Improved Hard-Decision Quantization with Decision Tree for HEVC Video Compression

Data Compression Conference

Motong Xu and Byeungwoo Jeon
Dept. of Electrical and Computer Engineering, Sungkyunkwan University, Korea.

교육과학기술부 국가지정연구실 National Research Lab
Digital Media Lab, Sungkyunkwan University, Korea

Introduction

▪Motivation

▪The soft-decision quantization (SDQ) achieves excellent coding gain but with high computational expense.

▪In general, SDQ is relatively not so friendly to hardware implementations.

▪Our approach

We design an improved HDQ scheme by analyzing the behavior of rate-distortion optimized quantization (RDOQ) [1] technique for High Efficiency Video Coding (HEVC).

Problem Formulation

▪RDOQ behavior in HEVC

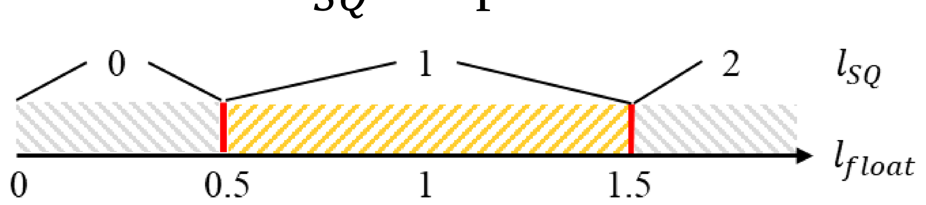
▪An initial uniform scalar quantization (SQ) level is firstly calculated as

$$l_{SQ} = \left\lfloor \frac{|c|}{\Delta q} + \theta \right\rfloor$$

c: unquantized transform coefficient
Δq: quantization step size
$\theta = 1/2$: quantization round offset

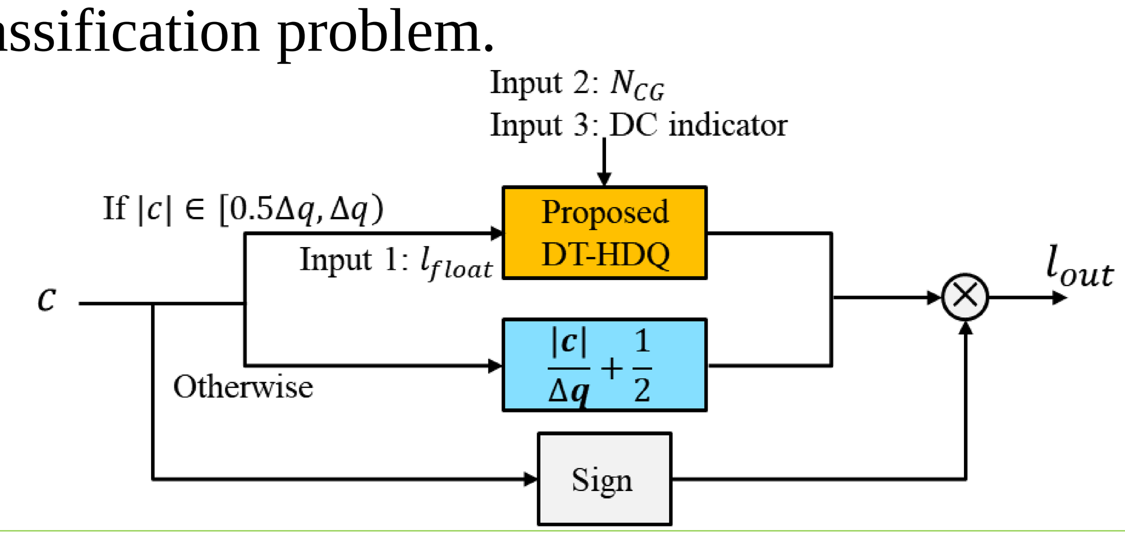
▪We denote the floating form of the quantization level as $l_{float} = |c|/\Delta q$.

▪RDOQ mostly achieves coding gain over SQ when a nonzero l_{SQ} is optimized to zero especially when the value of l_{SQ} is optimized from 1 to 0.



▪We also observe through statistical analysis that RDOQ makes level change mainly when $l_{float} \in [0.5,1)$.

Therefore, this work aims at designing **a decision tree-based HDQ (DT-HDQ) scheme** only for transform coefficient values belonging in this interval which can be modeled as a binary classification problem.



Selection of Inputs

It is essential to choose the appropriate input information given in the learning process to help building the desired HDQ model. Several statistical analyses on RDOQ behavior are carried out for selecting the input information.

▪Input 1: l_{float} value (calculated from $|c|$)

▪The floating form of scalar quantized level.

▪Input 2: N_{CG}

▪Total number of |c| values equal to or larger than 0.5Δq in a 4×4 coefficient group play.

