

## Low Complexity H.265/HEVC Coding Unit Size Decision for a Videoconferencing System

Zhaoqing Pan<sup>1,2,3</sup>, Liming Chen<sup>1,2</sup>, Wenjie Yan<sup>3</sup>

<sup>1</sup>School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

<sup>2</sup>Jiangsu Engineering Centre of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

<sup>3</sup>School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300401, Tianjin, China

Emails: zhaqingpan@nuist.edu.cn leegend\_chen@outlook.com wenjiyan@scse.hebut.edu.cn

**Abstract:** With the development of Internet and video capture techniques, videoconferencing becomes still more popular. However, with the increased size of the video resolution, the volume of the raw videoconferencing video data increases dramatically. The H.265/HEVC video compression standard can efficiently compresses the videoconferencing video, while the high compression rate is at the cost of heavy encoding complexity. Hence, reducing the complexity is vital to be used for the H.265/HEVC encoder in videoconferencing systems. In this paper we proposed a low complexity H.265/HEVC coding unit size decision algorithm for a videoconferencing system, by considering the coding content property of the coding unit, in which the coding unit size decision process is terminated if the prediction residuals of the coding unit are all transformed and quantized to zeroes. The experimental results show that the proposed algorithm achieves 62.79% on average in encoding time saving, while the rate distortion degradation is acceptable.

**Keywords:** Videoconferencing system, H.265/HEVC, Coding unit, quadtree, low complexity.

### 1. Introduction

VideoConferencing (VC) is the conduct of a videoconference by a group of telecommunication techniques which allows more than one locations to communicate by simultaneous two-way video and audio transmissions. With the increased size of the video resolution, the amount of raw video data needs to be

compressed and transmitted. However, the current computing and transmission capabilities are still limited. As a result, the high compression rate and low complexity are the key requirements for high definition videos to be widely used in VC. The H.265/HEVC (High Efficiency Video Coding) standard [1-3] can efficiently address the compression rate problem. However, the achieved compression rate is at the cost of high computational complexity of a group of advanced coding tools used in the H.265/HEVC, such as the quadtree structure based Coding Tree Unit (CTU). Therefore, reducing the computational complexity is essential for the H.265/HEVC encoder to be applied in the VC system.

To compress the videoconferencing video, the video is separated into a number of frames, and the frames are sequentially encoded. Furthermore, the frame is partitioned into slices, the slices are split into a group of CTUs, which are the basic compression unit of the H.265/HEVC encoder. According to the quadtree structure, the CTUs are further partitioned into Coding Units (CUs), and the HEVC supports flexible CU sizes from the size of  $64 \times 64$  to the size of  $8 \times 8$ , which are corresponding to the CU quadtree depth of 0 up to depth 3. In CU encoding process, the HEVC encoder checks  $4^i$ ,  $i \in [0, 1, 2, 3]$ , CU partitions for each quadtree depth level  $n$ , there are totally 85 CU partitions. Fig. 1 gives an example of the H.265/HEVC quadtree based CU partition process. At last, the best CU quadtree depth level  $d^*$  is determined according to the rate distortion cost function [4, 5], as shown:

$$(1) \quad \begin{cases} d^* = \arg \min_{d \in D} J_d, \\ J_d = D(d) + \lambda \bullet R(d), \end{cases}$$

where  $D$  is the candidate CU quadtree depth level,  $D(d)$  is the distortion between the original CU and its reconstructed CU which is obtained by encoding the original CU with a quadtree depth level  $d$ ;  $\lambda$  is the Lagrangian multiplier;  $R(d)$  represents the number of total bits for encoding the CU with depth  $d$ . This “try all and select the best” CU quadtree depth decision method consumes high computation, which could limit the H.265/HEVC encoder to be used in real time VC systems.

