International Journal of Control and Automation
Vol. 8, No. 2 (2015), pp. 83-94
http://dx.doi.org/10.14257/ijca.2015.8.2.09

Fast Algorithm with Early Termination CU Split and Mode Decision

Yong Gan, Xiaorong Zhao and Qiuwen Zhang

School of Computer and Communication Engineering, Zhengzhou University of Light Industry, 450002, Zhengzhou, China
ganyong@zzuli.edu.com

Abstract

HEVC (High Efficiency Video Coding) is the next-generation video coding standard, which introduced many new technologies such as tree structure coding unit and a variety of intra prediction mode. Compared with H.264/AVC, HEVC has achieved higher coding efficiency but requires a very high computational complexity. So this paper proposes a fast algorithm through early termination of the CU (Coding Unit) splitting and mode decision. Based on the research of image features, we use the variance of the input image to early terminate CU splitting. In addition, based on the analysis of candidate modes which were obtained by the rough mode decision, the adjacent mode is used to early terminate the RDO process to reduce the computational complexity. The experimental results show that the proposed algorithm saves averagely about 40% of encoding time with neglecting degradation of coding performance.

Keywords: *HEVC, mode decision, intra prediction*

1. Introduction

In recent years, digital video especially digital HD video has become a dominant format of media content in many scopes of consumption. Video coding is one of the extremely important support technologies, so as to facilitate transmission of digital video. In order to improve the coding efficiency of high resolution video, the Joint Collaborative Team on Video Coding (JCT-VC) developed a new generation of video coding standard HEVC [1, 2]. Compared with H.264/AVC, the latest video compression standard HEVC is designed to improve the coding efficiency, and it is reduced bit rate at the expense of increasing the computational complexity. Therefore, how to reduce the complexity of the HEVC encoder has become a significant research direction of HEVC video coding [3, 4].

In fact, researching the fast mode decision algorithm for HEVC has become one of hot spots in the field of video coding technology, which has broad prospects. Some fast intra prediction algorithms for HEVC have been proposed recently. Piao, *et al.,* [5] estimated the HAD cost by calculating the SATD (Sum of Absolute Transformed Difference) for 35 kinds of intra modes, and then selected N best modes out of 35 kinds of modes as candidate modes. In this way, we just needed to carry out the rate distortion optimization (RDO) algorithm in the N candidate modes (RDO) to find the optimal intra prediction mode on the HEVC software HM1.0. In addition to select N candidates, Zhao Liang, *et al.,* [6] made full use of the correlation between neighboring pixels in the HM2.0 and added the MPM (the best prediction mode of the upper and left PU) to the candidate sets, which improved the video quality. Jiang, *et al.,* [7] proposed a fast intra coding algorithm, detected whether the current coding unit belongs to the smooth region. According to the detection result, it can help determine how to adaptively skip the coding process of the coding unit sizes not suitable for the texture features.

ISSN: 2005-4297 IJCA

2. HEVC Intra Coding Technology

A new generation of video coding standard HEVC introduces many new technologies including recursive tree structure unit and the big block transformation [8]. Block structure is expanded on the basis of H.264/AVC, including three parts: Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU). The separation of the block structure of each unit is very favorable to the optimization of the algorithm. Starting from the Largest CU, each CU block can recursively split into 4 sub-blocks which have the same size until CU has been divided into the blocks whose size are 8. This process will develop a coding tree structure which consists of CU block based on adaptive content. The size of the best CU is decided by the intra prediction and the rate distortion cost function. The prediction unit PU is only at leaf nodes of quad-tree structure. For the intra case, both CU and PU are squared size with either one 2N×2N block or 4 N×N sub-blocks after further splitting, N∈ [4; 8; 16; 32]. In order to further improve the compression efficiency, HEVC also uses a recursive partition for transform unit. Figure 1 gives the recursive structure of the CU, PU, and TU [9].

