

RESEARCH

Open Access

Fast intermode decision algorithm based on general and local residual complexity in H.264/AVC

Jaeho Lee¹, Seongwan Kim¹, Kyungmin Lim¹, Jae Hyun Kim² and Sangyoun Lee^{1*}

Abstract

The state-of-the-art video coding standard H.264/AVC achieves significant coding performance by adopting variable block sizes for motion estimation (ME) and mode decision. However, this technique brings out high computational complexity since the optimal mode is determined by exhaustively performing rate-distortion optimization (RDO) on each coding mode with different block sizes. In this paper, the fast intermode decision algorithm is proposed to reduce the computational complexity. Based on the high correlation between the residual error of ME and the optimal block size, general residual complexity (GRC) and local residual complexity (LRC) are defined. According to MB activity evaluated on GRC and LRC, candidate intermodes are determined and RDO processes are only performed on selected intermodes. The experimental results demonstrate that the proposed algorithm achieves time saving by 63% on average with negligible degradation of coding efficiency.

Keywords: H.264/AVC, Rate-distortion optimization, Fast intermode decision, General residual complexity, Local residual complexity

1 Introduction

The state-of-the-art video coding standard H.264/AVC [1] was developed by the Joint Video Team of ITU-T Video Coding Experts Group and ISO/IEC Moving Picture Experts Group. Compared to previous the video coding standards, H.264/AVC achieves higher coding performance with the use of several advanced coding tools such as variable block sizes for motion estimation (ME) and mode decision, multi-reference frames, a motion vector (MV) with quarter-pixel accuracy, a deblocking filter, and context-based adaptive binary arithmetic coding (CABAC) [2]. However, computational complexity increases with these tools, especially by several block sizes for ME and mode decision [3].

In ME and mode decision, a macroblock (MB) can be divided into four block partitions: 16×16 , 16×8 , 8×16 , and 8×8 ; and each 8×8 subblock, denoted $P8 \times 8$, can be further divided into 8×8 , 8×4 , 4×8 , and 4×4 .

Figure 1 represents all block sizes and their relationships for ME and mode decision.

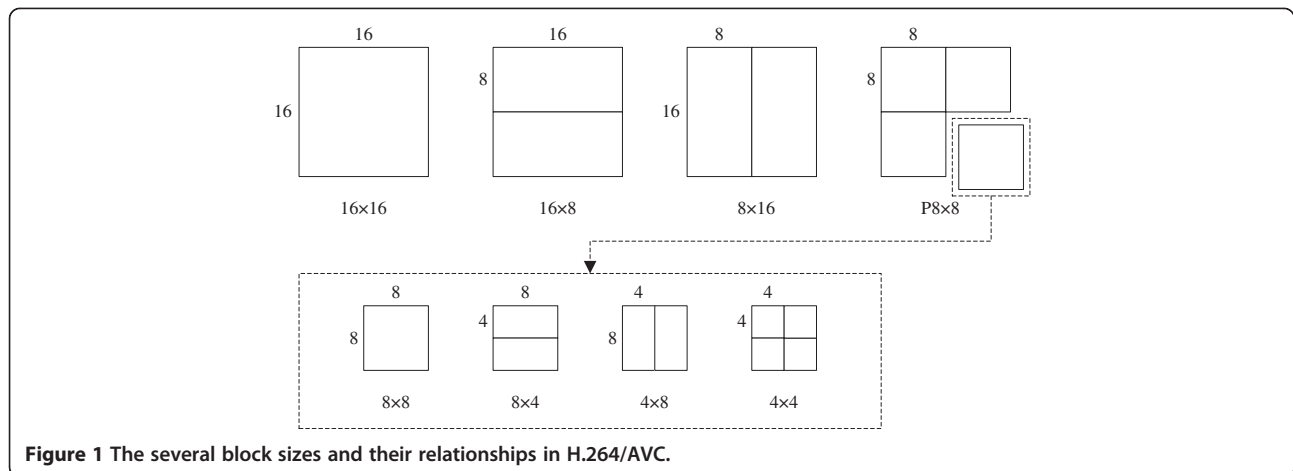
To obtain the optimal mode from candidate coding modes with various block sizes, H.264/AVC employs the rate-distortion optimization (RDO) technique [4]. In RDO, after the rate-distortion cost (RDcost) of each coding mode is calculated, the mode with the minimum RDcost is selected as the optimal mode. Since RDO requires encoding and decoding processes to obtain RDcost, it brings forth vast computational complexity. Therefore, it is highly aspired to reduce the encoder complexity without noticeable coding loss for the wide application of H.264/AVC, including video services (video conferencing, digital multimedia broadcasting, and IPTV), and consumer products (mobile devices and DVD or Blu-ray players).

A number of approaches [5-16] have been reported for H.264/AVC to reduce the encoder complexity. The main idea of algorithms is to eliminate unnecessary modes with different block sizes and to perform the RDO process for possible block sizes only. In some approaches [5-8], SKIP mode is focused in reducing the encoder complexity. The early SKIP mode decision

* Correspondence: syleee@yonsei.ac.kr

¹Department of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, South Korea

Full list of author information is available at the end of the article



algorithm [5,6] is proposed based on four conditions. If these conditions are verified, the SKIP mode is determined as the optimal mode, and the remaining modes are considered as unnecessary modes for the RDO process. This algorithm has been adopted as fast mode decision option in JM reference software. Additional conditions [7] for SKIP mode are proposed based on spatial and temporal neighborhood information to improve the performance of the early SKIP mode decision algorithm [6]. Saha et al. [8] employed the sum-of-absolute-transformed-differences to predict the SKIP mode. In these algorithms [5-8], since coding modes excluding the SKIP mode are considered as a single group, if the SKIP mode is not determined as the optimal mode, the RDO process is executed for all coding modes. Accordingly, the performance improvement is limited in sequences with fast motion or detailed regions. Therefore, additional approaches [9-16] for remaining coding modes are proposed to further reduce the encoding time. Wu et al. [9] proposed spatial homogeneity by using the Sobel operator and temporal stationary with the difference between current MB and its co-located MB in the reference frame. Based on homogeneity and stationary, only a small number of intermodes are selected for the RDO process. In [10], the bottom-up merge method is introduced for the fast intermode selection. In the bottom-up merge method, 16×16 block is split into 4×4 blocks, and then 4×4 blocks with the same class are merged based on MVs and the edge information. Ren et al. [11] proposed a fast adaptive early termination (FAT) mode selection algorithm. FAT consists of three steps: initial prediction, early termination, and refinement. In the initial prediction, the candidate modes are determined based on the mode histogram from neighboring MBs. Early termination and refinement for candidate modes are considered to trade-off computational efficiency and accuracy. The MB tracking scheme [12] is proposed based on the temporal

