Fast Intra Mode Decision for HEVC

Ban Doan[0000−0003−0900−6284] and Andrey Tropchenko[000−0001−9812−7947]

ITMO University, Saint Petersburg, 197101, Russian Federation
bandoan@itmo.ru, aatropchenko@itmo.ru

Abstract. For the higher coding performance than the previous video coding standards, High-Efficiency Video Coding (HEVC) adopts an intra prediction method with 35 modes, which requires heavy computational complexity. Intending to reduce this complexity, we analyzed the role of modes and proposed a scheme that contains two rough mode decision (RMD) processes with a customized set of modes to be tested in the first stage. The second stage of the RMD is calculated for a maximum of 4 modes. As compared to the default encoding scheme in HEVC test model HM-16.20, experimental results show that the proposed method reduces encoding time up to 22.74% with negligible loss of coding efficiency.

Keywords: HEVC/H.265 · Video compression · Intra prediction · Mode decision · Rate-distortion optimization.

1 Introduction

In recent years, there has been a growing interest in services related to the transmission and storage of high and ultrahigh definition videos. The video coding standard H.264/Advanced Video Coding (AVC) [1] published in 2003 has been unable to meet those requirements and the introduction of the HEVC [2] video coding standard as one of the solutions to the problem.

Mainly due to the new coding tools and the flexible data structures, HEVC provides a significant improvement in compression efficiency compared to its predecessors H.264, especially when operating on high-resolution video content [4, 3]. Similar to older video compression technologies, HEVC is based on a hybrid scheme of coding image blocks, which uses intra- and inter-frame prediction coding together with transform coding of residual data.

HEVC contains several elements improving the efficiency of intra prediction over earlier solutions. HEVC design supports a total of 35 intra prediction modes, including Planar, DC and 33 angular modes, as presented in Figure 1, which contribute to representing different texture and object edge direction more precisely [5]. Due to the significantly increased number of intra modes, more techniques are required to efficiently encode the mode, one of which is to divide the frame into segments called coding units (CU), prediction units (PU), and transformation units (TU). The encoder needs to try all the combinations of

CU, PU, and TU in the rate-distortion optimization (RDO) process to find the best mode with the lowest rate-distortion (RD) cost [5]. Such a process is very time-consuming.

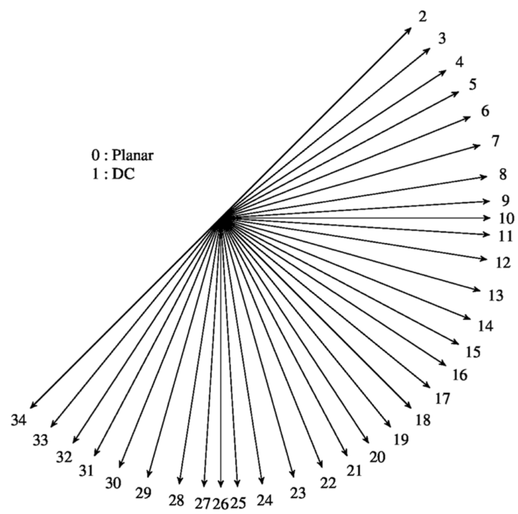


Fig. 1. Intra prediction modes in HEVC

While an increase in the number of intra prediction modes can provide substantial performance gains, it also makes the RDO process more complex. To reduce the computational load of intra prediction, the official HM software [6] uses a fast encoding algorithm [5, 7, 10] with two phases through a combination of RMD and RDO process. First, all 35 modes are evaluated with respect to a cost function. N modes with minimum cost J_{SATD} are then selected as the most promising candidate modes.

$$J_{SATD} = SATD + \lambda_{pred} \times R_{pred} \tag{1}$$

where $SATD$ represents the absolute sum of Hadamard transformed residual signal for a PU. λ_{pred} is a Lagrange multiplier, and R_{pred} represents the number of bits for the prediction mode. The number N is varied depending on the PU size. The N is set to $\{8, 8, 3, 3, 3\}$ for 4×4, 8×8, 16×16, 32×32, and 64×64 PU, respectively.

In the second step, three most probable modes (MPM), which are derived from the intra modes of the left and top neighboring PUs [5], are added to the list of candidates [8, 9]. The full RD costs with the reconstructed residual signal used for the actual encoding process are compared among those $(N + 3)$ modes, and the prediction mode with the minimum RD cost is selected as the final prediction mode. The RD cost (J_{RDO}) for each intra mode is computed by:

$$J_{RDO} = SSE + \lambda_{pred} \times R_{total} \quad (2)$$

where SSE represents the sum of the squared errors between the original CU and the reconstructed CU. R_{total} is the total number of bits used for encoding with this mode.