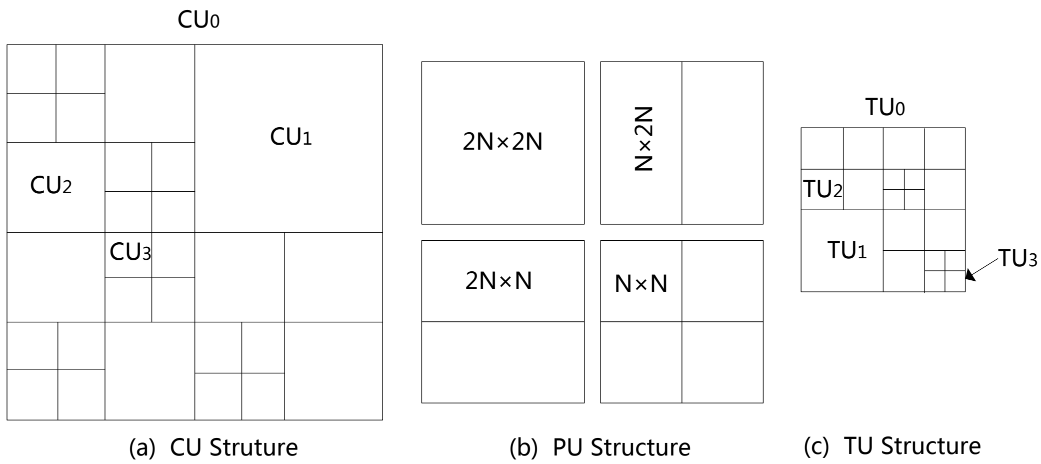
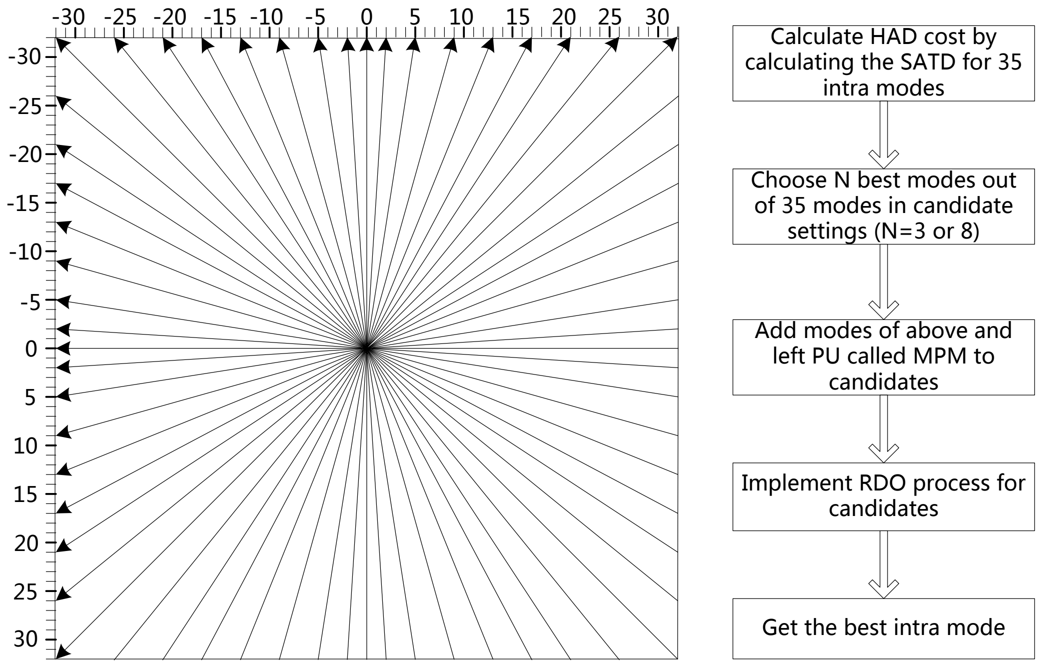


Figure 1. Recursive Block Structure for HEVC

In order to find the optimal split of block structure, the HEVC encoder needs to consider all combinations among CU, PU, and TU, which is an extremely time-consuming process. In addition, intra prediction supports 35 kinds of modes including Planar mode, DC mode and 33 kinds of mode in HEVC, as shown in Figure 2 (a). In order to find the best mode of the current coding block, it needs to traverse the all 35 kinds of modes, this progress will also produce tremendous computation. Therefore, we should find a fast mode selection algorithm to reduce the encoding time. Figure 2 (b) shows the algorithm flow of intra mode decision in HM.