correlation between successive frames. The candidate modes of the current MB are determined according to the optimal mode of co-located MB in the previous frame. Zeng et al. [13] proposed a method to select the candidate modes according to motion activity, generated by the MVs of spatially and temporally nearby MBs. Liu et al. [14] proposed an efficient intermode decision algorithm based on motion homogeneity evaluated on a normalized MV field, as determined by the MVs from ME of 4×4 block. Three-directional motion homogeneities are exploited to select candidate intermodes. The classified region algorithm [15] is proposed based on the spatial and temporal homogeneity of the block obtained by using 16×16 and 8×8 block pattern. The candidate intermodes for RDO process are reduced by using the spatial and temporal homogeneity. Martínez-Enríquez et al. [16] proposed an adaptive algorithm based on RDcost statistics to decrease the encoding time. The differences between the RDcost for each mode are used to attain successive adaptive thresholds that allow the reduction of the number of evaluated intermodes. Although previous works [9-16] are well designed, there is still a need to further develop a more efficient algorithm, especially for video sequences with high motion or detailed regions.

In this paper, fast intermode decision algorithm is proposed based on global residual complexity (GRC) and local residual complexity (LRC). MB activity is determined based on GRC and LRC, and RDO processes are only performed for candidate intermodes chosen according to MB activity.

The features of the proposed algorithm are as follows: First, three subsets of candidate intermodes are defined to represent different MB activities. Candidate intermodes of selected subset are only considered for ME and mode decision. Second, to determine MB activity, GRC and LRC are proposed based on the observation that smaller block sizes are likely to be selected as the

optimal mode in MB with high residual error of ME. Adaptive threshold for LRC according to GRC is designed for available QP ranges of H.264/AVC. Third, additional computation is only needed to obtain GRC, which can be easily and inexpensively calculated and performed only once for each frame. Since LRC is obtained in the process of ME for 16×16 , additional computation is not needed. These features contribute to the performance improvement for all sequences including ones with high motion or detailed regions.

The rest of this paper is organized as follows. In Section 2, the proposed algorithm is introduced in detail. The simulation results are provided in Section 3. Finally, the conclusion is represented in Section 4.

2 The proposed intermode decision algorithm

2.1 Motivation

In ME and mode decision, variable block sizes can reduce the prediction or residual error efficiently. For example, if a MB includes homogeneous regions with no or slow motion, it is appropriate to select large block sizes such as 16×16 , since large block sizes can result in sufficiently small residual error. In contrast, large block sizes lead to large prediction error for a MB with fast motion or detailed regions. In this case, smaller block sizes such as 8×8 , 8×4 , 4×8 , and 4×4 can be considered as the optimal mode. Accordingly, it is likely to select smaller block sizes when a MB includes detailed regions with fast motion.

It is observed that there is a high correlation between the optimal block size and residual error. Figure 2 shows one example, in which the 77th frame from the CIF sequence Silent is used. Figure 2a represents the real frame image, and Figure 2b represents the absolute difference between each pixel in the current MB and the reference block of 16×16 ME. The optimal intermodes are represented by using different-sized boxes overlying on the corresponding MB, and the absolute difference value is represented in a darker shade when the absolute difference value is larger and vice versa. As seen in Figure 2a, larger block sizes are selected for the homogeneous region, such as the white wall in the background. Also, the painting of the background is coded for large blocks, since it remains still, although this region includes non-homogeneous patterns. On the other hand, regions including motion boundary such as the hair and the right hand of the woman are coded in smaller blocks. Since in these regions, the residual error of ME for large block is large, smaller blocks are considered to achieve small residual error. It can be seen in Figure 2b that when a small block size is selected as the optimal mode for a MB, the residual error of ME for 16×16 block is large (gray or

