A fast geometric prediction merge mode decision algorithm based on CU gradient for VVC

Mengmeng Zhang, Siqi Deng* and Zhi Liu**

*North China University of Technology, BeiJing, China

muchmeng@126.com, 1345745636@qq.com, lzliu@ncut.edu.cn

Abstract

To address the problem of large computational redundancy in the geometric prediction merge mode with motion vector refinement (GPM with MMVD), a new decision algorithm based on CU gradient is proposed. By comparing the mean value of the gradient in four directions to determine whether GPM can be terminated early. The advance decision of GPM partition mode can be determined by the calculated gradient direction of CU. Meanwhile, the proposed algorithm saves coding time by reducing the number of candidate modes during the mode selection process.

The Proposed Strategy

By analyzing the characteristics of the CU using GPM, it has been found that there is a significant difference in the pixel values of the CU using GPM and the GPM partition mode is closely related to the CU texture boundary. A fast GPM decision algorithm using first-order differential Sobel edge detection operator to compute the gradient of each CU is proposed.

To process image texture features in four directions, the algorithm uses the Sobel operator template to calculate the average of the gradients in four directions: horizontal, vertical, 45° and 135°, and ranks them by magnitude. If the current image is uniformly smooth, it means that there is almost no difference between pixels and the calculated gradient means have small differences in the four directions, satisfying condition (A) and skipping the GPM prediction process. Otherwise, the GPM prediction process will continue.

$$(\overline{G_{\max}} - \overline{G_{\min}}) \le 0.85 * \overline{G_{\min}} \qquad \text{Condition (A)}$$

The frequency of GPM usage is mostly chosen in the object texture boundary region and is closely related to the direction of the edge texture, and the quantization angle can be considered as the direction of the texture boundary. The gradient direction is perpendicular to the texture boundary direction, so the GPM quantization angle φ can be quantified using the gradient direction to make advance decisions.

