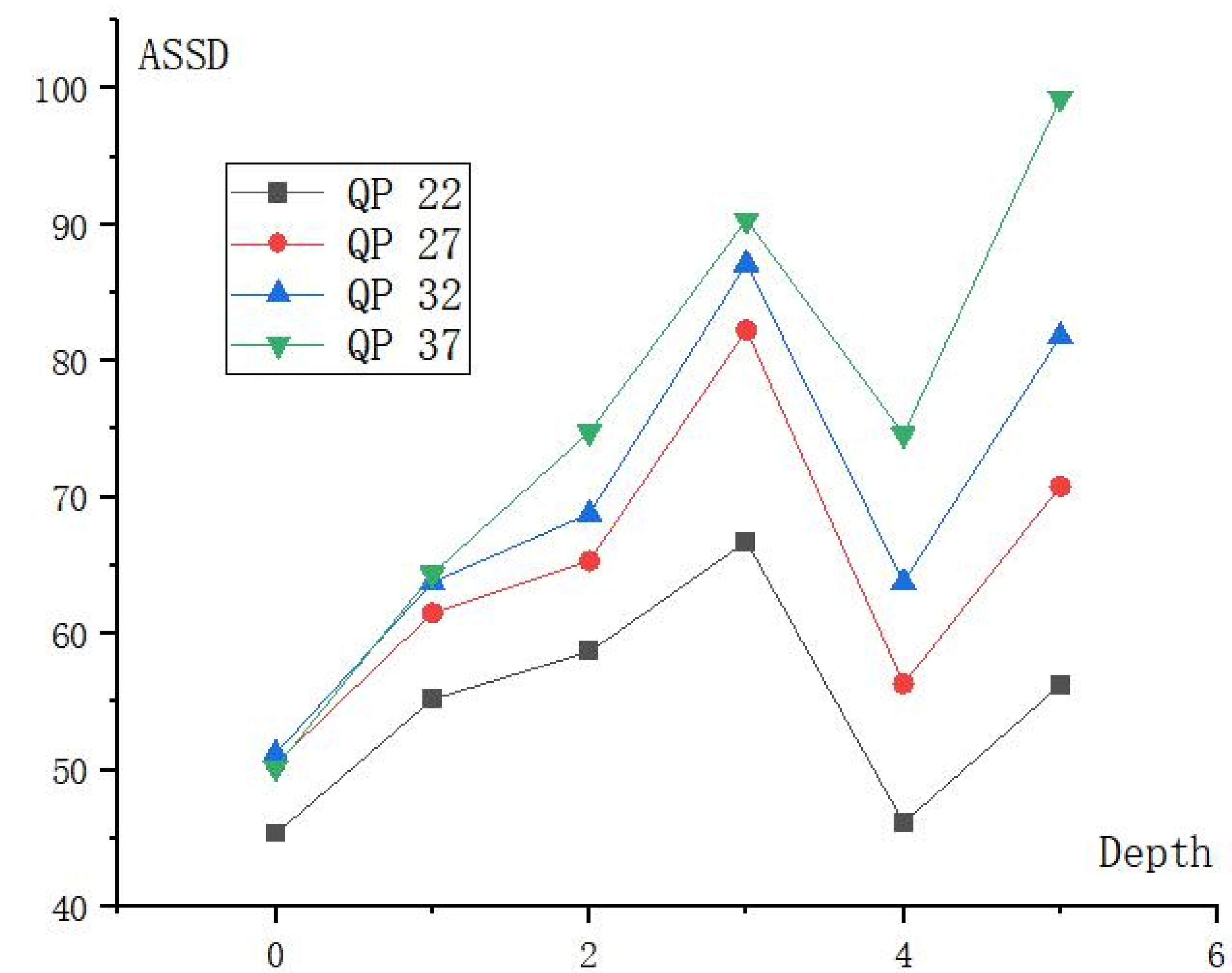
Areas with complex texture will be divided more finely