To reduce the computational complexity, many researchers have devoted their efforts to reducing the computational complexity of H.265/HEVC CU size decision process. Choi, Park and Jang [6] proposed a coding tree pruning based CU early termination method for a low complexity H.265/HEVC encoder, in which the CU size decision process is terminated, if the current CU selects the Skip mode as its best prediction mode. By using the spatial and temporal CU depth selection correlation, Pan et al. [7] have proposed a fast CU size decision method for H.265/HEVC encoder. Based on the texture complexity, Hou et al. [8] proposed a fast CU size decision algorithm for H.265/HEVC intra coding. Ahn, Lee and Kim [9] proposed a fast CU encoding method by using spatiotemporal encoding parameters for H.265/HEVC inter coding. Based on statistical analysis, Lee et al. [10] proposed a fast CU size decision algorithm by considering the Skip mode decision, CU skip estimation and early CU termination. Based on the depth information correlation between the spatial-temporal adjacent coding tree units and the current coding tree unit, Zhang, Wang and Li [11] proposed a fast coding

unit depth decision algorithm for interframe coding in H.265/HEVC. Mu et al. [12] proposed a fast CU depth decision method by exploiting the relationship between the rate distortion cost and CU depth. Huang et al. [13] proposed an adaptive depth search range for HEVC coding unit size selection according to the coded block flag and the information of the quantization parameter. However, these methods cannot efficiently remove the complexity of H.265/HEVC encoder of a videoconferencing system due to the fact that video content properties are not considered by these methods (Fig. 1).

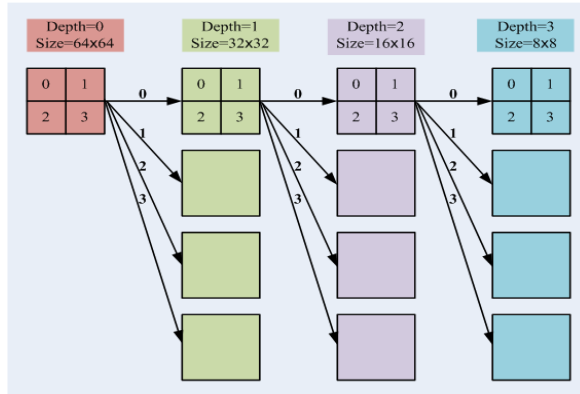


Fig. 1. An example of the H.265/HEVC quadtree based CU partition process

In this paper we propose a fast CU size decision method for the low complexity H.265/HEVC encoder of a videoconferencing system by considering the video content property. The rest of this paper is organized as follows. The motivations and statistical analyses are presented in Section 2. Then the details of the proposed fast CU size decision algorithm are given in Section 3. Experimental results are given in Section 4. Finally, Section 5 concludes the paper.

## 2. Motivations and statistical analyses

To analyze the encoding complexity of the best CU quadtree depth level decision process, three different HD video sequences with a resolution 1280×720, including “FourPeople”, “Johnny”, and “Vidyo1” are tested; detailed information of these three sequence is given in Table 1. The HEVC reference software HM12.0 is used as a software platform. The test conditions are listed as follows: the maximum CU size is 64×64, and the maximum CU quadtree depth is 4; the fast motion estimation method is TZSearch; four Quantization Parameters (QPs) are adopted. The statistical results are given in Table 2, where the CU encoding complexity,  $C_{\text{complexity}}$  is computed by

$$(2) \quad C_{\text{complexity}} = T_c / T_e,$$

where  $T_c$  represents the total encoding time of the CU size decision process,  $T_e$  indicates the total encoding time of the H.265/HEVC encoder.

Table 1. Test sequences

Sequence	Size	Frame rate, fps	Coding structure	Frames encoded
FourPeople	1280×720	60	Low-Delay-Main	65
Johnny				
Vidyo1				

Table 2. Statistical results of CU encoding complexity

Sequence	QP=22	QP=27	QP=32	QP=37
FourPeople	0.9980	0.9984	0.9985	0.9986
Johnny	0.9981	0.9983	0.9985	0.9988
Vidyo1	0.9982	0.9984	0.9986	0.9987

From Table 1 it can be seen that the encoding complexity of the CU size decision process is rather large, the value of  $C_{\text{complexity}}$  is from 0.9980 up to 0.9988. From these values, we can draw a conclusion that the “try all and select the best” CU size decision process consumes most of the total encoding time. Thus, if the CU size decision process could be optimized, much more encoding time could be saved.