The aforementioned fast algorithms for HEVC intra prediction in section 1 all can save the coding time in varying degrees. However, the correlation between candidate modes after rough mode decision (RMD) algorithm has not been studied. The first two modes which are chosen as the optimal mode in the candidate set have high probability. With the decrease of the ranking order in RMD, the probability which is selected from the rest also appears a decreasing trend. The correlation of adjacent mode can early terminate RDO process so as to reduce the computational complexity of the encoder. In addition, it can quickly determine the CU size based on the variance of the image [10]. Therefore, this paper presents a fast algorithm with early termination CU splitting and mode decision. The algorithm reduces the computational complexity and saves the coding time on the premise that keeps video quality.



(a) (b)

Figure 2. (a) 33 Kinds of Intra Prediction Modes in HEVC (b) Algorithm Flow of Intra Mode Decision in HM

3. Proposed Intra Prediction Algorithm

3.1. Early Terminate of CU Split based on the Variance of the Image

In HEVC, the coding block adopts quad-tree structure, so it will takes a very long time to determine the optimal CU size which needs to compute rate distortion cost function of CU on each depth. We propose a fast algorithm using the variance of the input image quickly to determine the optimal size of CU and early terminate of the CU division process. It can reduce the encoding time while maintaining the flexibility of CU.

The size of CU affects by the features of to-encode image. For example, most of the CU with 64×64 located in the smooth region of the to-encode image, and it can improve the compression efficiency. However, if these CUs are applied to the complex regions of the image, it cannot achieve the high compression effect because of using large size of CU can increase the prediction error. Therefore, it is very necessary to use small size of CU in HEVC when we encode complex regions. However, it is a complicated process to determine the size of CU, which needs to calculate the rate distortion cost for CU on each depth, which would take a very long time. In order to reduce the computational complexity of intra prediction, we propose a method of using the variance of image to quickly determine the size of CU.

The proposed algorithm uses the variance as a feature value of an image, calculated as shown in formula one. If the calculated variance of the image block is small, we can deem that the image blocks locate in the flat region, otherwise, we can deem that the image blocks locate in the complex region. It is based on this idea to quickly determine the size of CU.

$$\text{var} = \frac{1}{N^2}\sum_{i=0}^{N}\sum_{j=0}^{N}\left(ave - I_{(i,j)}\right)^2 \qquad (1)$$

Formula one shows the calculated variance of the image block, and the result is the variance of the CU block with N × N pixel. In formula one, var represents the variance of the image block, ave represents the average of the image luma blocks, and $I_{(i,j)}$ represents the luma value of the image blocks in the position (i, j). We use the formula one to calculate the variance for CU on each depth. If the calculated variance is great, we will use the small size of CU. Conversely, if the calculated variance is tiny, we will use the large size of CU. In order to calculate the variance of CU on each depth, formula one would be incorporated into the HM.

The size of CU is decided by comparing the variance of the image block with the threshold value in our proposed algorithm, so the threshold selection determines the size of CU, and thereby affecting image quality. Through the experiment and analysis of a large number of video sequences, we find that only small changes in the image quality occur when the threshold value sets between 100 and 200. We would set the threshold value equals 100 in this paper. To-encode image is first divided into multiple LCUs, and their depth is set to 0, and then we will calculate the variance of CU. Finally, comparison of the calculated variance and the threshold value are given. If the variance is greater than the threshold value, it will skip the calculation of the current CU and directly split it with the depth plus one. Conversely, if the variance is not greater than the threshold value, it will directly use the current CU for the optimal CU. Since it does not need to implement all the depth of intra prediction in our proposed algorithm, compared with the traditional method, it reduces the time need to determine the size of CU.

3.2. Early Terminate Mode Decision using the Correlations of the Modes

We jointly use RMD process and RDO process to select the optimal intra prediction mode of the current prediction unit in our proposed intra mode decision algorithm. HM decides to select the first N best candidate modes based on the RMD process where all modes are tested by minimum absolute sum of Hadamard transformed coefficients of residual signal and the mode bits [11]. RDO process only uses in the N candidate modes which have selected. It is unnecessarily to calculate the rate-distortion cost on the 35 kinds of modes for prediction unit, reducing intra prediction time. The number of N best modes in RMD process is shown in Table 1.