The Flowchart of Proposed Strategy

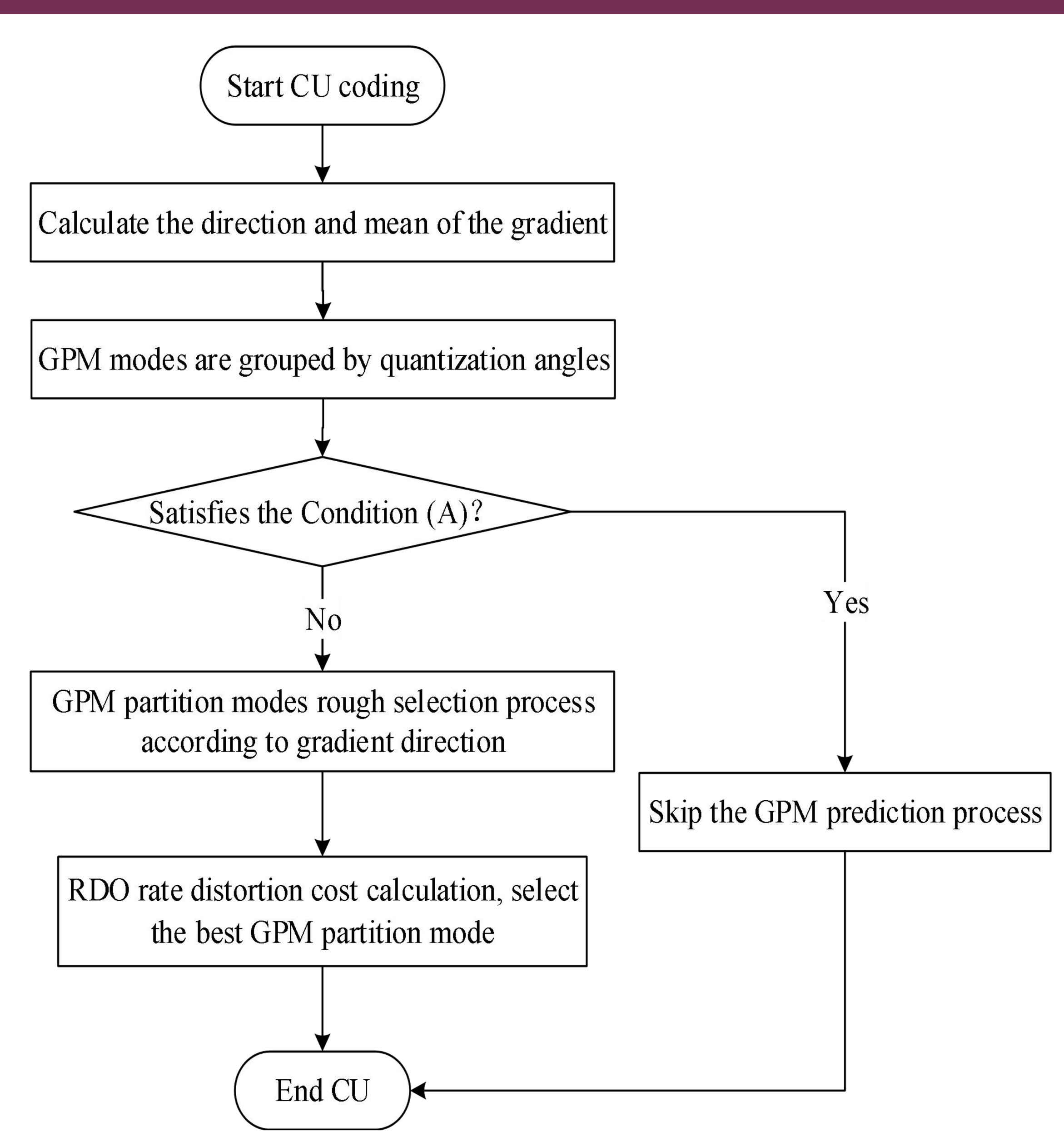


Figure 1: The algorithm flow of the proposal

Table.1: Classification of candidate modes sets

Candidate mode set categories	Quantization angles	Number of modes
horizontal	φ=0、10	4
vertical	φ=5、15	4
45°	φ=1~4	16
135°	φ=6~9	16
-135°	φ=11~14	12
-45°	φ=16~19	12

The 64 GPM partition modes are divided into six sets of candidate pattern sets according to the quantization angles, and the quantization angles are corresponded to the texture orientation of the video images, as shown in Table 1.

$$G_k = \sum_{i=1}^{w}\sum_{j=1}^{h} A \times S_k \qquad (k = x, y, 45, 135)$$

$$\overline{G_k} = \sum_{x=0}^{W-1}\sum_{y=0}^{H-1} G_k / (W \times H) \qquad (k = x, y, 45, 135)$$

$$\theta = \arctan \frac{G_y}{G_x}$$

By calculating the CU gradient direction θ, the appropriate set of GPM candidate patterns is selected and the GPM partition patterns in the unnecessary candidate pattern set are skipped. Then RDO is calculated for the partition patterns in the selected GPM candidate pattern set, and the GPM best partition pattern with the lowest cost is selected.

Experimental Results

The proposed algorithm is implemented in the VVC reference software VTM8.0 and tested for each sequence at different QPs (22, 27, 32, 37) under random access (RA) configuration. BD-rate is used to measure the rate distortion performance of the proposed algorithm. The coding time saving percentage ΔET is used to measure the reduction in computational complexity. Experimental results show that the proposed algorithm can reduce the coding time by 13.5% on average while BD-rate is almost the same as the reference algorithm with no significant loss. The experimental results are provided in the table 2 below.

Table 2. Performance comparison between the reference algorithm and the proposed algorithm

Class	Sequence	Random access Main10 BD-rate			△ET
		Y	U	V	
Class A 4K	Traffic	0.06%	0.11%	0.36%	11.5%
	NebutaFestival	-0.02%	-0.22%	-0.12%	15%
	PeopleOnStreet	0.14%	0.74%	0.63%	15.75%
Class B 1080p	Kimono	-0.03%	0.07%	0.51%	7.75%
	ParkScene	0.26%	0.19%	1.16%	17%
	Cactus	0.10%	0.52%	0.42%	12.25%
	BasketballDrive	0.42%	-0.40%	0.20%	12.5%
	BQTerrace	0.00%	0.04%	0.03%	11.5%
Class C WVGA	BasketballDrill	0.20%	0.04%	0.31%	13.5%
	BQMall	0.21%	-0.13%	0.06%	14.5 %
	PartyScene	0.11%	0.32%	0.13%	14.75 %
	RaceHorsesC	0.38%	0.48%	0.04%	13.5 %
Class D WQVGA	BasketballPass	-0.11%	0.93%	-0.02%	12.75 %
	BQSquare	0.16%	-0.20%	0.63%	17 %
	BlowingBubbles	0.45%	0.43%	0.61%	14.25 %
	RaceHorses	0.39%	1.10%	0.82%	14.25 %
ClassF	BasketballDrillText	0.38%	-1.13%	0.54%	15.5%
	ChinaSpeed	0.24%	0.62%	0.36%	14.5%
	SlideEditing	-0.18%	-0.46%	-0.51%	8%
	Average	-0.34%	-0.46%	0.18%	12%
Average		0.14%	0.13%	0.31%	13.5%