### 3. The proposed fast CU size decision algorithm for a low complexity H.265/HEVC encoder

#### 3.1. The proposed all-zero block based fast CU size decision algorithm

The videoconferencing video has many homogenous regions, such as the background, since a big size CU could encode the prediction residual in a few number of symbols than possible under the condition of many small size CUs; the homogenous regions are rather suitable for encoding in a big size CU partition. In addition, the interprediction residuals of these homogenous regions are easy transformed and quantized into zeroes. Hence, designing an early termination for the CU size decision process based on the quantization coefficients is practicable. In video coding, the All-Zero Block (AZB) means that the prediction residuals of this block are all transformed and quantized into zeroes. To find the correlation between the AZB and the optimal CU quadtree depth, the event  $\mathbf{M}$  indicates after encoding the CUs of depth level  $i$  with the merge and inter  $2N \times 2N$  PU mode, they are all AZBs. The event  $\mathbf{N}$  represents that the CU quadtree depth  $i$  is selected as the optimal CU quadtree depth level, the conditional probabilities of  $P(\mathbf{N}|\mathbf{M})$  are tested. The statistical results are shown in Table 3.

Table 3. The conditional probabilities of  $P(\mathbf{N}|\mathbf{M})$ 

Sequence	QP=22	QP=27	QP=32	QP=37	Average
FourPeople	80.10%	86.97%	89.79%	91.82%	87.17%
Johnny	80.56%	86.01%	89.20%	91.69%	86.87%
Vidyo1	78.34%	85.70%	88.66%	90.70%	85.85%
Average	79.67%	86.23%	89.22%	91.40%	86.63%

From Table 3 it can be observed that the CU has a quite large probability to choose the quadtree depth level  $i$  as its best CU quadtree depth level, when the CUs with depth  $i$  are all AZBs. The conditional probability of  $P(\mathbf{N}|\mathbf{M})$  holds an average of 86.63%, and is from 78.34% up to 91.82%. It can also be seen that the conditional probability of  $P(\mathbf{N}|\mathbf{M})$  increases as the QP value gets bigger, this is because more inter prediction residuals are transformed and quantized to zeroes due to the fact that the coding content becomes simple as the QP increases. Therefore, according to the above analyses, the best CU quadtree depth level selection process is completed if

$$(3) \quad \sum_n \text{QM}_n + \sum_n \text{QI}_n = 0,$$

where  $n$  indicates the number of total CU partitions,  $n = 4^i$ ,  $i \in [0, 1, 2, 3]$ ,  $i$  is the CU quadtree depth level. After encoding the CU with the merge mode, if its prediction residuals are transformed and quantized into zeroes, then the value of  $\text{QM}_n$  is equal to 0, otherwise, the value of  $\text{QM}_n$  is bigger than 0; after encoding the CU with the inter  $2N \times 2N$  mode, if its prediction residuals are transformed and quantized into zeroes, then the value of  $\text{QI}_n$  is equal to 0; otherwise, the value of  $\text{QI}_n$  is bigger than 0.

### 3.2. The overall algorithm

According to the above analysis, the low complexity H.265/HEVC CU size decision algorithm for videoconferencing video coding is summarized in Algorithm 1.

## 4. Experimental results

In order to evaluate the efficiency of the low complexity CU size decision algorithm, the HEVC reference software HM12.0 is used as a software platform. The hardware platform is Intel Xeon CPU E3-1241 v3 @ 3.50 GHz and 3.50 GHz, 4.00 GB RAM with Microsoft Windows 7 64-bit operating system. The test conditions are indicated in Table 4. Six videoconferencing video sequences, including “FourPeople”, “Johnny”, “KristenAndSara”, “Vidyo1”, “Vidyo3”, “Vidyo4” are adopted. These six video sequences are shown in Fig. 2. The information of these six sequences is that the resolution is  $1280 \times 720$ , the frame rate is 60 fps, and the number of encoded frames is 97.