In this way, the RDO process has to check only a maximum of 11 modes instead of all 35, and so the computational load can be reduced. However, the complexity is still high, since, in the RMD step all the 35 modes need to perform the cost calculation, the number of modes for RDO is still large. In this paper, a fast intra mode decision is proposed to further reduce the complexity of HEVC intra coding while maintaining the RD performance.

2 Analysis mode selection probability

Theoretically, 35 intra modes play the same role and their probability of choice is equal. However, the results of analyzing some videos of various categories have given a different perspective. To perform the statistical analyses of frequently chosen modes, HEVC reference software HM-16.20 was used to encode a set of video sequences of different classes and resolutions. Statistical results for test sequences in class B and the sequence "PeopleOnStreet" are shown in tables 1 and 2 with the four most frequent modes are represented in bold.

Table 1. Average distribution of intra prediction modes for B-class test sequences.

Mode	Frequency, %	Mode	Frequency, %	Mode	Frequency, %
Planar	**23.31**	12	1.95	24	2.02
DC	**13.31**	13	1.65	25	2.75
2	1.01	14	1.54	**26**	**8.45**
3	0.89	15	1.35	27	2.32
4	0.89	16	1.13	28	1.74
5	1.18	17	1.25	29	1.55
6	2.00	18	1.23	30	1.33
7	2.02	19	1.27	31	1.00
8	2.13	20	1.22	32	0.92
9	2.97	21	1.32	33	0.98
10	**6.05**	22	1.45	34	0.98
11	3.11	23	1.74		

Figure 2 shows diagrams reflecting the probability (P) of the choice of specific prediction modes when encoding test video sequences.

Statistics show that Planar and DC prediction modes are most likely for all video sequences. However, the probability of being the best choice for the vertical and horizontal modes (Angular10 and Angular26) is much greater than the other angular modes.

Table 2. Distribution of intra prediction modes for test sequence "PeopleOnStreet".

Mode	Frequency, %	Mode	Frequency, %	Mode	Frequency, %
Planar	**20.27**	12	1.05	24	3.09
DC	**12.11**	13	1.24	25	2.62
2	1.61	14	1.83	**26**	**7.72**
3	1.35	15	1.06	27	2.24
4	1.70	16	0.81	28	2.35
5	2.62	17	0.91	29	1.94
6	3.77	18	0.78	30	1.61
7	2.68	19	0.97	31	1.11
8	2.22	20	0.98	32	0.94
9	3.85	21	1.23	33	1.21
10	**4.69**	22	1.83	34	1.57
11	1.27	23	2.75		

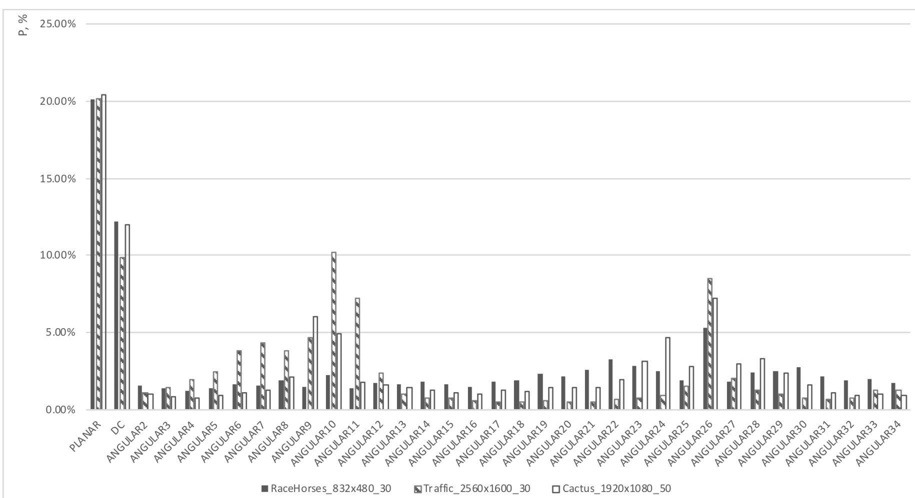


Fig. 2. Mode selection probability.

There has been significant work to speed up the intra mode decision process. Based on the above analysis, this paper proposes a scheme for reducing intra modes in RMD and hence increasing the encoding speed, the descriptions of which will be described below.