Table 1. Number of N in the RMD Process

The size of prediction unit	Number N
64× 64	3
32 × 32	3
16 × 16	3
8 × 8	8
4 × 4	8

However, the computational load of the encoder still keeps high. Because of the correlation between pixels, MPM reflects direction information of the current coded block. According to this point, we improve the intra mode decision algorithm. We use the correlation between the best intra prediction modes to early terminate the RDO process and thereby reduce the computational complexity of the encoder. Details of the algorithm are described as follows.

Experimented on a large number of video sequences, we can observe that candidate modes selected from RMD process present a descending trend to be the RDO-optimal mode according to their rank in candidates. The probability is approximately 60% which the first candidate mode with minimum cost is chosen as the best mode, as shown in

Figure 3. Meanwhile, the adjacent blocks have the similar texture characteristics in the one image. Therefore, the best intra prediction mode of the current CU has a strong correlation with the adjacent CU. We respectively estimate the probability of the first candidate mode from RMD process and MPM from the spatially adjacent CU as the best intra prediction mode of the current CU. MPM from the spatially adjacent CU are acquired from the left, top, upper left and upper right CU. According to the result, we know that the probability of the first candidate mode from RMD as the best intra prediction mode of the current CU is greater than 55%, and the probability of the second candidate mode from RMD as the best intra prediction mode is approximately 15%. Generally, the probability of MPM from the spatially adjacent CU as the best intra prediction mode is approximately 34%. So the first candidate mode from RMD process and MPM from the spatially adjacent CU as the best intra prediction mode has a high percentage.

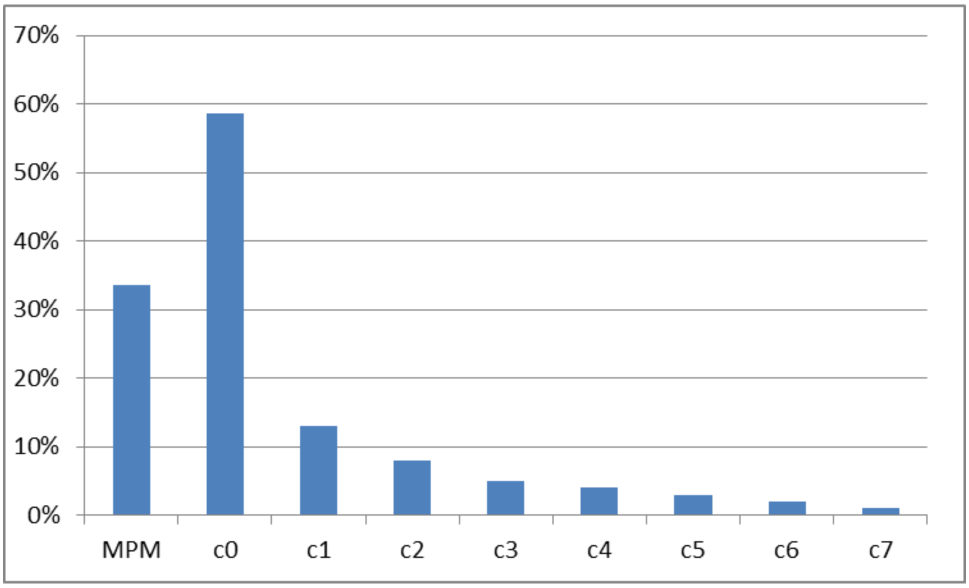


Figure 3. The Possibility to be Selected as the Best Prediction Mode

Thus, we must first check whether MPM from the spatially adjacent CU and the first candidate mode from RMD process have the same mode in our proposed algorithm. If they have the same intra prediction mode, we will choose the first candidate mode from RMD process as the best mode, and skip the other modes. In this way, it will save a lot of encoding time.

3.3. Early Terminate the RDO Process based on the Adjacent Mode

As in this heading, they should be Times New Roman 11-point boldface, initially capitalized, flush left, with one blank line before, and one after.