#### **Algorithm 1. The low complexity H.265/HEVC CU size decision algorithm**

**Step 1.** *Input:* The maximum CU size:  $64 \times 64$ , the maximum quadtree depth: 4

**Step 2.** *for* Quadtree depth level=0 to 3 *do*

**Step 3.** Encode the current CU with the Merge mode

**Step 4.** Encode the current CU with the Inter  $2N \times 2N$  mode

**Step 5.** *if*  $\sum_n \text{QM}_n + \sum_n \text{QI}_n = 0$  *Then*

- Step 6.** The CU size decision process is terminated  
**Step 7.** *else*  
**Step 8.** Encode the current CTU with the next quadtree depth level  
**Step 9.** *end if*  
**Step 10.** *Output:* The best CU quadtree depth level  
**Step 11.** *end for*  
**Step 12.** Process the next coding tree unit

Table 4. Test conditions

The maximum CU size	64×64
The maximum quadtree depth	4
Motion estimation method	TZSearch
Motion estimation search range	±64
Quantization parameters	22, 27, 32, 37

We compared the encoding performance of the proposed algorithm with Choi, Park and Jang [6] in terms of the Peak Signal to Noise Ratio (PSNR), Bit Rate (BR) and total encoding time saving. The experimental results are summarized in Table 5. In the table  $\Delta\text{PSNR}$ ,  $\Delta\text{BR}$  and  $\Delta T$  denote the PSNR decrease, BR increase and the total encoding CPU time saving [14, 15], and they are computed as given below:

$$(4) \quad \begin{cases} \Delta\text{PSNR} = \text{PSNR}_{\phi} - \text{PSNR}_{\text{HM}}, \\ \Delta\text{BR} = \frac{\text{BR}_{\phi} - \text{BR}_{\text{HM}}}{\text{BR}_{\text{HM}}}, \\ \Delta T = \frac{T_{\phi} - T_{\text{HM}}}{T_{\text{HM}}}, \end{cases}$$

where  $\text{PSNR}_{\phi}$ ,  $\text{BR}_{\phi}$ ,  $T_{\phi}$  present the PSNR, BR and total encoding time of the algorithm  $\phi$ ,  $\phi \in \{\text{Choi [6], Proposed}\}$ ;  $\text{PSNR}_{\text{HM}}$ ,  $\text{BR}_{\text{HM}}$ ,  $T_{\text{HM}}$  denote the PSNR, BR and total encoding time of the original HM12.0, respectively.

From Table 5 we can see that Choi's method reduces the total encoding time from 32.54 up to 70.17%, 55.47% on average. Meanwhile, the PSNR decreases from 0.024 up to 0.071 dB, 0.038 dB on average; and the BR decreases from 0.38 up to 1.99%, 1.04% on average. The proposed method saves the total encoding time 41.71 up to 76.36%, 62.79% on average. At the same time, the PSNR degrades from 0.027 up to 0.110 dB, 0.055 dB on average, and the BR decreases from 0.30 up to 2.12%, 1.07% on average. From these values we can see that the proposed method achieves a quite similar rate distortion performance with Choi's method, while it obtains more encoding time saving than Choi's method. Thus, we can conclude that the proposed method efficiently removes the encoding complexity of the H.265/HEVC encoder, meanwhile maintaining a comparable rate distortion performance.

