

The Study on Decision Tree Classification and Three-Step Search Algorithm Based Compression Technology for High Efficiency Video Coding

Chang Liu

Beihang University, Beijing, China

lc76magic@buaa.edu.cn

Keywords: High Efficiency Video Coding; Decision Tree; Three-Step Search

Abstract. The new generation video coding algorithm HEVC reduces the bitrate by half compared to the previous standard H.264 with the same video quality, but this also brings the increasing of encoding time. In this paper, we mainly improve the intra frame prediction to solve this question: Combined with the decision tree method in machine learning, the reference pixel smoothing characteristics of the prediction unit of each partition mode are quickly determined, and some intra modes with the same prediction result are adaptively skipped, which greatly saves coding time. Experimental results show that the proposed decision tree and three-step search based intra prediction algorithm achieves better performance compared with the HM15.0 encoder. The BD-Rate is increased by 1.19%, the BD-PSNR is decreased by 0.058dB, and the encoding time is reduced by 37.87% on average.

Introduction

ISO/IEC and ITU have instituted a series of video compression standards [1] in the past few decades, in order to lower the code rate of video transmission while guaranteeing the video quality. A great many new complicated algorithms are introduced into HEVC, which makes the code stream of the compressed video with the same video quality reduce by half compared with the standard of the last generation H.264, yet also increases the coding time. Thus, the key tasks for optimization of HEVC are to reduce the complexity of the algorithm and decrease the compression time on the precondition of guaranteeing the compression quality.

Decision Tree-Based Intraframe Prediction Algorithm

The HEVC standard adopts 35 intra- frame prediction modes, which need to be traversed for each image block for each to-be-compressed frame of image, then the RD cost with each mode is calculated to determine the optimal intraframe mode of each block [5,6,7]. Yet we have observed that the video usually teems with flat regions of pixels. These blocks are usually coded using mode 0 or mode 1. Besides, the same result can be obtained by adopting some modes for coding when the reference pixel possesses certain smooth rules. Thus, we can classify the image block into three classifications according to the smoothness of reference pixel of the image block. The smoothness of the reference pixel can be measured by its variance. We have defined three variances to jointly determine the classification of the image block to determine the candidate intraframe mode needing to be traversed for it. We obtain a classification decision tree using the offline training for fear of calculating the threshold for each classification mode, then the classification of the image block can be determined just according to the decision tree, thereby significantly saving the coding time [9].

We can classify PUs into three classifications according to the smoothness of PU reference pixel. The reference pixel of PUs in the first classification is very smooth, so only the modes 0 and 1 need to be checked for PUs now. The reference pixels of the upper part and left part in the second classification of PUs respectively have higher smoothness, and now only 19 modes (0, 1, 10, ..., 26) are usable. The rest are the third classification of PUs, now all the 35 modes need to be checked. We adopt the variance of the following three reference pixels to represent its smoothness.

$$Var_{all} = \frac{1}{4N+1} \left[\sum_{i=0}^{2N} (R_{0,i} - \mu_{all})^2 + \sum_{i=1}^{2N} (R_{i,0} - \mu_{all})^2 \right] \tag{1}$$

$$Var_{above} = \frac{1}{2N} \sum_{i=0}^{2N} (R_{i,0} - \mu_{above})^2 \tag{2}$$

$$Var_{left} = \frac{1}{2N} \sum_{i=0}^{2N} (R_{0,i} - \mu_{left})^2 \tag{3}$$

Where, $R_{i,j}$ ($0 \leq i, j \leq 2N$) corresponds to the reference pixels around the to-be-calculated pixel, and μ_{all} , μ_{above} and μ_{left} respectively correspond to the average of all the reference pixels, above reference pixels and left reference pixels.

The training of the decision tree is realized via the C4.5 algorithm [12-14], which determines the attribute via the information gain ratio. It can prune in the process of constructing the decision tree to avoid the existence of some nodes with large error that leads to error accumulation and time waste in its subsequent branching. The C4.5 algorithm is implemented in WEKA via the J48 classifier. 66% of the data obtained previously is used to train the decision tree and the rest 34% will be used to evaluate the precision of the decision tree. The finally constructed decision tree is as shown in Fig.1.

The decision tree has eight leaf nodes. Wherein, 1, 2, 3 respectively represent the three types of PUs defined above. After the prediction over the residual 34% training data via this decision tree, it can be seen from the classification output window that about 85.29% of data has been correctly classified, and 14.71% not. The precision of decision tree constructed for intraframe PU division in literature [15] is 86%. It follows that the constructional precision of the decision tree herein is high.