2022 Data Compression Conference

Geometric Prediction Merge Mode(GPM)

The CU of GPM is divided into two non-rectangular blocks of overlapping partitions MV0 and MV1

4 distances and 20 quantization angles

64 GPM partition modes

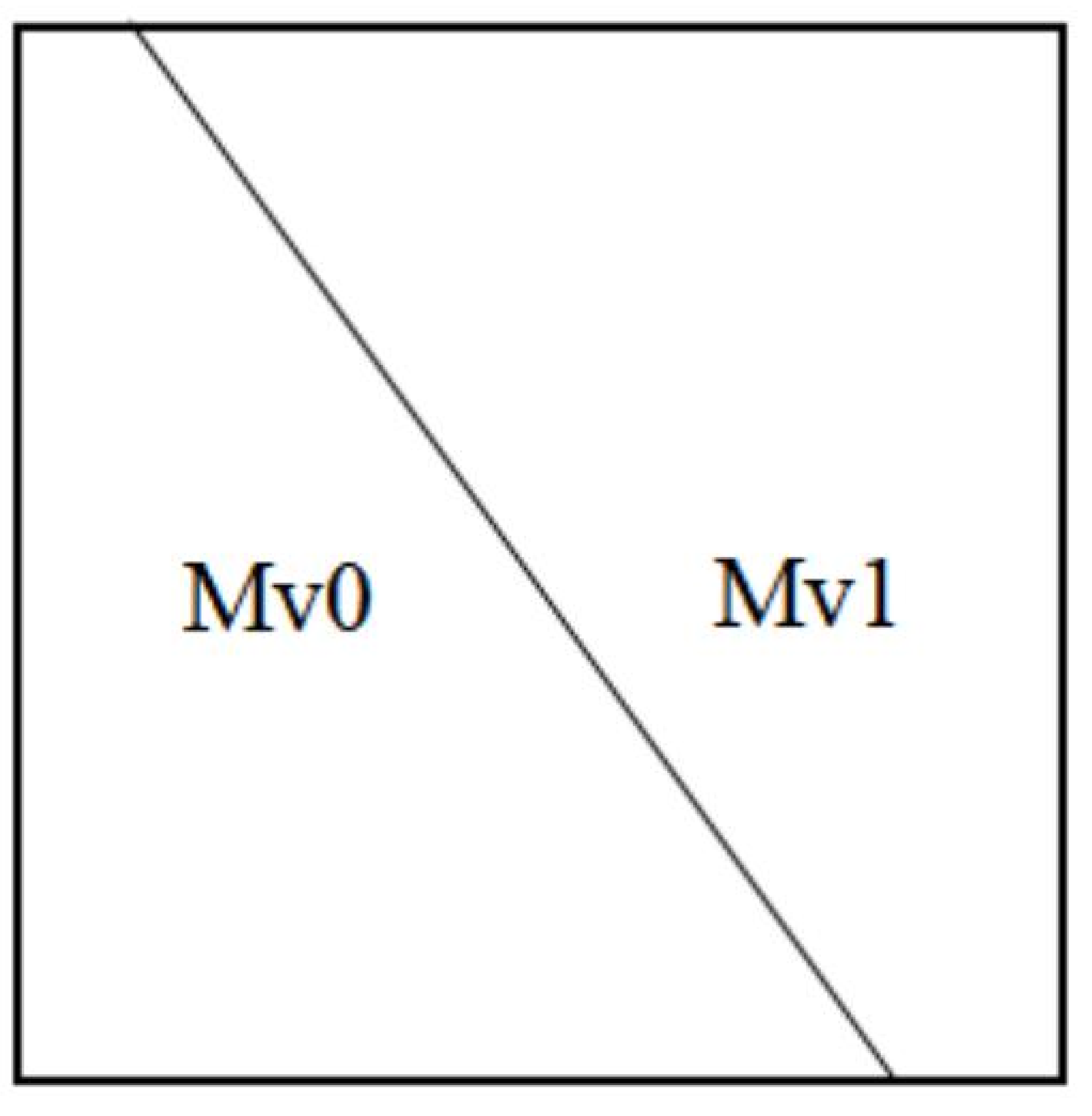


Figure 2: CU divided into non-rectangular block partitions

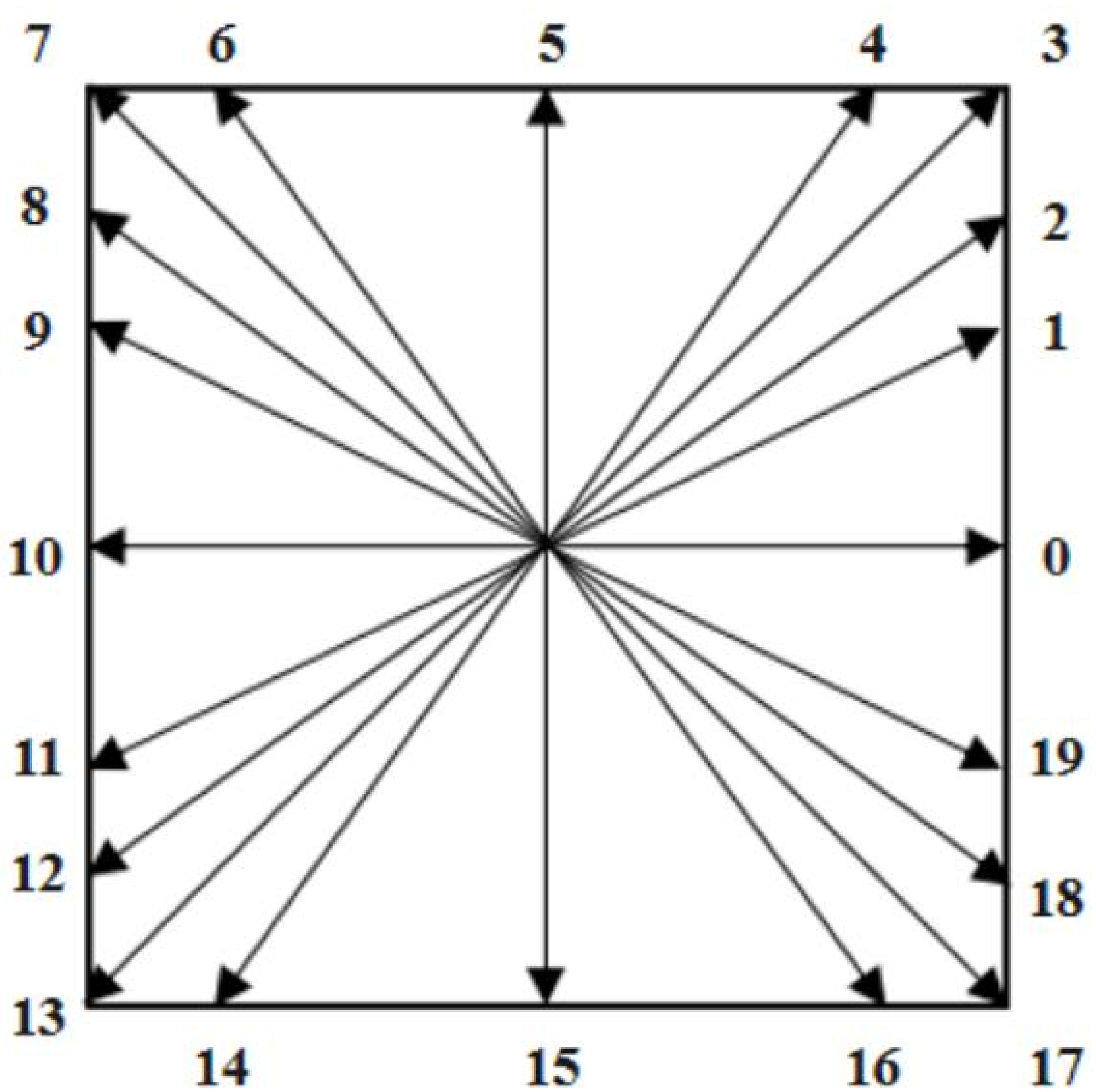


Figure 3: 20 quantitative perspectives of GPM

Geometric Prediction Merge Mode with Motion Vector Refinement(GPM with MMVD) allows for motion vector refinement on base of GPM, a CU traverses 20 × 64 × 32 = 40960 cases, resulting in a large increase in coding time. Therefore, optimization algorithms to reduce the encoding time are necessary.