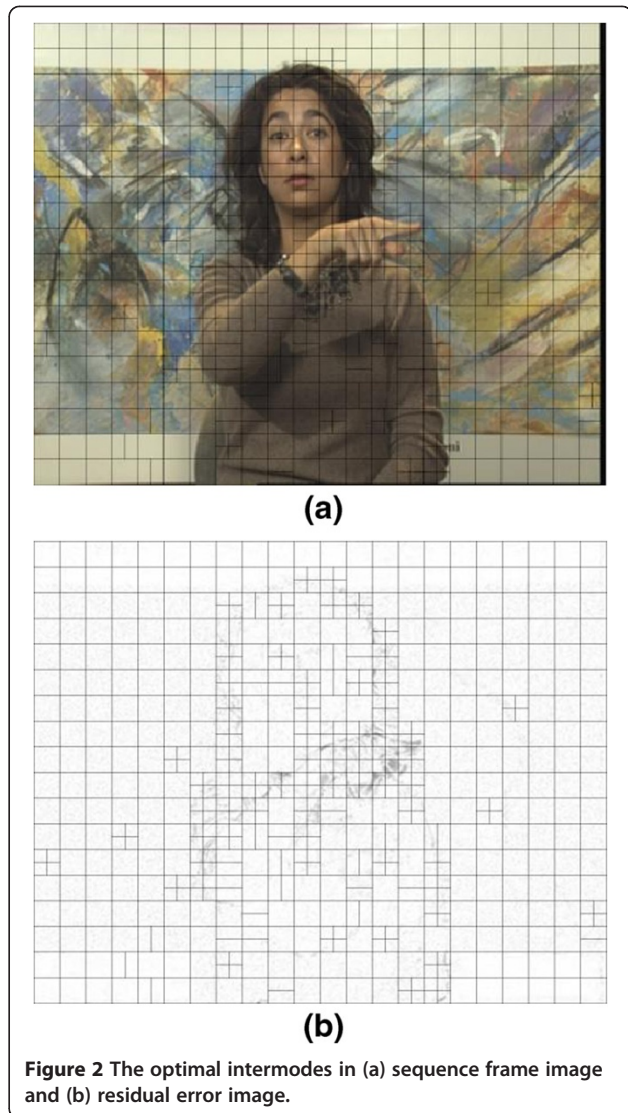


Figure 2 The optimal intermodes in (a) sequence frame image and (b) residual error image.

dark gray region), whereas if the residual error is small (white region), MBs are coded for a large block.

According to this observation, it is concluded that the optimal block size can be predicted based on the MB residual error. If the residual error is small when ME for large block is performed, a large block size can be used as the optimal mode. In contrast, if the residual error is not small enough to select a large block, smaller block sizes are considered to determine the optimal block size mode.

2.2 Intermode decision based on LRC and GRC

Based on the above analysis, LRC and MB activities are proposed to determine candidate intermodes. After ME is performed on the 16×16 block size, LRC is calculated as the sum of absolute difference between the current MB and the reference block of 16×16 ME, as follows:

Table 1 Classes and candidate intermodes according to MB activity

Class	MB activity	Candidate intermodes
0	Low	16×16
1	Medium	16×16 , 16×8 , 8×16
2	High	16×16 , 16×8 , 8×16 , $P8 \times 8$

$$LRC = \sum_{x=0}^{15} \sum_{y=0}^{15} |s(x, y) - r(x, y)|, \quad (1)$$

where s denotes the current MB, r indicates the reference block of 16×16 ME, and x and y are the horizontal and vertical positions of the MB. The MB activity is determined according to LRC as follows:

$$MB_{activity} = \begin{cases} \text{Low} & \text{if } LRC \leq L_0 \\ \text{Medium} & \text{if } L_0 < LRC \leq L_1 \\ \text{High} & \text{if } L_1 < LRC \end{cases} \quad (2)$$

Based on the selected MB activity, the candidate intermodes summarized in Table 1 are considered to determine the optimal mode. For example, if LRC is greater than L_0 and less than or equal to L_1 , MB activity is selected as medium, and, henceforth, RDO processes are performed only for 16×16 , 16×8 , and 8×16 . To design the adaptive thresholds L_0 and L_1 in (2), GRC is defined as in (3):

$$GRC = \frac{1}{N \times M} \times \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |S(s, y) - R(x, y)| + 0.5 \quad (3)$$

where $\lfloor x \rfloor$ is the largest integer value less than or equal to x , N and M are the frame height and width, respectively. S is the current frame and R is the previous frame in display order, which means the nearest candidate reference frame to the current frame. The thresholds L_i are obtained according to each GRC and QP as follows.

$$L_i = \alpha_i \times \mu_i \quad (4)$$

where $i \in \{0, 1\}$, α_i are defined to provide a trade-off between visual quality and time saving, and μ_i is the average of LRCs of MBs for class i from Table 1. For example, μ_0 is obtained as the average of LRCs from MBs in which the optimal mode is 16×16 , and LRCs of MBs in which the optimal mode is 16×16 , 16×8 , or 8×16 are considered to obtain μ_1 .

The exhaustive experiments are performed to obtain μ_i according to each GRC and QP, and afterwards these values are modeled as a function of GRC and QP. For this, the training sequences from Table 2 have been used with the ten QPs, 12, 16, 20, 24, 28, 32, 36, 40, 44, and 48. The training sequences have been selected in considering resolutions and sequence characteristics. Empirically, we have selected as $\alpha_0 = 1.1$ and $\alpha_1 = 1.5$.

Figure 3 represents the thresholds L_i , $i \in \{0, 1\}$, according to GRC from the exhaustive experiment. Figure 3a,b shows the results for QP = 24 and QP = 36,

Table 2 Summary of the training sequences and the non-training sequences

Format	Size (pixel)	Training sequence	Non-training sequence
QCIF	176×144	Carphone	Claire
		Coastguard	Highway
		Hall	Miss-America
		Soccer	News
CIF	352×288	Akiyo	Container
		Mother-Daughter	Football
		Silent	Foreman
		Mobile	Paris
WQVGA	416×240	BasketballPass	BlowingBubbles
		BQSquare	RaceHorses
WVGA	832×480	BasketballDrill	BQMall
		PartyScene	RaceHorses
HD (720 p)	$1,280 \times 720$	FourPeople	KristenAndSara
		Johnny	
HD (1,080 p)	$1,920 \times 1,080$	Kimono	ParkScene
		Cactus	BasketballDrive
			BQTerrace
UHD (4 K)	$2,560 \times 1,600$	PeopleOnStreet	Traffic