In the HM, selecting the first N minimum cost modes to compose a set Ψ after the process of RMD. These N candidate modes are further encoded using the RDO process to get the best prediction mode. In the Ψ, N candidate modes are ordered in accordance with the Hadamard cost from low to high, for example, $Mode_{m_{(i)}} \leq Mode_{m_{(i+1)}}, (1 \leq i \leq N)$. Finally, it will select the best mode from the N candidate modes. This paper proposes an algorithm based on adjacent mode and Hadamard cost to early termination RDO process. We can see that, from Figure 3, the first two candidate modes have great possibility are chosen as the best mode in set Ψ. Therefore, the two modes are always confirmed by the full RDO process in this paper. In addition, using $\left|m_i - m_j\right|$ represents the distance of the two modes between m_i and m_j. If it meets the following two conditions, m_i and m_j are considered to be adjacent mode:

(1)the two modes are not the Planar or DC mode.

(2) |mi-mj|<=2

If these two conditions are not satisfied, mi and mj are not the adjacent mode. According to these definitions, we propose the following early termination strategy:

(1) Using set S to save the modes which have already implemented RDO process. Initialize $s = \{m_1, m_2\}$ and implement the full RDO process in this mode.

(2) For the rest of the N-2 modes, such as m_i, if it is the adjacent mode for any mode in S, and $RD\cos t_{m_{(i)}} > \alpha \cdot RD\cos t_{m_{(1)}}$, the RDO process of this mode will skip. Otherwise, encode it using RDO process, and add it to S. In here, the m_1 is the corresponding mode of the minimum Hadamard cost. Parameter α is set to 1.05 in this paper.

3.4. Integrated Fast Intra Prediction Algorithm

In this section, we integrate the above three schemes into a full fast intra prediction method, as shown in Figure 4. Such integrated fast algorithm is implemented into the HEVC reference software HM13.0 to demonstrate the performance.

```
//Step1. Early termination of CU split
for k = 0, 1, 2, 3
  Th=100 //The threshold value sets 100
  for i = 0, 1,........, N-1
    for j = 0, 1,........, N-1
        sum += I(i, j)
    end
  end
  ave = sum / (N*N) //The average of coding
  block pixels
  for i = 0, 1,........, N-1
    for j = 0, 1,........, N-1
        Var = (ave - I(i, j)) * (ave - I(i, j)) / (N*N)
         //Calculate the variance of image block
    end
  end
  if ( var > Th )
    skip CU level k  //skip to the next CU
    jump to CU level k+1 //the depth plus 1
  else
    break;
  end
end
```

```
//Step2. Early termination of mode
decision
for m is the first candidate mode in Ψ
  if any mode of MPM is the same as m
    m is the best mode
  else
        jump to 3
  end
end
```

```
//Step3. Early RDO termination
S = {m1, m2} //the initial setting S
δ = 1.05
for m in Ψ-S
  if m is adjacent with any mode in S
    adj(m, S) = True
  else
    adj(m, S) = False
  end
  if adj(m, S) && HC(m) >δ * HC(m1)
  // HC(m1) is the minimum of the
  Hadamard cost
  skip mode m
  else
    perform RDO for m
    update S = S + {m}
  end
end
```

Figure 4. The Fast Algorithm using the Integrated Schemes

4. Experimental Results

In order to verify the performance of the proposed fast algorithm, it is implemented on the latest HEVC reference software HM13.0 [12]. The hardware configuration of experimental platform is Pentium (R) Dual-Core CPU with frequency 2.10GHZ and 2GB of memory; the operating system is Windows 7 32-bit operating system; development tool is Microsoft Visual Studio 2010. Experiments are carried out for all I-frames sequences, and CABAC is used as the entropy coder. In this experiment, we will use 100 frames of each video sequence that recommended by JCT-VC in four resolutions which are Class A (4K×2K), B (1080p), C (WVGA), D (QWVGA), and E (720p) to test the performance of proposed algorithm, and every frame is intra coded. The proposed algorithm is evaluated with QPs 22, 27, 32 and 37 [13]. In the above-mentioned test conditions, the proposed algorithm in this paper is compared with HM13.0 and the fast algorithm proposed in literature [7] in the three aspects consisted of the encoding time, PSNR, and bitrate [14]. The experimental results are shown in Table 2, Table 3.

As can be seen from Table 2, the performance of the proposed fast algorithm is compared with the original algorithm HM13.0. The proposed algorithm can reduce encoding time by about 38%, and the bit rate increases slightly only about 0.72% than

HM13.0. On the other hand, coding efficiency loss is negligible in Table 2. Therefore, the proposed algorithm can reduce the coding time, and effectively improve the coding efficiency while maintaining the video quality unchanged as HM13.0.