Analyzing the CU characteristics of GPM

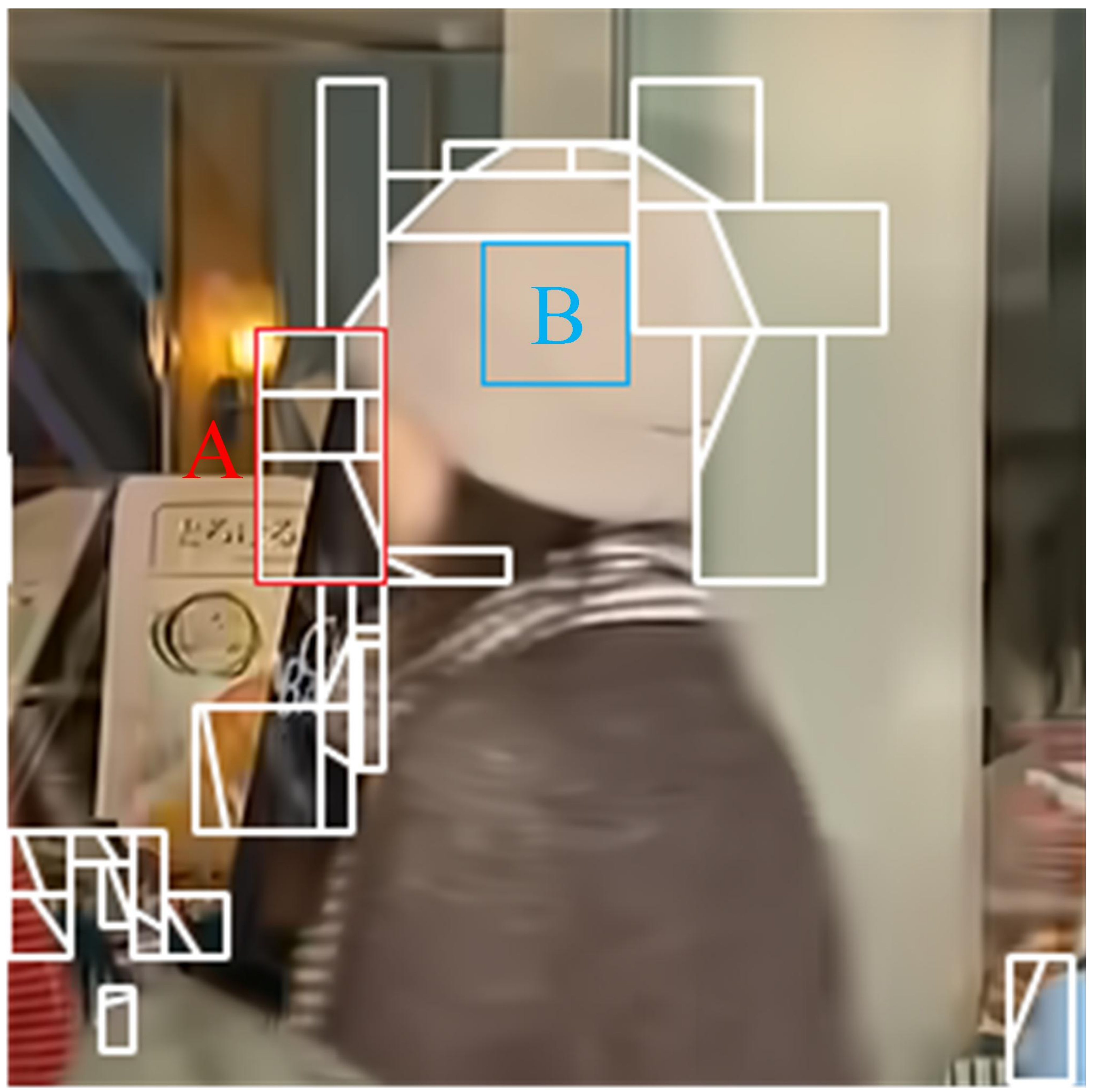


Figure 4: Geometric prediction merge mode partition

Figure 4.A is the CU of GPM, which is divided two sub-blocks by GPM. It can be found that there exists obvious differences between the left and right parts of the CU and partitioning is highly dependent on texture edges.

Figure 4.B is the CU without GPM, which is uniformly flat images with almost no difference between pixel values.

Calculation of CU gradient mean and gradient direction

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \tag{1}$$

$$S_{45} = \begin{pmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \qquad S_{135} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} \tag{2}$$

$$G_k = \sum_{i=1}^{W} \sum_{j=1}^{H} A \times S_k, (k = x, y, 45, 135) \tag{3}$$