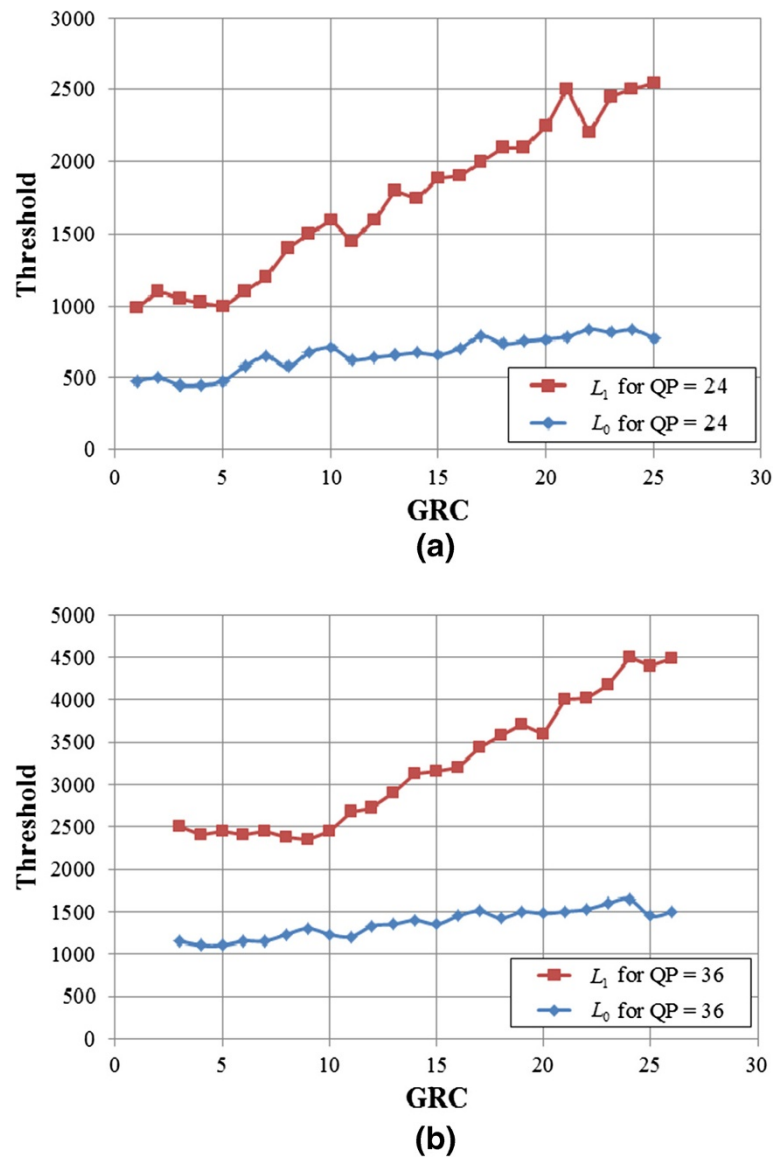


Figure 3 Thresholds L_0 and L_1 according to GRC in (a) QP = 24 and (b) QP = 36.

Table 3 Parameters for QPs, 12, 16, 20, 24, 28, 32, 36, 44, and 48

Parameter	QP									
	12	16	20	24	28	32	36	40	44	48
a_0	242	317	423	482	672	891	1,186	1,578	2,065	2,802
b_0	10	12	14	16	18	21	25	30	34	40
c_0	234	304	396	452	592	788	1,104	1,228	1,815	2,111
a_1	283	422	581	852	1,512	1,914	2,343	3,707	5,963	8,471
b_1	37	43	56	68	86	110	135	182	234	271
c_1	176	238	397	514	663	906	1,408	1,910	2,405	3,266

respectively. It is observed from Figure 3 that the thresholds are divided into two regions according to the tendency of the slopes. For example, in Figure 3a, L_i , $i \in \{0, 1\}$ can be divided into two regions, $GRC \leq 4$ and $4 < GRC$. In $GRC \leq 4$, L_i , $i \in \{0, 1\}$ can be approximated as constants because the slopes are close to zero. In $4 < GRC$, L_i , $i \in \{0, 1\}$ are approximated as linear functions. A similar tendency can be observed

from the slope in Figure 3b. Consequently, L_i is modeled as follows:

$$L_i = \begin{cases} \alpha_i & \text{if } GRC \leq G \\ b_i \times GRC + c_i & \text{if } G < GRC \end{cases}, \quad (5)$$

where parameters α_i , b_i , and c_i are determined in Table 3, and G is defined according to the tendency