▪Input 3: DC indicator

▪The DC coefficient has a lower possibility to be changed by RDOQ since it represents low frequency information of the current transform block.

▪It is found also useful to consider whether a coefficient is at DC location or not as an additional feature.

Inputs	Values	DT-HDQ output
l_{float} value	[0.5,1)	0 or 1
N_{CG}	0 ~ 16	
DC indicator	0 or 1	

Decision Tree-based HDQ Model

▪A binary decision tree-based HDQ (DT-HDQ) model is trained to mimic the quantized level decisions in the RDOQ process with the three selected input information.

▪Therefore, the final quantization output with the proposed DT-HDQ is obtained as

$$l_{out} = \begin{cases} l_{DT-HDQ}, & if\ c \in [0.5\Delta q, \Delta q) \\ l_{SQ} = \left\lfloor \frac{|c|}{\Delta q} + \frac{1}{2} \right\rfloor, & otherwise \end{cases}$$

l_{DT-HDQ}: the predicted quantized level by the proposed DT-HDQ model.

▪The proposed DT-HDQ is implemented on top of the HEVC reference software [3] by partially substituting the scalar dead-zone quantizer.

This work is supported in part by the National Research Foundation of Korea (NRF) grant (2017R1A2B2006518) and by the Grand Information Technology Research Center support program (IITP-2017-2015-0-00742) supervised by the IITP (Institute for Information & Communications Technology Promotion), both funded by the Ministry of Science and ICT.

Experimental Results

▪Experiment condition

▪Test software: HEVC reference software HM 16.15 [3] with RDOQ tool disabled.

▪Test sequences: totally 20 sequences from the JVET common test conditions [4].

▪QP value 22, 27, 32, and 37 under RA-Main encoding configuration.

▪Anchor: The original HM 16.15 [3] with RDOQ tool turned off.

▪Performance measurement

▪BDBR and quantization time increment $t_{quant} = \frac{T^{*} - T_{anchor}}{T_{anchor}} \times 100$.

▪T^{*}: quantization time of the method to be evaluated; T_{anchor}: quantization time of anchor.

Table 3: Performance comparison of RDOQ and the proposed method against Anchor

Sequences		RDOQ		Quantization with DT-HDQ	
		BDBR (%)	t_{quant} (%)	BDBR (%)	t_{quant} (%)
Class B	Kimono	-9.07	122.59	-6.07	47.68
	ParkScene	-6.12	148.24	-3.54	50.40
	Cactus	-8.90	169.34	-5.54	52.95
	BasketballDrive	-8.25	170.96	-4.75	52.74
	BQTerrace	-7.48	166.07	-3.56	53.21
Class C	BasketballDrill	-2.36	178.23	-1.14	54.93
	BQMall	-5.37	168.96	-3.47	53.40
	PartyScene	-4.23	231.34	-2.39	60.32
	RaceHorses	-5.51	223.38	-3.20	62.44
Class D	BasketballPass	-4.96	210.28	-2.98	59.96
	BQSquare	-5.71	218.84	-4.47	58.08
	BlowingBubbles	-4.25	227.59	-2.26	60.76
	RaceHorses	-4.74	234.14	-2.82	64.23
Class E	FourPeople	-4.13	105.16	-2.14	45.01
	Johnny	-4.83	97.89	-2.52	43.60
	KristenAndSara	-5.31	100.07	-3.02	44.28
Class F	BasketballDrillTest	-2.57	182.67	-1.34	55.39
	ChinaSpeed	-4.53	176.60	-3.00	57.55
	SlideEditing	-3.20	93.75	-2.22	42.71
	SlideShow	-2.70	107.18	-1.71	44.81
Average		**-5.21**	**166.71**	**-3.11**	**53.22**

The proposed DT-HDQ scheme is more meaningful for practical usage to achieve approximately 60% of the RDOQ coding efficiency with only 30% of its additional processing time complexity.

Conclusion

This paper proposed a decision tree-based HDQ scheme for video coding to improve the coding efficiency of the conventional HDQ structure. The proposed quantization scheme with DT-HDQ is reported to provide an average of 3.11% bit-rate saving in a BDBR sense without complex RD cost estimation.

[1] M. Karczewicz, et al., Rate Distortion Optimized Quantization, Q.6, document ITU-T SG16, VCEG-AH21, January 2008.
[2] H. Wang, S. Yu, Y. Zhang, Z. Kuang and L. Yu, "Hard-Decision Quantization Algorithm Based on Deep Learning in Intra Video Coding," in Proc. Data Compression Conference (DCC), March 2019.
[3] High Efficiency Video Coding Test Model Software, Version 16.15. [Online] Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.15.
[4] F. Bossen, "Common HM Test Conditions and Software Reference Configuration," Joint Collaborative Team on Video Coding, document JCTVC-L1100, July 2012.