A Bandwidth-Adjustable Motion Estimation Design for Portable Video Coding

Jui-Hung Hsieh^a, King-Chu Hung^a, and Hung-Ren Wang^a

^aDepartment of Computer and Communication Engineering,
National Kaohsiung First University of Science and Technology,
No.1, University Rd., Yanchao Dist., Kaohsiung City 824, Taiwan

*Corresponding Author: jhhsieh@ncku.edu.tw

Abstract

Motion estimation (ME) is the most computation- and power-demanding module of portable video coding owing to the numerous calculations of rate-distortion (R-D) optimization operation and a large number of frame memory access. Nevertheless, how to minimize the power consumption of this component is the main research topic with regard to the H.264/AVC video coding standard. More severe problems arose in the latest video coding standard, called high-efficiency video coding (HEVC) because of the more R-D calculations and frame data access owing to utilization of larger block sizes and numerous block partitions for better coding bit-rate reduction than H.264/AVC. Accordingly, in this paper, we introduce a bandwidth-adjustable ME design, which is capability of not only awareness of bus bandwidth change but lower memory access times. Compared with the HEVC test model (HM-16.6), the proposed algorithm can save memory access under the common test conditions with negligible rate-distortion degradation.

Keywords: Mobile device, video coding, memory access, HEVC.

1. Introduction

Giving the rising trend of high-definition (HD) video services for mobile terminals, the previous video coding standard, called H.264/AVC⁽¹⁾, cannot meet the current coding requirement because coding bit-rate reduction is the most important design aims. The emerging video coding standard HEVC⁽²⁾ provides almost the 50% bit-rate reduction⁽³⁾ than H.264/AVC due to the abundant coding

tools. Conventional H.264/AVC adopts fixed 16×16 coding block, called macroblock (MB) coded with 16×16 , 16×8 , 8×16 , 8×8 , 8×4 , 4×8 , and 4×4 sub-block sizes to finish an entire frame coding. Oppositely, HEVC provides quadtree-based structure defined as coding tree unit (CTU) to code a frame, and the CTU can be split into small coding unit (CU) based on video coding contents. The CU size includes four depths of 64×64 (defined as depth 0), 32×32 (defined as depth 1), 16×16 (defined as depth 2), 8×8 (defined as depth 3). In theory, the larger CU settings dramatically make coding bit-rate reduction and coding efficiency enhancement. While CU settings chosen, the corresponding CU size can be further partitioned into small prediction unit (PU) for mode decision of inter or intra prediction. HEVC provides both symmetric and asymmetric PU partition size from 64×64 down to 4×4 . Owing to such a variety of CU block size and PU partition size settings, the coding complexity for mode decision of rate-distortion cost calculations in HEVC are more complicated than in H.264/AVC, especially on ME of inter prediction. For ME manipulation, HEVC provide some fast search schemes of coding tool to enhance the coding speed, such as test zonal (TZ)⁽⁴⁾ search and advanced motion vector prediction (AMVP). TZ search comprises with diamond search pattern and raster search to decrease computational complexity and speed up coding time. Another search scheme of ME was provided by adopting initial search center form origin to AMVP. This search method speedup coding time by reducing search points due to the minimum rate-distortion cost frequently around AMVP. Besides, some stage-of-the-art ME algorithms^(5,6) for HEVC in the literature are proposed and demonstrated the outstanding coding efficiency.

However, the video processor inside the portable devices nowadays is developed by employing with a smart power monitoring technique to achieve a low-power VLSI chip. The dynamic voltage-frequency scaling (DVFS) mechanism^(7,8) is widely employed for dynamically adjusting working frequency and supply voltage of video processor to meet the modern mobile system applications. The smart power management method in mobile devices gives us an important consideration, that is, for frame memory access in ME, the system provided an unfixed memory bandwidth rather than permanent bandwidth. Therefore, a bandwidth-adjustable ME algorithm, which monitors the available memory bandwidth and then adjusts the amount of memory access is an emerging design topics. In this paper, we are interested in developing a low-memory access ME algorithm and simultaneously awareness of memory bandwidth variation for HEVC.

The organization of this paper is as follows. In Section 2, we present our proposed framework for a bandwidth-adjustable ME design. Section 3 presents the simulation results and comparisons with the traditional approaches. Finally, Section 4 concludes this paper.