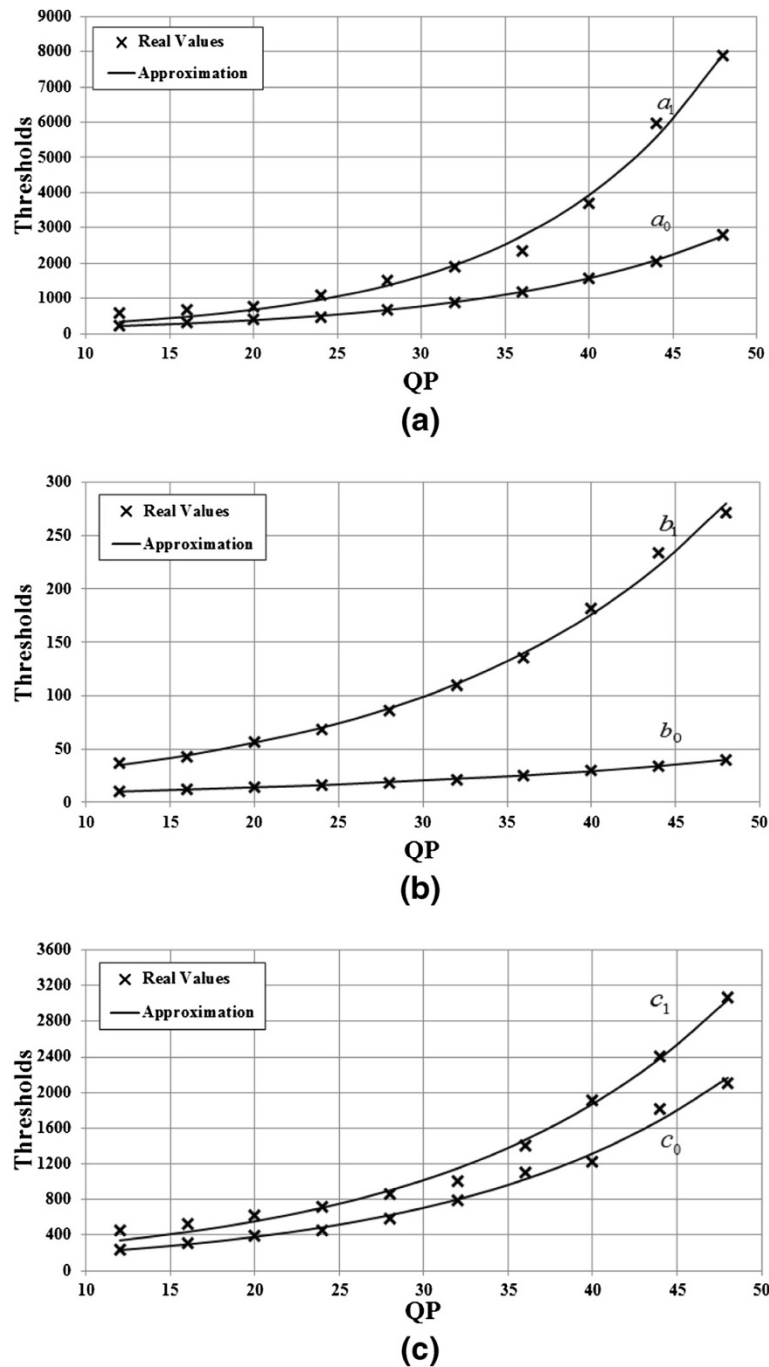


Figure 4 Parameter value approximations for (a) a_0 and a_1 , (b) b_0 and b_1 , (c) c_0 and c_1 .

of the slopes for L_i as follows:

$$G = \max\left(0, \frac{\lfloor QP-16 \rfloor}{4}\right) + 2. \quad (6)$$

By plotting the parameter values of the ten QPs given in Table 3 onto Figure 4, it can be observed that these values are estimated as the exponential function, $f(QP) = p \times e^{q \times QP}$. Accordingly, the parameter values in Table 3, α_i , b_i , and c_i , $i \in \{0, 1\}$ are determined as follows:

$$\begin{cases} a_0 = 93.76 \times e^{0.07060 \times QP} \\ b_0 = 6.312 \times e^{0.03842 \times QP} \\ c_0 = 110.0 \times e^{0.06210 \times QP} \\ a_1 = 118.5 \times e^{0.08757 \times QP} \\ b_1 = 17.65 \times e^{0.05755 \times QP} \\ c_1 = 165.2 \times e^{0.06070 \times QP} \end{cases} \quad (7)$$

2.3 Overall algorithm

Figure 5 represents the flowchart of the proposed intermode decision algorithm. In each frame, GRC is calculated, and then thresholds L_0 and L_1 are determined. For each MB, LRC is calculated and MB activity is determined. Finally, according to MB activity, candidate intermodes for the RDO process are determined for each MB. The overall procedure of the proposed algorithm is described as follows:

Step 1: Calculate GRC according to (3) for current frame.

Step 2: Determine two thresholds, L_0 and L_1 , based on GRC according to (5), (6), and (7).

Step 3: Perform ME of 16×16 block size, and then calculate LRC according to (1) in each MB for current frame.

Step 4: Determine MB activity of current MB based on LRC and two thresholds L_0 and L_1 , according to (2).