Table 2. Results of the Proposed Algorithm Compared to HM13.0 with QPs at 32

Class & Video Sequences		BDPSNR(dB)	BDBR(%)	ΔT(%)
Class A (2560×1600)	NebutaFestival	-0.08	0.96	-43.26
	SteamLocomotiveTrain	-0.10	0.88	-42.89
	PeopleOnStreet	-0.10	0.85	-43.45
	zTraffic	-0.09	0.91	-42.96
Class B (1920×1080)	ParkScene	-0.08	0.78	-40.21
	BasketballDrive	-0.10	0.72	-41.02
	BQTerrace	-0.08	0.69	-39.88
	Cactus	-0.08	0.81	-39.91
	Kimono1	-0.08	0.75	-40.62
Class C (832×480)	BasketballDrill	-0.08	0.92	-36.78
	BQMall	-0.07	0.84	-36.85
	PartyScene	-0.06	0.71	-36.43
	RaceHorses	-0.08	0.58	-35.38
Class D (416×240)	BasketballPass	-0.06	0.55	-34.14
	BQSquare	-0.06	0.59	-34.57
	BlowingBubbles	-0.08	0.56	-34.12
	RaceHorses	-0.10	0.62	-34.28
Class E (1280×720)	FourPeople	-0.08	0.54	-38.36
	Johnny	-0.09	0.60	-37.92
	KristenAndSara	-0.08	0.62	-38.19
	vidyo1	-0.08	0.62	-38.14
	vidyo3	-0.08	0.68	-37.96
	vidyo4	-0.09	0.72	-38.21
Average		-0.08	0.72	-38.50

Table 3 shows performances of proposed fast prediction algorithm compared to the fast algorithm in the literature [7]. The proposed algorithm can reduce coding time, on average, about 22.8% with the maximum of 26% in "PeopleOnStreet (2560×1600)" and the minimum of 20% in "BasketballPass (416×240)". At the same time, the proposed fast prediction algorithm achieves better RD performance, where the average coding efficiency loss in term of PSNR is only 0.04dB and in term of bit rate is about 0.34% compared with the reference algorithm in the [7].

Table 3. Results of the Proposed Algorithm Compared to Literature [7] with QPs at 32

Class & Video Sequences		BDPSNR(dB)	BDBR(%)	ΔT(%)
Class A	NebutaFestival	-0.04	0.45	-26.12

(2560×1600)	SteamLocomotiveTrain	-0.04	0.49	-26.04
	PeopleOnStreet	-0.04	0.46	-26.21
	zTraffic	-0.05	0.51	-25.83
Class B (1920×1080)	ParkScene	-0.05	0.38	-24.37
	BasketballDrive	-0.04	0.34	-23.96
	BQTerrace	-0.02	0.31	-24.49
	Cactus	-0.02	0.43	-24.65
	Kimono1	-0.02	0.39	-24.78
Class C (832×480)	BasketballDrill	-0.03	0.52	-22.59
	BQMall	-0.04	0.47	-22.68
	PartyScene	-0.03	0.35	-23.10
	RaceHorses	-0.03	0.26	-21.21
Class D (416×240)	BasketballPass	-0.05	0.14	-20.02
	BQSquare	-0.05	0.22	-20.16
	BlowingBubbles	-0.04	0.21	-20.35
	RaceHorses	-0.04	0.26	-20.22
Class E (1280×720)	FourPeople	-0.05	0.19	-21.03
	Johnny	-0.05	0.23	-23.01
	KristenAndSara	-0.04	0.25	-22.65
	vidyo1	-0.03	0.25	-20.17
	vidyo3	-0.03	0.28	-21.03
	vidyo4	-0.03	0.36	-21.11
Average		-0.04	0.34	-22.86

Figure 5 shows the rate distortion curve of the four test sequences including "PeopleOnStreet", "BQTerrace", "PartyScene" and "BasketballPass" with three algorithms. We can observe that, compared with the original algorithm HM13.0 and the fast algorithm proposed in literature [7], our proposed algorithm has almost the same coding efficiency with them. At the same time it can also reduce the computational complexity, save the coding time, and improve the encoding speed.