Background regions with slight motion and simple textures will be represented by large blocks

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Describe of motion characteristics of coded blocks with time correlation

$$ASSD = \frac{1}{2N \times 2N} \sum_{i=0}^{2N-1} \sum_{j=0}^{2N-1} (CU_{Cur(i,j)} - CU_{Col(i,j)})^2$$



Set the ASSD threshold by the statistical result

If current block ASSD < threshold (smooth motion)

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Description of texture features of coded blocks with spatial correlation

Scharr operator is used to calculate the texture complexity of CU:

$$S_x = \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix}, S_y = \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix} \quad (1)$$

$$G_x = \sum_{i=1}^{W} \sum_{j=1}^{H} A \times S_x \ , G_y = \sum_{i=1}^{W} \sum_{j=1}^{H} A \times S_y \quad (2)$$

$$G_{scharr} = |G_x| + |G_y| \quad (3)$$

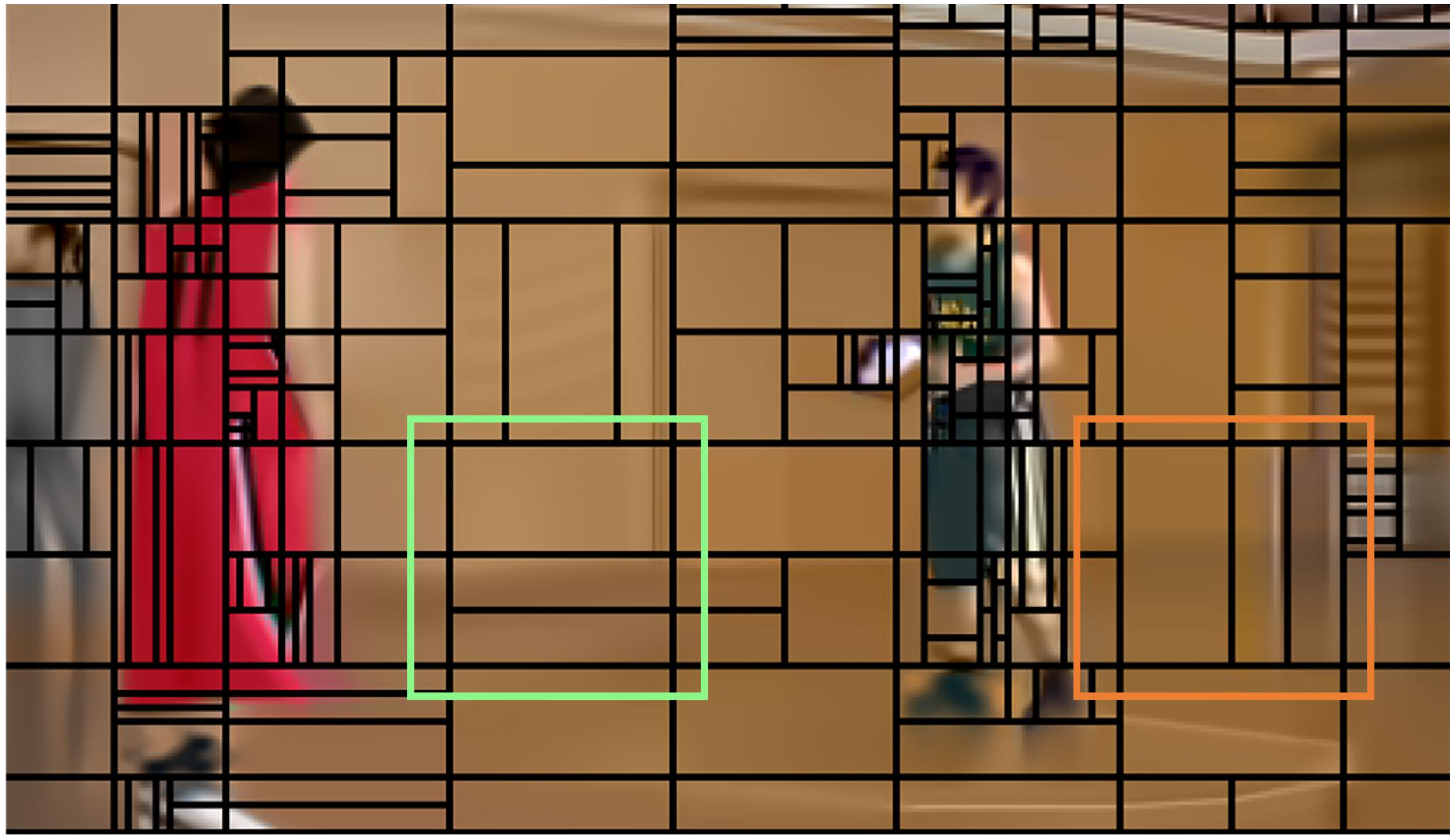
Depth \ QP	22	27	32	37
0	53180. 21	62871. 15	75609. 64	87442. 02
1	36343. 70	44287. 11	53252. 4	63494. 44
2	13262. 64	20486. 54	39929. 97	46495. 69
3	9191. 94	10969. 32	13399. 94	21759. 45
4	28360. 99	31301. 94	34304. 89	36689. 83
5	34998. 02	40600. 41	46873. 67	54070. 43