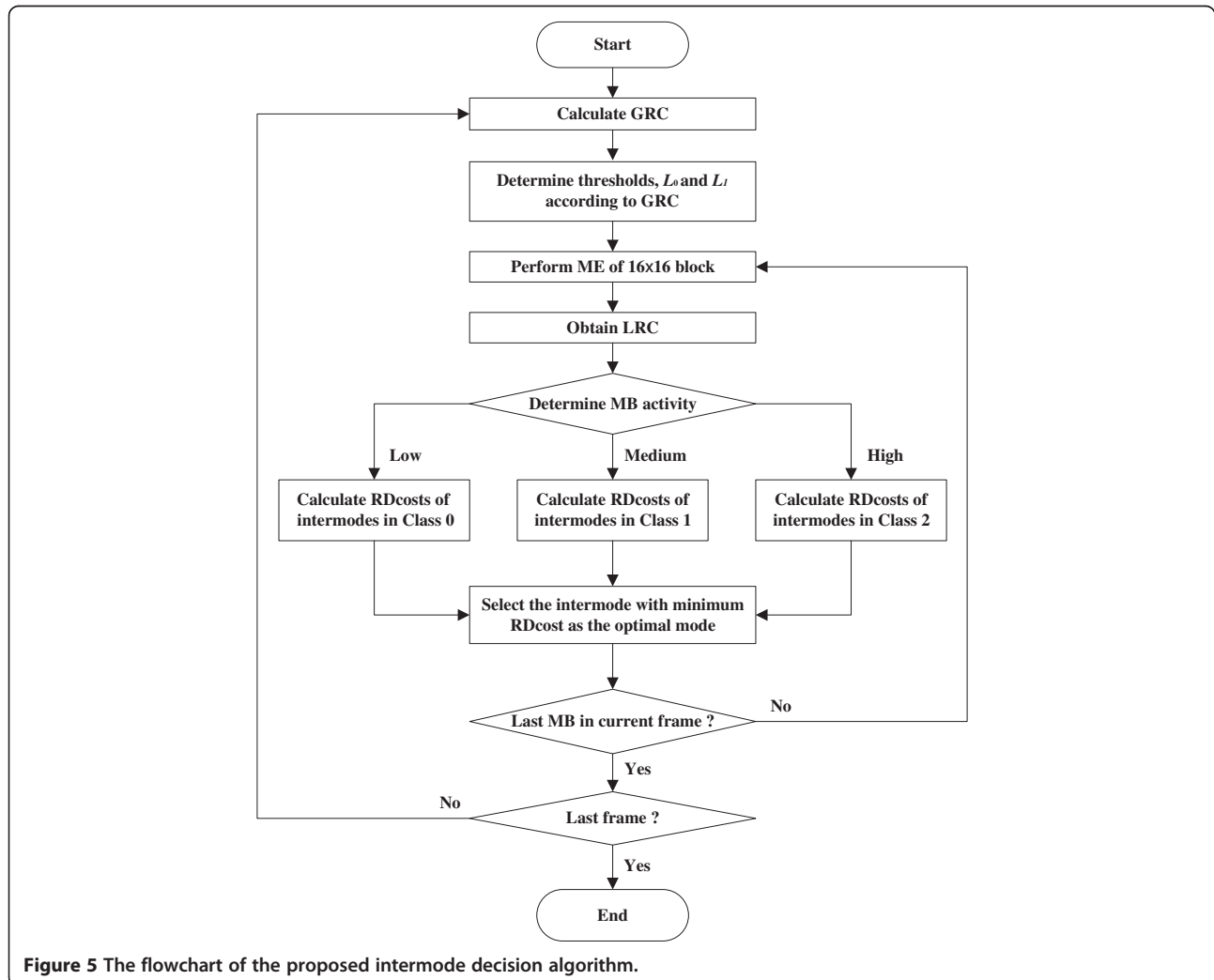


Figure 5 The flowchart of the proposed intermode decision algorithm.

Table 4 Performance of the proposed algorithms in terms of BDPSNR, BDBR, and TS

	Sequence	Format	BDPSNR	BDBR	TS
Training sequences	Carphone	QCIF	-0.09	0.10	-64
	Coastguard		0.00	0.06	-47
	Hall		-0.01	-0.63	-67
	Soccer		-0.01	1.30	-53
	Akiyo	CIF	-0.03	-0.33	-76
	Mother-Daughter		-0.02	-0.30	-70
	Silent		-0.01	-0.14	-60
	Mobile		-0.04	0.87	-56
	BasketballPass	WQVGA	-0.02	0.29	-60
	BQSquare		-0.01	0.50	-56
	BasketballDrill	WVGA	-0.02	0.29	-59
	PartyScene		-0.01	0.38	-50
	FourPeople	HD (720 p)	-0.02	-0.05	-74
	Johnny		-0.04	-0.50	-77
	Kimono	HD (1,080 p)	-0.06	-0.47	-67
	Cactus		-0.01	0.11	-60
	PeopleOnStreet	UHD (4 K)	-0.04	0.40	-56
Non-training sequences	Claire	QCIF	-0.08	-0.02	-74
	Highway		0.01	1.50	-73
	Miss-America		-0.06	-0.36	-75
	News		-0.05	-0.14	-63
	Container	CIF	-0.02	-0.67	-69
	Football		-0.04	0.67	-48
	Foreman		-0.04	0.34	-62
	Paris		-0.02	0.53	-54
	BlowingBubbles	WQVGA	-0.01	0.10	-47
	RaceHorses		-0.03	1.11	-58
	BQMall	WVGA	-0.01	-0.22	-52
	RaceHorses		-0.04	1.50	-55
	KristenAndSara	HD (720 p)	-0.03	-0.15	-77
	ParkScene		-0.02	0.23	-64
	BasketballDrive		-0.05	-0.04	-64
	BQTerrace		-0.02	0.11	-68
	Traffic	UHD (4 K)	-0.02	-0.12	-69
Average	-	-	-0.03	0.18	-63

Step 5: Determine the class of the current MB based on its MB activity from Table 1.

Step 6: Perform RDO process and obtain RDcosts for only candidate intermodes that belong to the selected class.

Step 7: Determine the mode with minimum RDcost as the optimal mode.

Step 8: If current MB is the last MB in current frame, go to the step 9. Otherwise, go to the step 3 for the next MB in current frame.

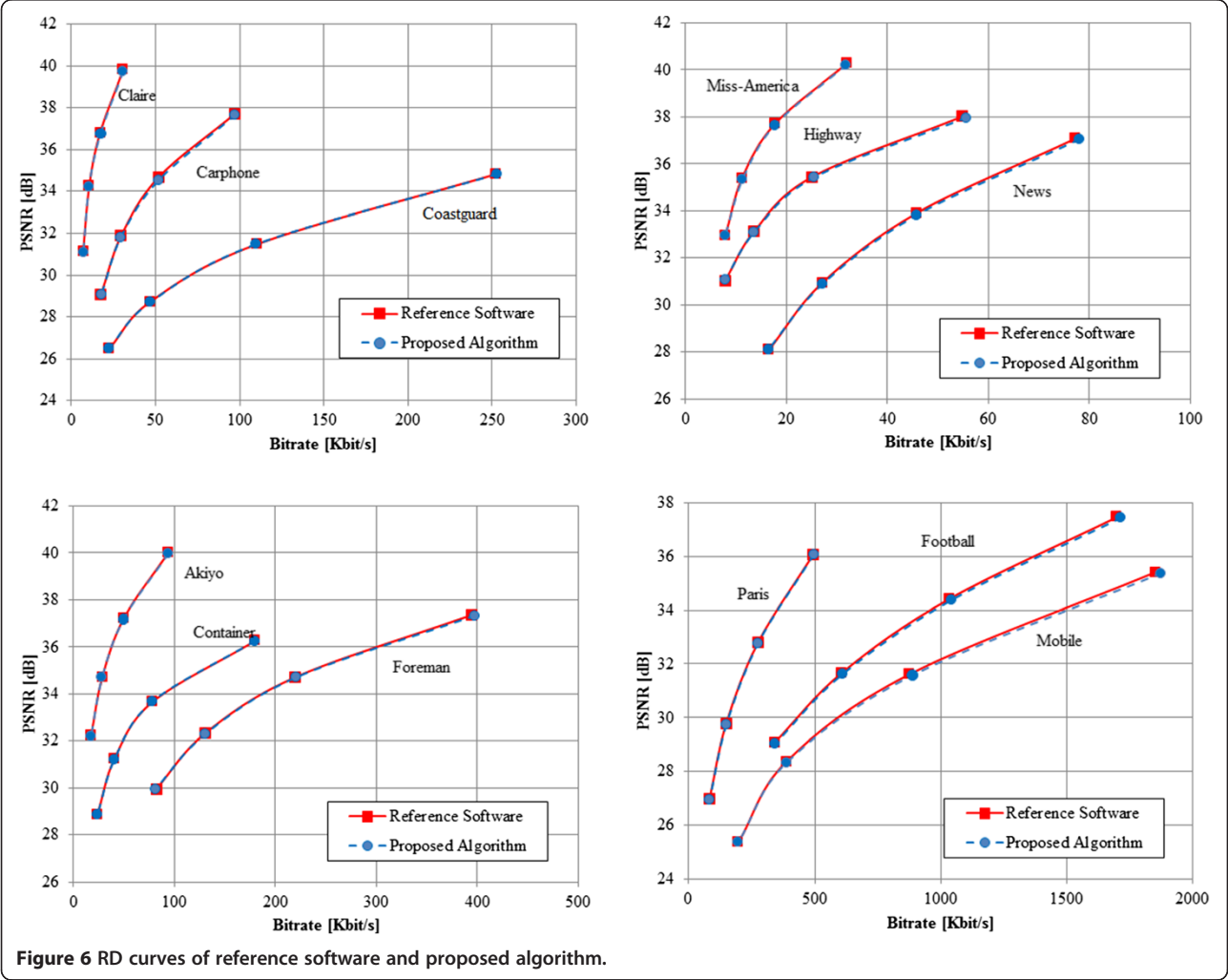
Step 9: If current frame is the last frame, finalize the encoding process. Otherwise, go to the step 1 for the next frame.