$$\overline{G_k} = \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} G_k / (W \times H), (k = x, y, 45, 135) \tag{4}$$

$$\theta = arctan \frac{G_y}{G_x} \tag{5}$$

Early Termination of GPM

$$(\overline{G_{max}} - \overline{G_{min}}) \leq 0.85 * \overline{G_{min}} \qquad Condition\ (A)$$

Rank the gradient means from largest $\overline{G_{max}}$ to smallest $\overline{G_{min}}$.

Whether condition A is satisfied to determine the existence of pixel discrepancy in CU for early termination of GPM.

If condition (A) can be satisfied, the GPM prediction process is skipped.

If condition (A) cannot be satisfied, the GPM prediction process can be continued.

GPM Partition Mode Early Decision

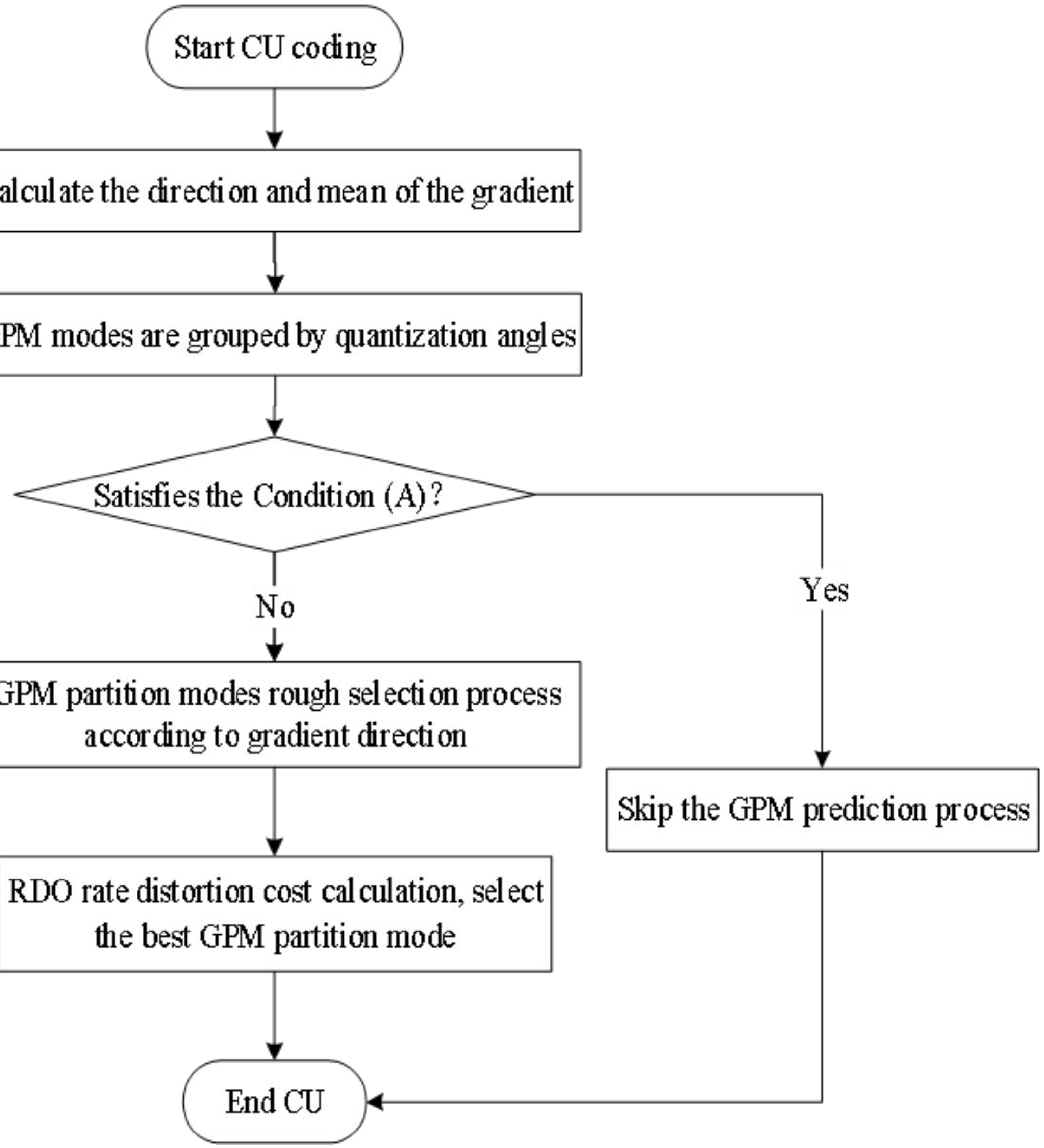
Table.1: Classification of candidate modes sets