3 Fast mode selection

Instead of all 35, we prefer to test the most probability modes. For this purpose, a set of modes was created, including DC, Planar and Angular modes $(2 + 4i)$ where $0 \leq i \leq 8$. As a result, only 11 modes are tested using the RMD process to find modes with the lowest cost. Let's call the 2 best modes FM (first mode) and SM (second mode). After that, a flexible step (RMD2) was added after the first. The input data for RMD2 depends on what the FM and SM are. The general scheme is shown in figure 3 with the mode selection algorithm for the second step is presented as follows:

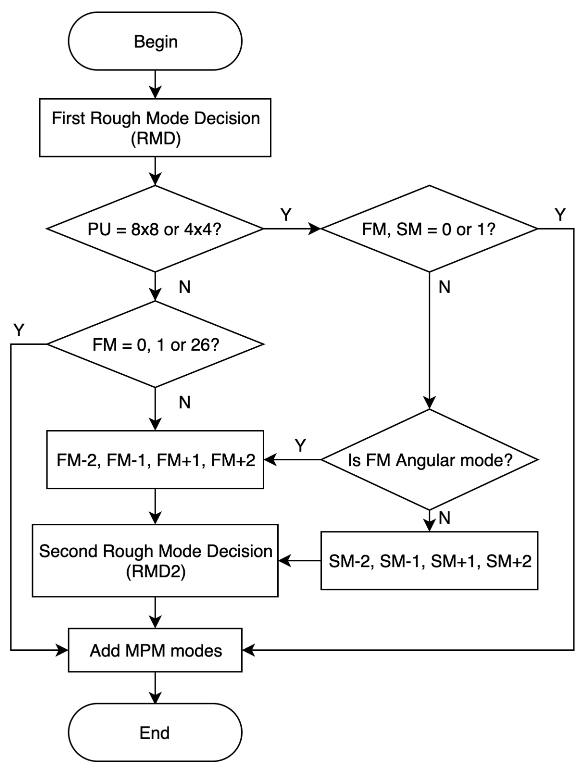


Fig. 3. Block diagram of the proposed algorithm

- For PU 16×16, 32×32 and 64×64: check if FM is not Planar, DC or vertical mode (0, 1, 26) then perform the second step of calculating RMD2 for

4 adjacent FM modes: $FM-2$, $FM-1$, $FM+1$, $FM+2$, after that update the list of candidates. Otherwise, RMD2 is skipped.

- For PU 8×8 and 4×4: if FM and SM are Planar or DC modes, the encoder will add MPM modes (if not already included in the candidate list) and perform the RD cost calculation step. In other cases, the second step RMD2 will check some other modes $FM-2$, $FM-1$, $FM+1$, $FM+2$ (if FM is an angular mode) or $SM-2$, $SM-1$, $SM+1$, $SM+2$ (if FM is DC or Planar).

According to the proposed scheme, in the second stage of calculating cost RMD2, we can calculate a maximum of 4 more modes surrounding FM and SM. The minimum and the maximum number of checked modes is 11 and 15, respectively, which reduces computational load. After that, the RDO process will be performed with MPM modes added to the candidate list.

4 Experimental results

The proposed scheme has been implemented on top of the HEVC reference software HM-16.20. A set of standard video sequences in five classes covering a wide range of resolutions and use cases (see Table 3) was tested using the All Intra-Main configuration and four values of the quantization parameter QP 22, 27, 32 and 37 as specified by [11].

Table 3. Test sequences.

Class	***Resolution***	***Sequence***	***Frame Rate***
A	2560×1600	Traffic	30 Hz
		PeopleOnStreet	30 Hz
B	1920×1080	Kimono	24 Hz
		ParkScene	24 Hz
		Cactus	50 Hz
		BasketballDrive	50 Hz
		BQTerrace	60 Hz
C	832×480	BasketballDrill	50 Hz
		BQMall	60 Hz
		PartyScene	50 Hz
		RaceHorses	30 Hz
D	416×240	BasketballPass	50 Hz
		BQSquare	60 Hz
		BlowingBubbles	50 Hz
		RaceHorses	30 Hz
E	1280×720	FourPeople	60 Hz
		Johnny	60 Hz
		KristenAndSara	60 Hz

To evaluate the efficiency of the algorithm, comparisons were made in terms of the Bjontegaard peak signal-to-noise ratio (BD-PSNR) and Bjontegaard bitrate (BD-Bitrate) [12], and time-saving $\Delta T(\%)$.

$$\Delta T = \frac{T_{HM-16.20} - T_{prop}}{T_{HM-16.20}} \tag{3}$$

where $T_{HM-16.20}$ denotes the time consuming of the default HM-16.20 and T_{prop} represents the time consumed by the proposed algorithm.