3 Simulation results

The proposed intermode decision algorithm is implemented into the H.264/AVC reference software, JM 13.2 [17]. The experiments are performed on a PC with 3.2 GHz CPU and 6 GB RAM by using the training sequences used for the threshold modeling and the non-training sequences as seen in Table 2. The test conditions are set as follows: encoding frame is 100; ME search range is 32; five reference frames are used; motion vector resolution is 1/4 pel; RDO and CABAC are used; fast ME [18] is used; GOP structure is IPPP; and four QPs (28, 32, 36, 40) are used. Bjontegaard delta

Table 5 Performance comparison of three intermode decision algorithms in terms of BDPSNR, BDBR, and TS

Method		Proposed			Martínez-Enríquez's			Jeon's		
Sequence	Format	BDPSNR	BDBR	TS	BDPSNR	BDBR	TS	BDPSNR	BDBR	TS
Carphone	QCIF	−0.09	0.10	−64	−0.06	0.84	−39	0.01	−0.27	−9
Claire		−0.08	−0.02	−74	0.01	−0.17	−65	0.01	0.1	−20
Coastguard		0.00	0.06	−47	−0.04	1.74	−28	−0.01	0.12	−7
Highway		0.01	1.50	−73	−0.08	2.55	−53	−0.02	0.67	−12
Miss-America		−0.06	−0.36	−75	0	0.01	−60	0	−0.14	−20
News		−0.05	−0.14	−63	−0.03	0.65	−57	−0.02	0.24	−19
Akiyo	CIF	−0.03	−0.33	−76	0	−0.14	−62	0	−0.29	−22
Container		−0.02	−0.67	−69	−0.01	0.18	−55	0.01	−0.34	−23
Football		−0.04	0.67	−48	0	0.03	−23	0	0.01	−10
Foreman		−0.04	0.34	−62	−0.03	0.49	−36	0	−0.05	−10
Mobile		−0.04	0.87	−56	−0.02	0.34	−29	−0.01	0.17	−6
Paris		−0.02	0.53	−54	−0.05	0.89	−40	0	0.03	−15
Average		−0.04	0.21	−63	−0.02	0.61	−46	0	0.02	−14



peak signal-to-noise ratio (BDPSNR), Bjontegaard delta bit rate (BDBR) [19], and time saving (TS) are used in order to evaluate the performance. TS is defined as follows.

$$TS = \frac{T_{\text{proposed}} - T_{\text{ref}}}{T_{\text{ref}}} \times 100\%, \quad (8)$$

where T_{ref} and T_{proposed} are the encoding times of the reference software and proposed algorithm, respectively. For BDBR, BDPSNR, and TS, the positive values represent an increase whereas negative values indicate a decrease.

The simulation results are represented based on the training sequences and the non-training sequence, together in Table 4. Also, the proposed algorithm is compared with two preceding works, Jeon's [5,6] and Martínez-Enríquez's [16] in Table 5. Jeon's and Martínez-Enríquez's methods are two of the most-cited fast mode decision algorithm. Especially, Jeon's method has been adopted as fast mode decision algorithm in JM reference software. Martínez-Enríquez's method has been reported recently, and its performance is outstanding.

It is observed from Table 4 that the proposed algorithm reduces the encoding time by 63% on average with negligible coding loss in terms of BDPSNR and BDBR, with the values of -0.03 dB and 0.19%, respectively. The proposed algorithm also outperforms Jeon's and Martínez-Enríquez's, in terms of TS, by -49% and -17% on average, respectively, from Table 5. In particular, the improvement of TS for sequences with high motion or detailed regions, such as Coastguard, Football, and Mobile, is outstanding compared to the two works. For example, in the Mobile sequence, TS of the proposed algorithm is -56%, while TS of Martínez-Enríquez's and Jeon's algorithms are -29% and -6%, respectively.

To evaluate rate-distortion (RD) performance, the RD curves of the reference software and proposed algorithm are represented in Figure 6. The RD curves of the reference software refer to the optimal performance bounds. From Figure 6, it can be seen that the proposed fast intermode decision algorithm achieves a similar RD performance as in that of the exhaustive mode decision of the JM reference software.