Table 5. Summary of the encoding results

Sequence	QP	Choi [6]			Proposed		
		$\Delta$ PSNR, dB	$\Delta$ BR, %	$\Delta T$ , %	$\Delta$ PSNR, dB	$\Delta$ BR, %	$\Delta T$ , %
FourPeople	22	-0.030	-1.32	-41.84	-0.034	-1.27	-46.45
	27	-0.039	-1.28	-55.15	-0.048	-1.35	-59.15
	32	-0.025	-0.65	-62.45	-0.037	-0.61	-65.89
	37	-0.029	-0.62	-67.37	-0.056	-0.48	-69.87
	Average	-0.031	-0.97	-56.70	-0.044	-0.93	-60.34
Johnny	22	-0.024	-1.22	-41.67	-0.027	-1.13	-46.64
	27	-0.030	-1.16	-56.50	-0.041	-1.48	-61.95
	32	-0.046	-1.21	-64.89	-0.051	-1.08	-69.19
	37	-0.039	-1.30	-70.17	-0.065	-1.19	-73.16
	Average	-0.035	-1.22	-58.31	-0.046	-1.22	-62.74
KristenAndSara	22	-0.035	-1.12	-41.23	-0.049	-1.30	-48.47
	27	-0.047	-1.44	-54.02	-0.064	-2.12	-64.85
	32	-0.050	-1.99	-62.69	-0.084	-2.02	-71.64
	37	-0.053	-1.27	-67.84	-0.071	-1.07	-76.02
	Average	-0.046	-1.46	-56.45	-0.067	-1.63	-65.25
Vidyo1	22	-0.034	-1.40	-43.63	-0.044	-1.51	-51.81
	27	-0.042	-1.28	-55.67	-0.045	-1.35	-65.25
	32	-0.043	-1.06	-62.99	-0.063	-0.94	-72.02
	37	-0.036	-1.40	-68.12	-0.064	-1.18	-76.36
	Average	-0.038	-1.29	-57.60	-0.054	-1.25	-66.36
Vidyo3	22	-0.038	-0.80	-40.47	-0.044	-1.16	-50.23
	27	-0.045	-0.69	-51.98	-0.067	-0.94	-62.22
	32	-0.059	-0.84	-59.36	-0.096	-0.72	-69.08
	37	-0.071	-0.75	-65.03	-0.110	-0.44	-74.24
	Average	-0.053	-0.77	-54.21	-0.080	-0.82	-63.94
Vidyo4	22	-0.029	-0.69	-32.54	-0.034	-0.64	-41.71
	27	-0.022	-0.72	-46.57	-0.035	-0.88	-57.01
	32	-0.026	-0.38	-56.17	-0.044	-0.30	-62.19
	37	-0.028	-0.40	-62.89	-0.043	-0.57	-71.48
	Average	-0.026	-0.55	-49.54	-0.039	-0.60	-58.10
Average		<b>-0.038</b>	<b>-1.04</b>	<b>-55.47</b>	<b>-0.055</b>	<b>-1.07</b>	<b>-62.79</b>

Fig. 2 shows the test sequences.

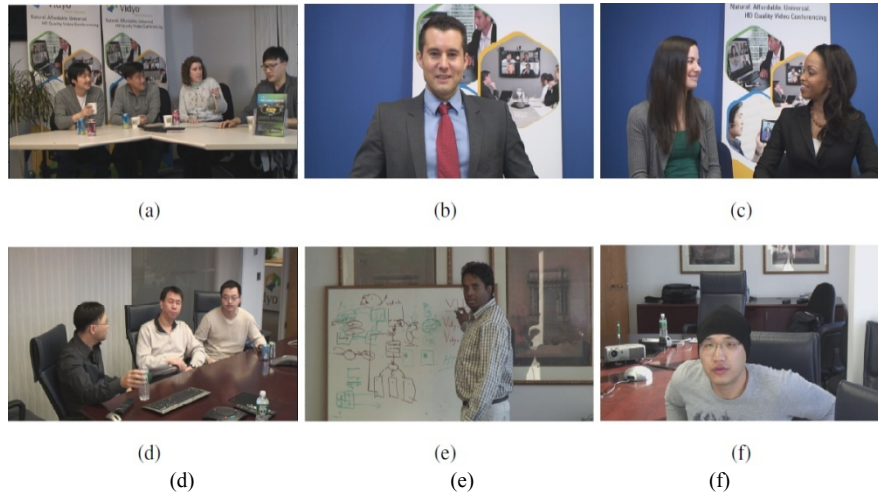


Fig. 2. Test sequences: FourPeople (a); Johnny (b); KristenAndSara (c); Vidyo1 (d); Vidyo3 (e); Vidyo4 (f)

## 5. Conclusion

The CU size decision process consumes the greatest proportion of the total encoding time of H.265/HEVC encoder, so in order to make the H.265/HEVC encoder meet the complexity requirements of a videoconferencing system we proposed a low complexity 265/HEVC CU size decision algorithm. The experimental results show that the proposed algorithm can efficiently reduce the computational complexity of the H.265/HEVC encoder, while the rate distortion performance degradation is acceptable.