Table 4 shows that compared to HM-16.20, the proposal can reach up to 22.74% of time-saving. This is because it reduces the number of the candidate modes for both the RMD and RDO processes compared to the original HM software. On the other hand, this happens with a slight decrease in bitrate and negligible degradation in video quality (see Table 5).

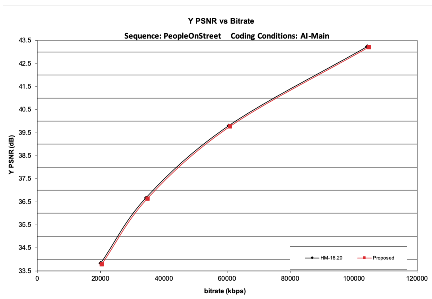
Different sequences are obtained different results, due to different detail and complexity. The RD curves of the proposed algorithm and the HM for some sequences are shown in Figures 4 and 5. It can be seen that the proposed algorithm achieves almost the same PSNR on different bitrates.

Table 4. Encoding time reduction compared to HM-16.20

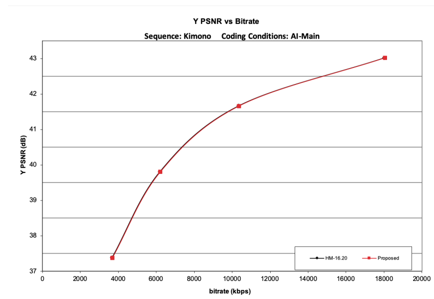
Sequence	ΔT, %					
	QP=22	**QP=27**	**QP=32**	**QP=37**	**Average**	
Traffic	10.11	12.09	15.41	21.69	14.83	13.49
PeopleOnStreet	14.36	12.30	10.53	11.45	12.16	
Kimono	33.81	40.26	13.24	15.36	25.67	16.77
ParkScene	34.31	12.82	17.96	14.29	19.85	
Cactus	16.84	12.09	14.03	12.44	13.85	
BasketballDrive	9.83	10.99	12.03	10.05	10.73	
BQTerrace	13.15	11.86	12.76	17.27	13.76	
BasketballDrill	34.33	34.30	37.50	39.13	36.32	16.74
BQMall	9.06	12.91	10.05	13.06	11.27	
PartyScene	7.80	8.39	7.75	8.63	8.14	
RaceHorses	8.54	10.07	10.19	16.07	11.22	
BasketballPass	22.22	22.36	27.79	24.24	24.15	22.74
BQSquare	27.57	16.37	29.89	36.34	27.54	
BlowingBubbles	33.93	28.95	27.29	26.03	29.05	
RaceHorses	10.86	9.32	8.94	11.74	10.22	
FourPeople	11.56	10.30	14.82	11.57	12.06	12.58
Johnny	11.77	14.90	13.61	11.72	13.00	
KristenAndSara	13.30	11.65	12.43	13.32	12.68	

Table 5. Coding efficiency comparisons between proposed intra prediction and HM-16.20 software.

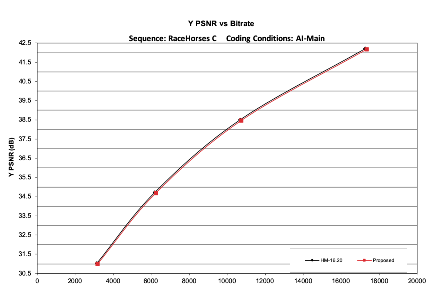
Class	***BD-PSNR (dB)***				***BD-Bitrate (%)***			
	QP=22	QP=27	QP=32	QP=37	QP=22	QP=27	QP=32	QP=37
A	0.026	0.024	0.027	0.033	-0.39	-0.72	-1.09	-1.36
B	0.018	0.015	0.019	0.017	-0.28	-0.69	-1.02	-1.50
C	0.039	0.041	0.044	0.041	-0.74	-0.89	-1.15	-1.46
D	0.046	0.047	0.049	0.051	-0.85	-1.16	-1.59	-2.03
E	0.024	0.028	0.040	0.037	-1.20	-1.91	-2.50	-3.39
Average	0.031	0.031	0.036	0.036	-0.69	-1.07	-1.47	-1.95