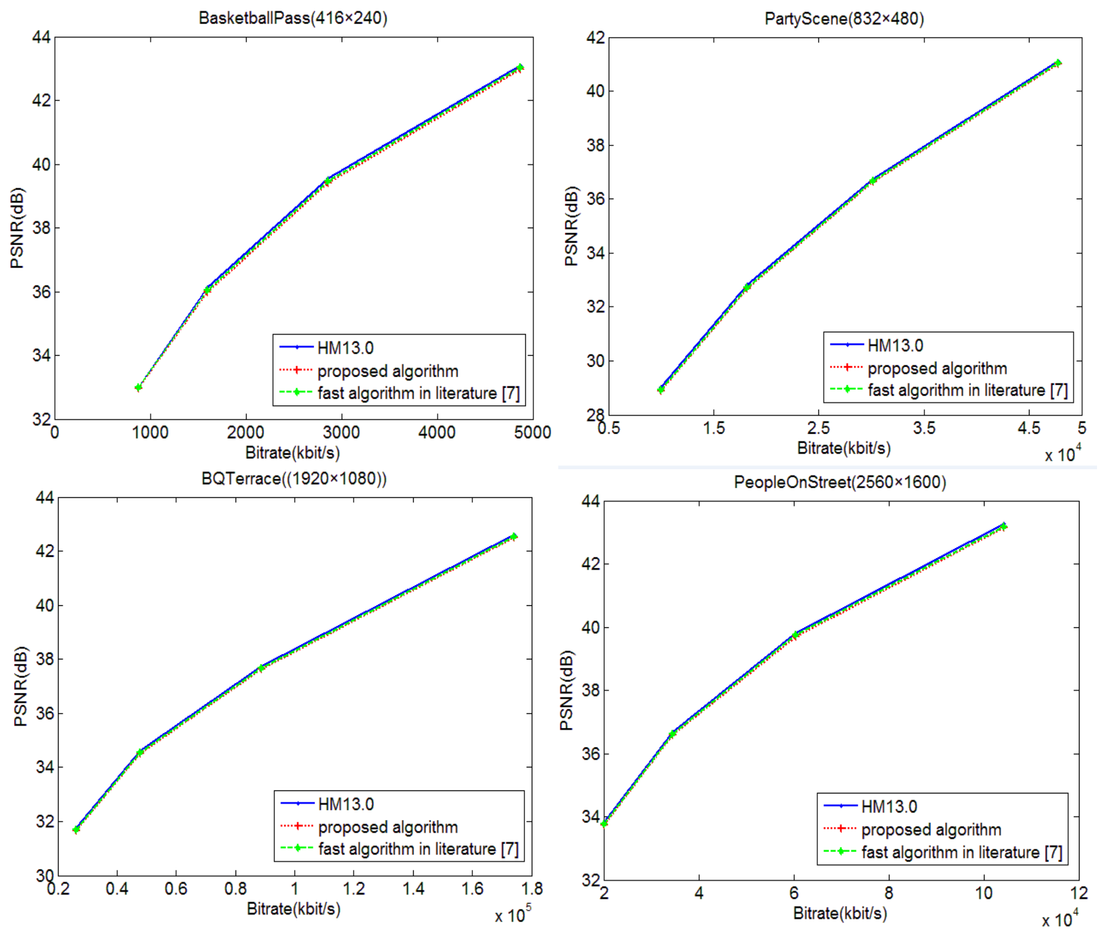


Figure 5. RD Curves of Four Test Sequences ("BasketballPass", "PartyScene", "BQTerrace", and "PeopleOnStreet")

5. Conclusions

This paper presents a fast algorithm with early termination CU splitting and mode decision. It uses the variance of the image block to determine the texture direction of the current image block, and early terminates mode decision process based on the correlation between the first candidate mode from RMD process and MPM from the spatially adjacent CU, and thereby reduces the amount of calculation and decreases encoding time. Experimental results show that, the proposed algorithm achieves considerable encoding time reduction with negligible bitrate increase and PSNR loss. Meanwhile, compared with the fast algorithm proposed in the [7], this algorithm has achieved good results.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China under grant NO. 61302118.