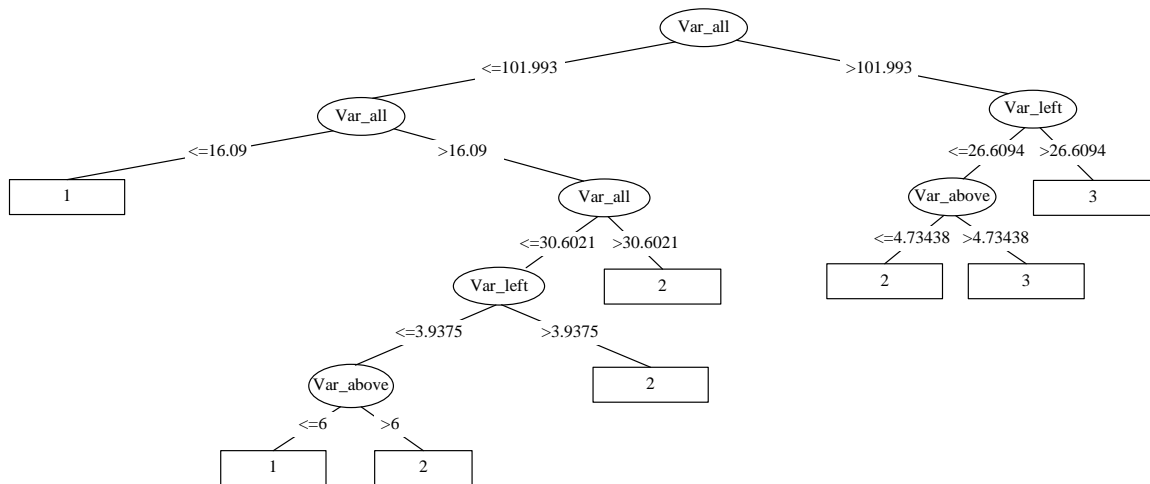


Figure 1. Mode selection decision tree

Optimization Algorithm Based on Three-Step Search

To completely traverse all the intraframe modes still consumes a lot of time, for the PUs classified into the classifications 2 and 3 respectively have 19 and 35 candidate intraframe modes. Thus, we will further cut the mode quantity by a three-step search algorithm for the two classification of PUs [16-20], three-step search algorithm 1 for the PUs of the classification 2 and three-step search algorithm 2 for the PUs of the classification 3. The two search algorithms are described in detail as below:

Three-step search algorithm 1:

Step 1: Define the initial candidate list $L1=\{0,1,10,14,18,22,26\}$ according to the 19 candidate intraframe prediction modes [21], calculate the SATD of each mode in list L1, rank the intraframe modes in L1 in ascending order of SATD, select four modes with the smallest SATD to constitute the new candidate list L2, and assume now $L2=\{1,10,14,18\}$.

Step 2: Select the neighboring intraframe modes by step 2 according to the modes in list L2 to constitute list L3={8,12,16,20}, calculate the SATD of each mode in list L2, select four modes with the smallest SATD from list L2 and list L3 to constitute list L4, and assume now L4={12,14,16,18}.

Step 3: Select the neighboring intraframe modes by step 1 according to the modes in list L4 to constitute list L5={11,13,15,17,19}. Calculate the SATD of each mode in list L5 in like manner, and select four modes with the smallest SATD from list L4 and list L5 to constitute list L6.

Three-step algorithm 2:

Step 1: Define the initial candidate list L1={0,1,2,6,10,14,18,22,26,30,34} according to the 35 candidate intraframe prediction modes [21], calculate the SATD of each mode in list L1, select five modes with the smallest SATD to constitute the new candidate list L2, and assume now L2={0,10,14,18,22}.

Step 2: Select the neighboring intraframe modes by step 2 according to the modes in list L2 to constitute list L3={8,12,16,20,24}, calculate the SATD of each mode in list L2, select five modes with the smallest SATD from list L2 and list L3 to constitute list L4, and assume now L4={10,12,14,16,18}.

Step 3: Select the neighboring intraframe modes [24-26] by step 1 according to the modes in list L4 to constitute list L5={11,13,15,17,19}. Calculate the SATD of each mode in list L5 in like manner, and select five modes with the smallest SATD from list L4 and list L5 to constitute list L6.