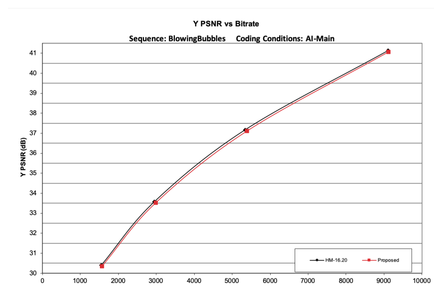
(a)



(b)



(c)



(d)

Fig. 4. RD-curves for some test sequences

5 Conclusion

By interfering with the mode selection for the RMD process, the number of modes to be tested has been significantly reduced. It can be argued that the proposed scheme requires less computational load, while the coding performance

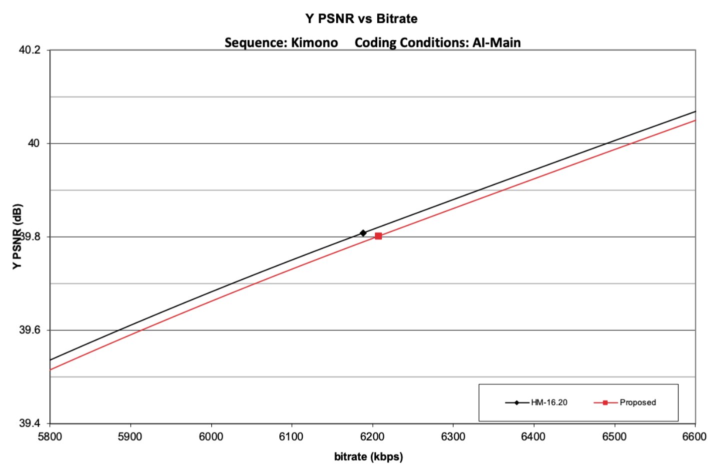


Fig. 5. RD-curve for sequence "Kimono" around QP=27

remains almost at the same level compared to the original HEVC encoder. It may be recommended to be combined with an algorithm that optimizes the CU separation process.

References

1. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Transactions on Circuits and Systems for Video Technology **17**(7), 560–576 (2003) https://doi.org/10.1109/TCSVT.2003.815165
2. Sullivan, G.J., Ohm, J.-R., Han, W.-J., Wiegand, T.: Overview of the high efficiency video coding (HEVC) standard. IEEE Transactions on Circuits and Systems for Video Technology **22**(12), 1649–1668 (2012) https://doi.org/10.1109/TCSVT.2012.2221191
3. Vanne, J., Viitanen, M., Hamalainen, T. D., Hallapuro, A.: Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Codecs. IEEE Transactions on Circuits and Systems for Video Technology. **22**(12), 1885–1898 (2012) https://doi.org/:10.1109/tcsvt.2012.2223013
4. Ohm, J.-R., Sullivan, G.J., Schwarz, H., Tan, T.K., Wiegand, T.: Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC). IEEE Transactions on Circuits and Systems for Video Technology **22**(13), 1669-–1684 (2013) https://doi.org/10.1109/TCSVT.2012.2221192
5. Lainema, J., Bossen, F., Han, W.-J., Min, J., Ugur, K.: Intra Coding of the HEVC Standard. IEEE Transactions on Circuits and Systems for Video Technology **22**(12), 1792–1802 (2012) https://doi.org/10.1109/TCSVT.2012.2221525
6. HEVC reference software, https://hevc.hhi.fraunhofer.de/. Last accessed Nov 2019
7. Hosseini, E., Pakdaman, F., Hashemi, M.R., Ghanbari, M.: A computationally scalable fast intra coding scheme for HEVC video encoder. Multimed Tools Appl. **78**(9), 11607–11630 (2019) https://doi.org/10.1007/s11042-018-6713-y

8. Zhao, L., Zhang, L., Ma, S., Zhao, D.: Fast mode decision algorithm for intra prediction in HEVC. 2011 Visual Communications and Image Processing (VCIP) https://doi.org/10.1109/VCIP.2011.6115979
9. Zhao, L., Fan, X., Ma, S., Zhao, D.: Fast intra-encoding algorithm for High Efficiency Video Coding. Signal Processing: Image Communication. **29**(9), 935–944 (2014) https://doi.org/10.1016/j.image.2014.06.008
10. Piao, Y., Min, J., Chen, J.: Encoder improvement of unified intra prediction. Document JCTVC-C207, JCT-VC. Guangzhou, CN (2010)
11. Bossen, F.: Common test conditions and software reference configurations. Document JCTVC-L1100, JCT-VC. Geneva, CH (2013)
12. Bjontegaard, G.: Calculation of average PSNR differences between RD-curves. Document VCEG-M33, ITU-T. Austin, Texas (2001)