2. Proposed Framework

In this study, we intend to construct a bandwidth-adjustable ME framework that combines the video coding-quality and DVFS system operating conditions on machine-learning approach for HEVC. The major steps of this paper are summarized as follows. First, we will analyze how to monitor system operating frequency. Next, we will analyze the video coding performance relationship between the rate-distortion cost and memory bandwidth. After that, we will introduce how to train, predict, and analyze the video coding data within the ME design. Finally, we will establish a bandwidth-adjustable ME formulation with the applicable coding performance for real-time portable video applications. The whole machine-learning ME scheme is described as following subsection.

2.1 Operating Frequency Detection

In general, the operating performance points (OPPs)⁽⁹⁾ within the power-aware mobile applications can be predefined in design kickoff stage. Therefore, detecting the operating frequency is easy to accomplish fulfilling this way as⁽⁹⁾ under such systems.

2.2 Relationship of Video Coding Performance

The amount of reference frame access is a function of search range (SR)⁽¹⁰⁾ within ME operations; meanwhile, the SR is a function of memory bandwidth. Theoretically, the larger SR the ME, the excellent the coding quality it usually is. However, the rate-distortion cost⁽¹¹⁾ (RD Cost) is the mostly used criterion to judge the performance of video coding. The rate-distortion cost employed in this paper is defined as follows.

$$\text{RD Cost} = D + \lambda R \quad (1)$$

where D indicates the coding distortion, λ denotes the Lagrange multiplier, and R stands for the video motion data and the required coded bit length.

2.3 Video Coding Data Analysis

Data training is an important step for this study. We adopt a variety of class-C video sequences to train the proposed ME scheme. Consequently, we will obtain a non-linear formulation of RD cost and SR as the general form shown in (2) for the representative “BasketballDrill” video sequence. As for video prediction phase, we use the criteria of RD cost of temporally collocated PU in previous frame to predict the current coded PU because the RD cost of current coded PU is highly correlated with the previous collocated PU⁽¹²⁾. Finally, the coded PU needs to further classify into pre-defined classes. In this study, we adopt some thresholds in RD cost senses to classify the coded PU into three classes. The threshold selection is based on analysis and statistic of numerous offline training videos.

$$\text{RD Cost} = i \times SR^k + j \times SR^l \quad (2)$$

where i, j, k, l are training parameters obtained with a linear regression scheme⁽¹³⁾. Consequently, based on (2), the proposed scheme provides several classifications to allocation distinct bandwidth for video coding.

2.4 Constrained Formulation

The final step of our proposed bandwidth-adjustable ME scheme is to build up a relationship between RD cost and memory bandwidth in a constrained formulation. To optimally enhance the RD performance and reduce the number of memory access under different OPPs⁽⁹⁾, we solve

it with optimization method⁽¹³⁾.

3. Simulation Results

To evaluate the coding performance, we have integrated our proposed algorithm into HM-16.6⁽¹⁴⁾ to encode HEVC video. The common test condition⁽¹⁵⁾ for low-delay P main settings and HEVC test sequence are used to test the performance with one reference frame, full-search algorithm in ME, IPPP prediction sequence.

Tables I to II present the performance comparison using QP value 37 between the proposed algorithm and the anchor over test sequences of class C (832×480 -sized videos); in this comparison, the PSNR difference of luma pixels in dB (ΔY_PSNR) and the bit-rate (BR) change of luma pixels in percentage (ΔY_BR) are computed as (5) and (6), respectively.

$$\Delta Y_PSNR \text{ (dB)} = Y_PSNR_{Ours} - Y_PSNR_{HM} \quad (5)$$

$$\Delta Y_BR \text{ (%) } = \frac{(Y_BR_{Ours} - Y_BR_{HM})}{Y_BR_{HM}} \times 100\% \quad (6)$$

where Y_PSNR_{HM} and Y_BR_{HM} respectively denote the PSNR and BR of luma samples obtained from HM-16, and Y_PSNR_{Ours} and Y_BR_{Ours} respectively stand for the PSNR and BR of luma samples in our proposed ME scheme.