Experimental Result

The proposed algorithm is implemented on HEVC standard software HM15.0 on computer with @3.4GHz Intel Core i7-3770 and memory of 8G. The coding quality is evaluated by BD-Rate BD-PSNR and coding efficiency by coding time. The Table 1 provides the comparison between result obtained by using the proposed algorithm to code 20 frames of 18 test video sequences with 5 different resolutions and the coding result by HM15.0.

It is observed from the table that the BD-PSNR of proposed algorithm averagely reduces 0.058dB, BD-Rate averagely increases 1.19% and time averagely saves by 37.87%, compared with the HM15.0. The proposed algorithm delivers very high coding efficiency and has finely lowered coding quality. The coding efficiency of the proposed algorithm is even better than that of the setting S1 which delivers the highest coding efficiency in [29], being about two times that of S1.

Table 1 Experimental result of The proposed algorithm

Video sequence		BD-PSNR (dB)	BD-Rate (%)	TimeSaving(%)
Class A	Traffic	-0.068	1.27	38.96
	PeopleOnStreet	-0.058	1.03	38.06
Class B	Kimono	-0.080	2.22	39.66
	ParkScene	-0.043	1.00	37.90
	Cactus	-0.028	0.73	34.83
	BasketballDrive	-0.042	1.69	40.57
	BQTerrace	-0.061	1.00	38.50
Class C	BasketballDrill	-0.079	1.68	37.99
	BQMall	-0.030	0.48	36.93
	PartyScene	-0.025	0.32	36.01
	RaceHorses	-0.044	0.71	38.67
Class D	BasketballPass	-0.091	1.54	34.29
	BQSquare	-0.056	0.65	40.31
	BlowingBubbles	-0.038	0.63	29.68
	Racehorses	-0.079	1.14	30.69
Class E	FourPeople	-0.023	0.39	40.75
	Johnny	-0.107	2.62	45.75
	KristenAndSara	-0.099	1.92	42.15
Average		-0.058	1.19	37.87

Summary

The high efficiency video intraframe compression algorithm based on decision tree classification and three-step search algorithm proposed in this paper hugely reduces the calculation amount of intraframe prediction and improves the prediction velocity. In order to raise the coding velocity, we firstly adopt intraframe candidate modes of different quantities for different intraframe prediction PUs in the light of the idea of decision tree classification, to judge the classification of intraframe prediction PUs via the decision tree. The decision tree can be defined via the smoothness of PU's reference pixels, and the concrete smoothness can be measured via calculation of variance. Thus, this algorithm can hugely lower the algorithm complexity of intraframe prediction. After the decision tree, we propose a three-step search algorithm for the intraframe mode to further reduce the mode quantity of PUs that still have substantive candidate modes, thereby further raising the coding velocity. The experimental result proves that the proposed algorithm can significantly raise the coding velocity and cut the coding time hardly with any influence on the coding quality.

Reference

- [1] G.J. Sullivan et al.: Overview of the high efficiency video coding (HEVC) standard, *IEEE Transactions on Circuits & Systems for Video Technology*, Vol. 22 (2012) No. 12, p.1649.
- [2] T.K. Tan, et al.: HEVC verification test results. Doc. JCTVC-Q204. *17th Meeting: Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11* (Valencia, España, 2014)
- [3] S.P. Zhu, R.D. Gao and Z. Li: Stereo matching algorithm with guided filter and modified dynamic programming, *Multimedia Tools and Applications*, Vol. 76 (2017) No. 1, p.199.
- [4] VCEG and MPEG. Terms of reference of the Joint Collaborative Team on video coding standard development. *Doc. VCEG-AM90. ITU-T SG16/Q6 VCEG* (Kyoto, Japan, 2010)
- [5] W. Jiang, H. Ma and Y. Chen: Gradient based fast mode decision algorithm for intra prediction in HEVC, *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet). IEEE* (2012) p.1836.
- [6] S.P. Zhu and R.D. Gao: A novel generalized gradient vector flow snake model using minimal surface and component-normalized method for medical image segmentation, *Biomedical Signal Processing and Control*, Vol. 26 (2016) p.1.
- [7] M. Zhang, J. Qu and H. Bai: Entropy-based fast largest coding unit partition algorithm in high efficiency video coding, *Entropy*, Vol. 15 (2013), No. 6, p.2277.
- [8] S.P. Zhu, Q. Zhou and R.D. Gao, A novel snake model using new multi-step decision model for complex image segmentation, *Computers and Electrical Engineering*, Vol. 51 (2016) p.58.
- [9] D. Gwon, H. Choi, J.M. Youn: HEVC fast intra mode decision based on edge and SATD cost, *Asia Pacific Conference on IEEE Multimedia and Broadcasting (APMediaCast)*, (2015) p.1.
- [10] S.P. Zhu, X.Y. Bu, and Q. Zhou: A novel edge preserving active contour model using guide filter and harmonic surface function for infrared image segmentation, *IEEE Access*, Vol. 6 (2018) p.5493.
- [11] L.H. Cancellier, I. Seidel: Energy-efficient hadamard-based SATD architectures, *2014 27th Symposium on IEEE Integrated Circuits and Systems Design (SBCCI)*, (2014) p.1.
- [12] S.P. Zhu, Y.S. Hou, Z.K. Wang, et al: Fractal video sequences coding with region-based functionality, *Applied Mathematical Modelling*, Vol. 36 (2012) No. 11, p.5633.
- [13] I.H. Witten, M. Hall, E. Frank, et al: The WEKA data mining software: An update, *Acm Sigkdd Explorations Newsletter*, Vol. 11 (2009) p.10.
- [14] J.R. Quinlan: *C4.5: programs for machine learning*, Morgan Kaufmann, 1993.
- [15] G. Correa, P. Assuncao, L. Agostini, L. Silva Cruz: A method for early-splitting of HEVC inter blocks based on decision trees, *Proceedings of the 22nd European on Signal Processing Conference (EUSIPCO)* (2014) p.276.