The average texture complexity of different multi tree depths under four QPS is counted. Set the texture complexity threshold from the statistical results.

If the ASSD and texture complexity of the current block are less than the threshold, the division is terminated

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Judgment of partition direction in advance



Green area:

Horizontal texture < Vertical texture

(Horizontal Partition)

Red area:

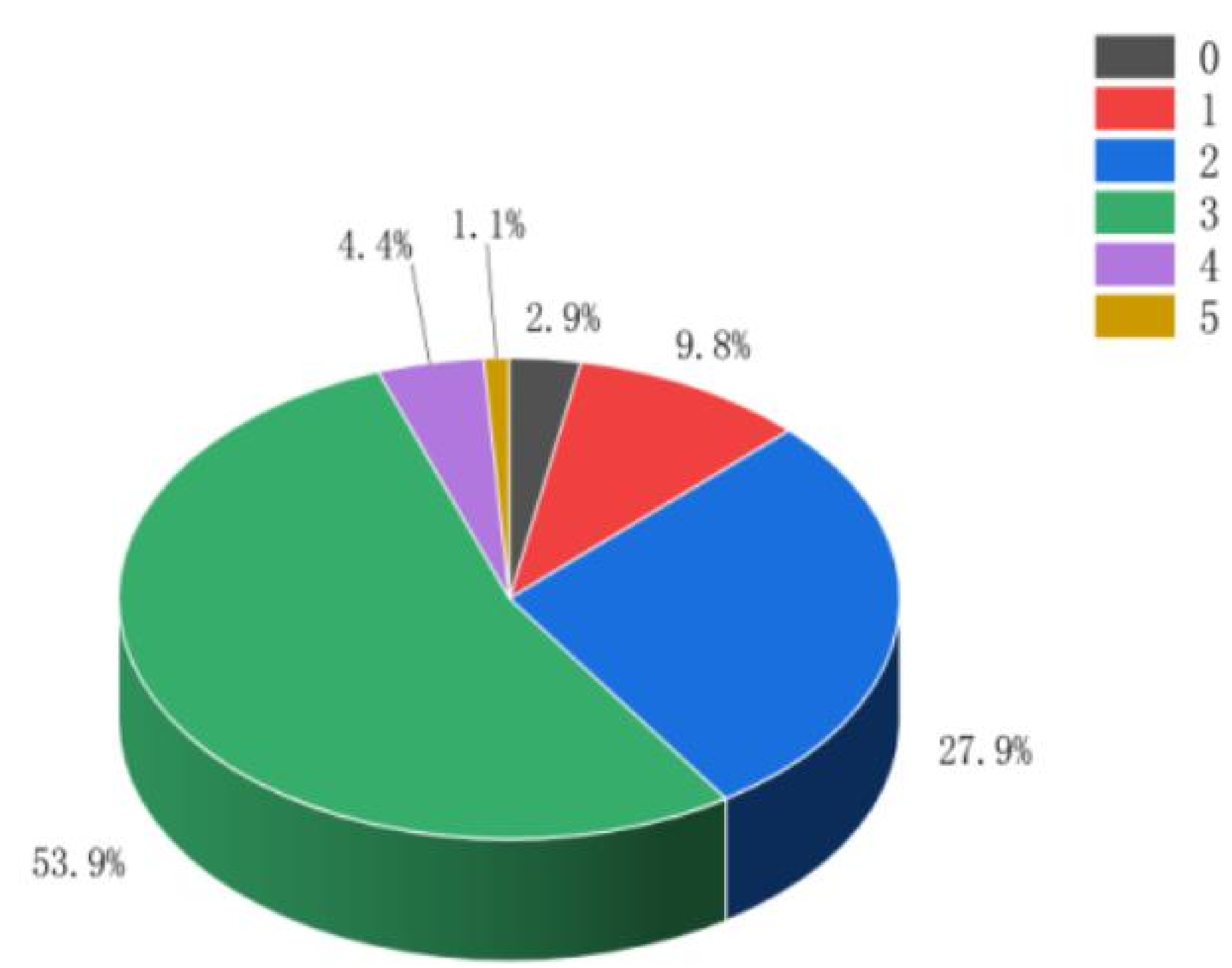
Horizontal texture > Vertical texture

(Vertical Partition)

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Early termination based on the depth of the multi-tree partition

Depth	
0	2.94%
1	9.75%
2	27.95%
3	53.89%
4	4.40%
5	1.06%



Depth 0-3:

94.56%

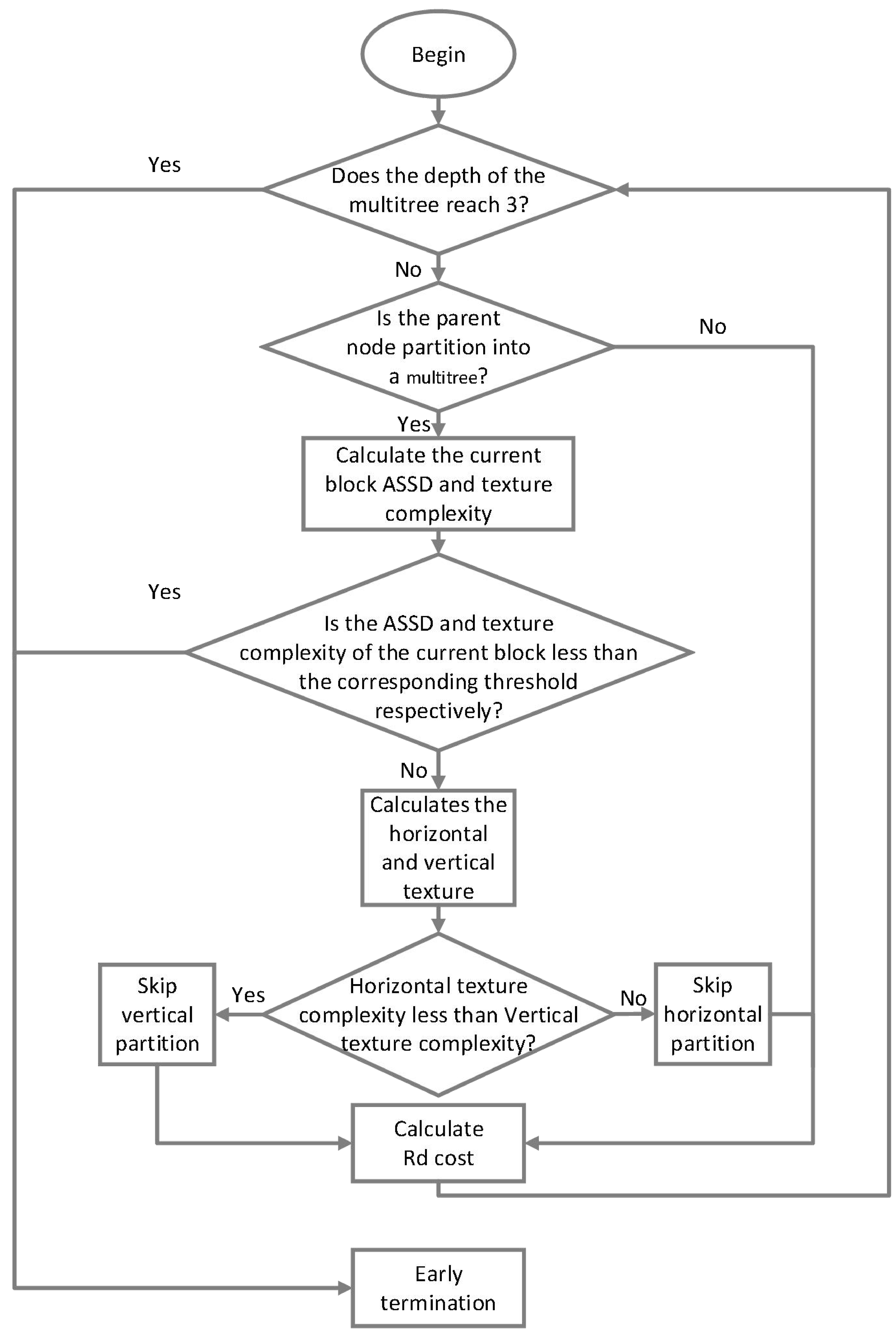
Depth greater than 3:

5.46%

A large amount of computational complexity is introduced by too deep partition. This paper proposes to terminate the partition early when the multi-tree division depth reaches 3

A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Algorithm flowchart:



A Fast Multi-tree Partition Algorithm Based on Spatial-temporal Correlation for VVC

Experimental Result

	Sequence	Random Access			TS	△PSNR（dB）
		Y	U	V		
Class B 1080p	Kimono	1.54%	-0.15%	1.60%	19.81%	-0.02
	ParkScene	1.12%	0.90%	1.22%	28.04%	-0.02
	Cactus	1.45%	1.81%	1.49%	28.04%	-0.02
	BasketballDrive	1.33%	1.58%	0.09%	24.46%	-0.02
	BQTerrace	1.49%	1.54%	0.04%	28.58%	-0.02
Class C WVGA	BasketballDrill	1.34%	1.07%	0.51%	30.46%	-0.04
	BQMall	1.52%	0.30%	1.28%	28.02%	-0.04
	PartyScene	0.57%	0.54%	0.25%	15.94%	-0.05
	RaceHorses	0.40%	1.30%	0.17%	8.61%	-0.02
Class D WQVGA	BasketballPass	2.20%	1.75%	2.93%	25.27%	0
	BQSquare	1.09%	1.55%	0.62%	18.40%	-0.05
	BlowingBubbles	1.51%	1.31%	1.06%	25.61%	-0.04
	RaceHorses	0.65%	-0.40%	0.51%	11.14%	-0.04
ClassE 720p	FourPeople	1.54%	0.70%	0.39%	32.17%	-0.03
	Johnny	2.44%	1.31%	1.54%	33.81%	-0.03
	KristenAndSara	1.99%	1.04%	1.27%	35.98%	-0.04
Class F	BasketballDrillText	1.22%	1.41%	1.33%	32.47%	-0.04
	ChinaSpeed	2.60%	2.42%	2.04%	37.32%	-0.1
	SlideEditing	0.18%	0.06%	-0.01%	12.72%	-0.01
	SlideShow	0.70%	4.88%	6.80%	11.52%	-0.01
	Average	1.34%	1.25%	1.26%	24.42%	-0.03

RD performance curves of different sequences