According to the experimental results, our proposed algorithm can attain similar coding performance of PSNR and BR with HM under different OPPs⁽⁹⁾. In Tables, the settings of OPP “3”, “2”, and “1” are respectively analogous to distinct target memory bandwidth. The proposed algorithm with regards to the class C test sequence can retrench memory bandwidth access under different OPPs with only 0.002dB to 0.005db PSNR degradation and average 0.25% BR increase of luma samples. Consequently, these results demonstrate that our proposed framework is applicable to ME operations for HEVC while exhibiting negligible quality loss; meanwhile, we anticipate the more memory bandwidth-savings will be achieved because of our optimal bandwidth allocation scheme.

4. Conclusions

In this paper, a bandwidth-adjustable ME design with capability of low memory access for HEVC was proposed. The presented ME framework exhibits that the enigma of ME coding performance subject to the system operating conditions is an optimization problem. The simulation results demonstrate that the proposed scheme based on machine-learning approach can reduce reference frame memory access; in addition, it has robust capability for preserving coding quality and diminishing reference frame memory activities in power-aware system conditions.

Acknowledgment

This work was supported by the Ministry of Science and Technology of the Republic of China under Contract MOST 104-2221-E-327-029.

References

- (1) T. Wiegand, G. J. Sullivan, G. Bjontegaad, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-575, Jul. 2003.
- (2) G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.
- (3) J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and 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, no. 12, pp. 1669-1682, Dec. 2012.
- (4) A. M. Tourapis, O. C. Au, and M. L. Liou, “New results on zonal based motion estimation algorithm -advanced predictive diamond zonal search,” in *Proc. IEEE ISCAS*, vol. 5, May 2001, pp.183-186.
- (5) N. Hu and E. H. Yang, “Fast motion estimation based on confidence interval,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1310-1322, Aug. 2014.
- (6) J. Vanne, M. Viitanen, and T. D. Hämäläinen, “Efficient mode decision schemes for HEVC inter prediction” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 9, pp. 1579-1593, Sep. 2014.
- (7) V. Hanumaiah, S. Vruthula, and K. S. Chatha,

- “Performance optimal online DVFS and task migration techniques for thermally constrained multi-core processors,” *IEEE Trans. Comput.-Aided Des.*, vol. 30, no. 11, pp. 1677-1690, Nov. 2011.
- (8) A. Mutapcic, S. Boyd, S. Murali, D. Atienza, G. D. Micheli, and R. Gupta, “Processor speed control with thermal constraints,” *IEEE Trans. Circuits Syst. I Reg. Papers*, vol. 56, no. 9, pp. 1994-2008, Sept. 2009.
- (9) “Power-management technique for OMAP35x applications processors,” TEXAS INSTRUMENTS Inc., TI. (2008). [Online]. Available: <http://www.ti.com/general/docs/lit/getliterature.tsp?baseLiteratureNumber=sprt495&fileType=pdf>.
- (10) H. Shim and C. M. Kyung, “Selective search area reuse algorithm for low external memory access motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 1044-1050, Jul. 2009.
- (11) G. J. Sullivan and T. Wiegand, “Rate-distortion optimization for video compression,” *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74-90, Nov. 1998.
- (12) K. I. Lee, A. C. Tsai, J. F. Wang, and F. F. Yang, “Efficient inter mode prediction based on model selection and rate feedback for H.264/AVC,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 6, pp. 708-716, Jun. 2011.
- (13) S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2014.
- (14) JCT-VC, HM 16.6, Reference Software. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6/
- (15) F. Bossen, *Common Test Conditions and Software Reference Configurations, Document JCTVC-K1100*, Oct, 2013.

Table 1. Performance comparison of ΔY_{PSNR} under distinct OPPs.

Size	Sequence	OPP		
		3	2	1
Class C (832x480)	RaceHorses	-0.002	-0.002	0.002
	BQMall	-0.005	-0.005	-0.003
	PartyScene	-0.004	-0.004	-0.001
	BasketballDrill	-0.009	-0.004	-0.006
	Average	-0.005	-0.004	-0.002

Table 2. Performance comparison of ΔY_{BR} under distinct OPPs.

Size	Sequence	OPP		
		3	2	1
Class C (832x480)	RaceHorses	1.0005	0.9984	1.0007
	BQMall	-0.0003	0.0004	-0.0003
	PartyScene	-0.0002	-0.0006	-0.0006
	BasketballDrill	0.0007	5E-05	0.0010
	Average	0.2502	0.2496	0.2502