- [16] S.P. Zhu and Z.Y. Xu, Spatiotemporal visual saliency guided perceptual high efficiency video coding with neural network, *Neurocomputing*, Vol. 275 (2018) p.511.
- [17] S.P. Zhu, S.P. Zhang, and C.H. Ran, An improved inter-frame prediction algorithm for video coding based on fractal and H.264, *IEEE Access*, Vol. 5 (2017), p.18715.
- [18] Y. Zhang, S. Kwong, X. Wang, et al: Machine learning-based coding unit depth decisions for flexible complexity allocation in high efficiency video coding, *IEEE Transactions on Image Processing*, Vol.24 (2015) No. 7, p.2225.
- [19] C. Zhu, X. Lin, L.P. Chau: Hexagon-based search pattern for fast block motion estimation, *IEEE Transactions on Circuits & Systems for Video Technology*, Vol.12 (2002) No. 5, p.349.
- [20] S.P. Zhu and X.Z. Zong: Fractal lossy hyperspectral image coding algorithm based on prediction, *IEEE Access*, Vol. 5 (2017) pp.21250.
- [21] S.P. Zhu, D.Y. Zhao and L.Y. Li: Adaptive fast intra prediction for high efficiency video coding, *Multimedia tools and applications*, Vol. 75 (2016) No. 13, p.7559.
- [22] I.K. Kim, J. Min, T. Lee, et al: Block partitioning structure in the HEVC standard, *IEEE Transactions on Circuits & Systems for Video Technology*, Vol. 22 (2012) No. 12, p.1697.
- [23] S.P. Zhu and C.Y. Zhang: A fast algorithm of intra prediction modes pruning for HEVC based on decision trees, *Multimedia Tools and Applications*, Vol. 76 (2017) No. 20, p.21707.
- [24] H. Lei, Z. Yang, H. Lei, et al: Fast intra prediction mode decision for high efficiency video coding, *Proceedings of International Symposium on Computer Communication Control & Automation*, Vol. 68 (2013) No. 12, p.34.
- [25] S.P. Zhu and L.N. Yan: Local stereo matching algorithm with efficient matching cost and adaptive guided image filter, *Visual Computer*, Vol. 33 (2017) No. 9, p.1087.
- [26] S.P. Zhu and Y. Gao: Noncontact 3-D coordinates measurement of cross-cutting feature points on the surface of a large-scale workpiece based on the machine vision method, *IEEE Transactions on Instrumentation and Measurement*, Vol. 59 (2010) No. 7, p.1874.
- [27] Information on https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/.
- [28] S.P. Zhu, J.C. Fang, R. Zhou, J.H. Zhao, W.B. Yu: A new noncontact flatness measuring system of large 2-D flat workpiece, *IEEE Transactions on Instrumentation and Measurement*, Vol. 57 (2008) No. 12, p.2891.
- [29] L. Zhao, L. Zhang, S. Ma, et al: Fast mode decision algorithm for intra prediction in HEVC, *2011 IEEE Visual Communications and Image Processing (VCIP)* (2011) p.1.