References

[1]. G. J. Sullivan, J. Ohm, Woo-Jin Han and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard”, in IEEE Transactions on Circuits and System for Video Technology, vol. 22, no. 12, **(2012)** December.

[2]. J. Lainema, F. Bossen, W-J Han, J. Min and K. Ugur, "Intra Coding of the HEVC Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, **(2012)** December, pp. 1792-1801.
[3]. C.-M. Fang, Y.-T. Chang and W.-H. Chung, "Fast Intra Mode Decision for HEVC Based on Direction Energy Distribution", in IEEE 17th International Symposium on Consumer Electronics (ISCE), **(2013)** June.
[4]. J. Jiang and B. Guo, "Fast intra coding algorithm using smooth region detection for HEVC", Journal of Xi Dian University, vol. 40, no. 3, **(2013)** June, pp. 194-199.
[5]. L. Shen, Z. Zhang and P. An, "Fast CU Size Decision and Mode Decision Algorithm for HEVC Intra Coding", in IEEE Transactions on Consumer Electronics, vol. 59, no. 1, **(2013)** February.
[6]. H. Zhang and Z. Ma, "Early termination schemes for fast intra mode decision in High Efficiency Video Coding", in IEEE International Symposium on Circuits and Systems (ISCAS), **(2013)** May.
[7]. T. Nishikori, T. Nakamura, T. Yoshitome and K. Mishiba, "A fast CU decision using image variance in HEVC intra coding", in IEEE Symposium on Industrial Electronics and Applications (ISIEA), **(2013)** September.
[8]. "HEVC Reference Model", [Online] Available: https://hevc.hhi.fraunhofer.de/trac/hevc /browser/tags.
[9]. B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan and T. Wiegand, "Working Draft 6 of High Efficiency Video Coding", document JCTVC-H1003, JCTVC of ISO/IEC and ITU-T, San José, CA, USA, **(2012)** February.
[10]. Y. Piao, J. Min and J. Chen, "Encoder improvement of unified intra prediction", JCTVC-C207, Guangzhou, **(2010)** October.
[11]. L. Zhao, L. Zhang, S. Ma and D. Zhao, "Fast mode decision algorithm for intra prediction in hevc", in The Visual Communications and Image Processing (VCIP) Conference, **(2011)** November.
[12]. A. S. Motra, A. Gupta, M. Shukla, P. Bansal and V. Bansal, "Fast intra mode decision for HEVC video encoder", in 20th International Conference on Software Telecommunication and Computer Networks, **(2012)** September.
[13]. F. Bossen, "Common test conditions and software reference configurations", Joint Collaborative Team on Video Coding (JCTVC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Document: JCTVC-B300, 2nd Meeting: Geneva, CH, **(2010)** July 21-28.
[14]. G. Bjontegaard, "Calculation of average PSNR difference between RD-curves", 13th VCEG-M33 Meeting, Austin, TX, **(2001)** April 2-4.

Authors

Yong Gan received the bachelor's degree in Semiconductor physics and device from Xi'an Jiao Tong University in 1986. He received Ph.D. Degree in computer science and technology from Xi'an Jiao Tong University in 2006.He is currently a Professor of computer science and technology in the department of computer and communication engineering, Zhengzhou University of Light Industry. His research interests include multimedia communications; image processing, coding; network engineering.

Xiaorong Zhao received the bachelor's degree in communication engineering from Zhengzhou University of Light Industry in 2009. She is currently a Master Degree Candidate of signal and information processing in the department of computer and communication engineering, Zhengzhou University of Light Industry. Her research interests include next-generation video coding standardization.

Qiuwen Zhang received the B.S. degree in Electrical and Computer Engineering from North China University of Water Conservancy and Electric Power, zhengzhou, China, in 2005, and the M.S. degree in Computer Engineering from Zhengzhou University of Technology, zhengzhou, China, in 2009, and the Ph.D. degree in communication and information systems from Shanghai University, Shanghai, China, in 2012. Since 2012, he has been with the faculty of the College of Computer and Communication Engineering, Zhengzhou University of Light Industry, where he is currently an Associate Professor. His major research interests include 3D video processing, multiview video and depth map coding, High Efficiency Video Coding (HEVC), video codec optimization, and multimedia communication.