4 Conclusions

In this paper, the fast intermode decision algorithm is proposed to reduce encoder complexity. In particular, based on the high correlation between the residual error and optimal mode, GRC and LRC, which can be obtained simply and inexpensively, are defined. The candidate intermodes are efficiently determined according to MB activity evaluated on the GRC and

LRC, and RDO processes can be skipped for unnecessary intermodes. Therefore, the proposed algorithm greatly reduces the encoding time with negligible degradation of coding performance and achieves significantly better results than that of the previous two works, Jeon's and Martínez-Enríquez's, for all sequences including ones with high motion or detailed region.

Abbreviations

BDBR: Bjontegaard delta bit rate; BDPSNR: Bjontegaard delta peak signal-to-noise ratio; CABAC: context-based adaptive binary arithmetic coding; FAT: fast adaptive early termination; GRC: global residual complexity; LRC: local residual complexity; MB: macroblock; ME: motion estimation (ME); MV: motion vector; RD: rate-distortion; RDcost: rate-distortion cost; RDO: rate-distortion optimization; TS: time saving.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by the MKE (The Ministry of Knowledge Economy), South Korea, under the 'ITRC (Information Technology Research Center)' support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2012-H0301-12-1008). Also, this work was supported by the National Research Foundation of Korea (NRF) with grant funded by the Korean government (MEST) (no. 2011-0016302).

Author details

¹Department of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, South Korea. ²Samsung Electronics, Suwon, Gyeonggi-do 442-600, South Korea.

Received: 20 December 2012 Accepted: 15 April 2013

Published: 10 May 2013

References

1. ISO/IEC, ISO/IEC 14496-10, Information Technology-Coding of Audio-Visual Objects-Part 10. Advanced Video Coding (ISO/IEC, 2003)
2. T Wiegand, GJ Sullivan, G Bjontegaard, A Luthra, Overview of the H.264/AVC video coding standard. *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 560-576 (2003)
3. J Ostermann, J Bormans, P List, D Marpe, M Narroschke, F Pereira, T Stockhammer, T Wedi, Video coding with H.264/AVC: tools, performance and complexity. *IEEE Trans. Circuits Syst. Mag.* **4**(1), 7-28 (2004)
4. T Wiegand, H Schwarz, A Joch, F Kossentini, G Sullivan, Rate-constrained coder control and comparison of video coding standards. *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 688-703 (2003)
5. B Jeon, Fast Mode Decision for H.264 (ISO/IEC, 2003) ISO/IEC JTC1/SC29/WG11 and ITU-T SG16, Input Document JVT-J033. http://wftp3.itu.int/av-arch/jvt-site/2003_12_Waikoloa/
6. I Choi, J Lee, B Jeon, Fast coding mode selection with rate distortion optimization for MPEG-4 part-10 AVC/H.264. *IEEE Trans. Circuits Syst. Video Technol.* **16**(12), 1557-1561 (2006)
7. C Crecos, MY Yang, Fast inter mode prediction for P slices in the H.264 video coding standard. *IEEE Trans. Broadcast.* **51**(2), 256-263 (2005)
8. A Saha, K Mallick, J Mukherjee, S Sural, SKIP prediction for fast rate distortion optimization in H.264. *IEEE Trans. Consumer Electron.* **53**(3), 1153-1160 (2007)
9. D Wu, F Pan, KP Lim, S Wu, ZG Li, X Lin, S Rahardja, CC Ko, Fast intermode decision in H.264/AVC video coding. *IEEE Trans. Circuits Syst. Video Technol.* **15**(7), 953-958 (Jul. 2005)
10. BD Choi, JH Nam, MC Hwang, SJ Ko, Fast motion estimation and inter mode selection for H.264. *EURASIP J. Adv. Signal Process.* (2006). doi:10.1155/ASP/2006/71643
11. J Ren, N Kehtarnavaz, M Budagavi, Computationally efficient mode selection in H.264/AVC video coding. *IEEE Trans. Consumer Electron.* **54**(2), 877-886 (2008)

12. B-G Kim, Novel inter-mode decision algorithm based on macroblock (MB) tracking for the P-slice in H.264/AVC video coding. *IEEE Trans. Circuits Syst. Video Technol.* **18**(2), 273–279 (2008)
13. H Zeng, C Cai, K-K Ma, Fast mode decision for H.264/AVC based on macroblock motion activity. *IEEE Trans. Circuits Syst. Video Technol.* **19**(4), 491–499 (2009)
14. Z Liu, L Shen, Z Zhang, An efficient intermode decision algorithm based on motion homogeneity for H.264/AVC. *IEEE Trans. Circuits Syst. Video Technol.* **19**(1), 128–132 (2009)
15. K Bharanitharan, BD Liu, JF Yang, Classified region algorithm for fast inter mode decision in H.264/AVC encoder. *EURASIP J. Adv. Signal Process* (2010). doi:10.1155/2010/150809
16. E Martínez-Enríquez, A Jiménez-Moreno, F Díaz-de-Maria, An adaptive algorithm for fast inter mode decision in the H.264/AVC video coding standard. *IEEE Trans. Consumer Electron* **56**(2), 826–834 (2010)
17. Joint Video Team (JVT), H.264/AVC reference software. <http://iphome.hhi.de/suehring/tml>
18. Z Chen, P Zhou, Y He, Fast motion estimation for JVT. JVT-G016.doc, in *7th Meeting of the Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Pattaya II, March 2003*
19. G Bjontegaard, Calculation of average PSNR differences between RD-curves, in *VCEG-M33, 13th Meeting, Austin, Texas, 2–4 April 2001*

doi:10.1186/1687-5281-2013-30

Cite this article as: Lee et al.: Fast intermode decision algorithm based on general and local residual complexity in H.264/AVC. *EURASIP Journal on Image and Video Processing* 2013 **2013**:30.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com