**Acknowledgements:** This work was supported in part by the National Natural Science Foundation of China under Grant 61501246, in part by the National Natural Science Foundation of Jiangsu Province of China under Grant BK20150930, in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 15KJB510019, in part by the Project through the Priority Academic Program Development of Jiangsu Higher Education Institutions, in part by the Startup Foundation for Introducing Talent of NUIST under Grant 2243141501012, in part by the Natural Science Foundation of Hebei Province of China under Grant F2015202311 and Grant F2015202214.

## References

1. Sullivan, G. J., J.-R. Ohm, W.-J. Han, T. Wiegand. Overview of the High Efficiency Video Coding (HEVC) Standard. – IEEE Trans. Circuits Syst. Video Technol., Vol. 12, 2012, No 12, pp. 1649-1668.
2. Information Technology-High Efficiency Coding and Media Delivery in Heterogeneous Environments-Part 2: High Efficiency Video Coding. ISO/IEC 23008-2:2013 ITU-T Rec. H.265, 2013.



3. Pan, Z., S. Kwong, M.-T. Sun, J. Lei. Early MERGE Mode Decision Based on Motion Estimation and Hierarchical Depth Correlation for HEVC. – IEEE Transactions on Broadcasting, Vol. **60**, 2014, No 2, pp. 405-412.
4. Ohm, J.-R., G. J. Sullivan, H. Schwarz, T. K. Tan, T. Wiegand. Comparison of the Coding Efficiency of Video Coding Standards-Including High Efficiency Video Coding (HEVC). – IEEE Trans. Circuits Syst. Video Technol., Vol. **22**, 2012, No 12, pp. 1669-1684.
5. Pan, Z., Y. Zhang, S. Kwong. Efficient Motion and Disparity Estimation Optimization for Low Complexity Multiview Video Coding. – IEEE Transactions on Broadcasting, Vol. **61**, 2015, No 2, pp. 166-176.
6. Choi, K., S.-H. Park, E. S. Jang. Coding Tree Pruning Based CU Early Termination. ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) Document JCTVC-F092, July 2011.
7. Pan, Z., S. Kwong, Y. Zhang, J. Lei, H. Yuan. Fast Coding Tree Unit Depth Decision for High Efficiency Video Coding. – In Proc. of 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, May 2014, pp. 3214-3218.
8. Hou, J., D. Li, Z. Li, X. Jiang. Fast CU Size Decision Based on Texture Complexity for HEVC Intra Coding. – In Proc. of 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), Shenyang, China, December 2013, pp. 1096-1099.
9. Ahn, S., B. Lee, M. Kim. A Novel Fast CU Encoding Scheme Based on Spatiotemporal Encoding Parameters for HEVC Inter Coding. – IEEE Transactions on Circuits and Systems for Video Technology, Vol. **25**, 2015, No 3, pp. 422-435.
10. Lee, J., S. Kim, K. Lim, S. Lee. A Fast CU Size Decision Algorithm for HEVC. – IEEE Transactions on Circuits and Systems for Video Technology, Vol. **25**, 2015, No 3, pp. 411-421.
11. Zhang, Y., H. Wang, Z. Li. Fast Coding Unit Depth Decision Algorithm for Interframe Coding in HEVC. – In Proc. Data Compression Conference (DCC), Snowbird, UT, Mar., pp. 53-62.
12. Mu, F., L. Song, X. Yang, Z. Luo. Fast Coding Unit Depth Decision for HEVC. – In Proc. of 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Chengdu, China, July 2014, pp. 1-6.
13. Huang, X., C. Kuo, C. Mao, Y. Ciou. Adaptive Depth Search Range for HEVC Coding Unit Size Selection. – In Proc. of 2014 Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association (APSIPA), Siem Reap, December 2014, pp. 1-6.
14. Pan, Z., S. Kwong, L. Xu, Y. Zhang, T. Zhao. Predictive and Distribution Oriented Fast Motion Estimation for H.264/AVC. – Journal of Real-Time Image Processing, Vol. **9**, 2014, No 4, pp. 597-607.
15. Pan, Z., S. Kwong. A Direction-Based Unsymmetrical-Cross Multi-Hexagon-Grid Search Algorithm for H.264/AVC Motion Estimation. – Journal of Signal Processing Systems, Vol. **7**, 2013, No 1, pp. 59-72.