Candidate mode set categories	Quantization angles	Number of modes
horizontal	φ=0、10	4
vertical	φ=5、15	4
45°	φ=1~4	16
135°	φ=6~9	16
-135°	φ=11~14	12
-45°	φ=16~19	12

The 64 GPM partition modes are divided into 6 groups according to the different quantization angles, such as vertical, horizontal, 45 degrees, 135 degrees, -45 degrees, and -135 degrees.

By calculating the CU gradient direction θ ,the appropriate GPM candidate modes set is selected, the GPM partition modes in the unnecessary candidate modes set are skipped, and then RDO is calculated for the partition modes in the selected GPM candidate modes set, finally the GPM best partition mode with the lowest cost is selected.

The Flowchart of Proposed Strategy



A fast geometric prediction merge mode decision algorithm based on CU gradient for VVC

Experimental Result

Table 2. Performance comparison between the reference algorithm and the proposed algorithm

Class	Sequence	Random access Main10 BD-rate			△ET
		Y	U	V	
Class A 4K	Traffic NebutaFestival PeopleOnStreet	0.06% -0.02% 0.14%	0.11% -0.22% 0.74%	0.36% -0.12% 0.63%	11.5% 15% 15.75%
Class B 1080p	Kimono ParkScene Cactus BasketballDrive BQTerrace	-0.03% 0.26% 0.10% 0.42% 0.00%	0.07% 0.19% 0.52% -0.40% 0.04%	0.51% 1.16% 0.42% 0.20% 0.03%	7.75% 17% 12.25% 12.5% 11.5%
Class C WVGA	BasketballDrill BQMall PartyScene RaceHorsesC	0.20% 0.21% 0.11% 0.38%	0.04% -0.13% 0.32% 0.48%	0.31% 0.06% 0.13% 0.04%	13.5% 14.5 % 14.75 % 13.5 %
Class D WQVGA	BasketballPass BQSquare BlowingBubbles RaceHorses	-0.11% 0.16% 0.45% 0.39%	0.93% -0.20% 0.43% 1.10%	-0.02% 0.63% 0.61% 0.82%	12.75 % 17 % 14.25 % 14.25 %
ClassF	BasketballDrillText ChinaSpeed SlideEditing Average	0.38% 0.24% -0.18% -0.34%	-1.13% 0.62% -0.46% -0.46%	0.54% 0.36% -0.51% 0.18%	15.5% 14.5% 8% 12%
Average		0.14%	0.13%	0.31%	13.5%

RD Performance curves of different sequences