2013년 6월 전자공학회 논문지 제 50 권 제 6 호
Journal of The Institute of Electronics Engineers of Korea Vol. 50, NO. 6, June 2013 http://dx.doi.org/10.5573/ieek.2013.50.6.260

논문 2012-50-6-30

Fast Enhancement Layer Encoding Method using CU Depth Correlation between Adjacent Layers for SHVC

Kyeonghye Kim, Seonoh Lee, Yongjo Ahn, and Donggyu Sim©

Abstract

This paper proposes a fast enhancement layer coding method to reduce computational complexity for Scalable HEVC (SHVC) which is based on High Efficiency Video Coding (HEVC). The proposed methoddecreases encoding time by simplifying Rate Distortion Optimization (RDO)for enhancement layers (EL). The simplification is achieved by restricting CU depths based on the correlation of coding unit (CU) depths between adjacent layers and scalability (spatial or quality) of EL. Comparing with the performance of SHM 1.0 software encoder, the proposed method reduces the encoding time by up to 31.5%.

Keywords: Scalable HEVC, HEVC, Fast encoder

Ⅰ.Introduction

Withneeds for high quality videos and recent advances of display techniques, high resolution and high quality services such as 4K and 8KUHDTV are emerging.To meet such requirements, ITU-T video coding experts group (VCEG) and ISO/IEC moving picture experts group (MPEG) have founded Joint Collaborative Team on Video Coding (JCT-VC), which has standardized HEVC[1]. On January 2013, HEVC final draft international standard (FDIS)[2] has led discussions concerning HEVC extensions such as three dimensional multi-view and scalable video coding standard.One of the extensions, SHVC has been standardized from 10th JCT-VC meeting on July 2012[3]. SHVC supports Ultra High Definition (UHD) video encoding, which was not supported in Scalable Video Codec (SVC) based on H.264/AVC,as well as scalable coding on spatial, temporal, and quality domains.Furthermore, SHVC is expected to support satellite broadcasting, digital cable TV (DCATV), and internet protocol TV (IPTV) in the coming years.The most critical issue of SHVC is high computational complexity derived from not only multi-layer encoding but also inter-layer prediction tools, especially in EL. This paper proposes a fast encoding method for SHVC using inter-layer CU depth correlation.This paper is organized as follows. Section Ⅱ presents SHVC standard encoder structure. In Section Ⅲ, an SHVC CU depth restriction method for fast enhancement layer encoding is proposed. Section Ⅳ shows performance evaluation of the proposed method and Section Ⅴ

* Dept. of Computer Engineering, Kwangwoon University
※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (NRF-2010-0025325)
© Corresponding Author(E-mail: dgsim@kw.ac.kr)

concludes the presented work with further research topics.

Ⅱ. Scalable HEVC

SHVC is an extension of HEVC to support technical scalabilities on temporal, spatial, and quality domains, which enables various multiple multimedia services for heterogeneous environment and conditions. SHVC supports inter-layer prediction tools to exploit redundancy between adjacent layers. Figure 1 illustrates a block diagram of a typical SHVC encoderwith inter-layer prediction tools, for double-layered video sequences.The base layer (BL) indicatesa video sequence to be codedin the lowest scale. For sake of the backward compatibility with HEVC, the BL can be encoded and decoded bya HEVC codec. EL is a video sequence to be coded in an enhanced spatial, temporal or quality scales. Not only encoding multi-layer sequences, but also supporting additive tools using inter-layer correlations, computational complexity of the SHVC encoder is much higher than single layer video encoder such as HEVC.

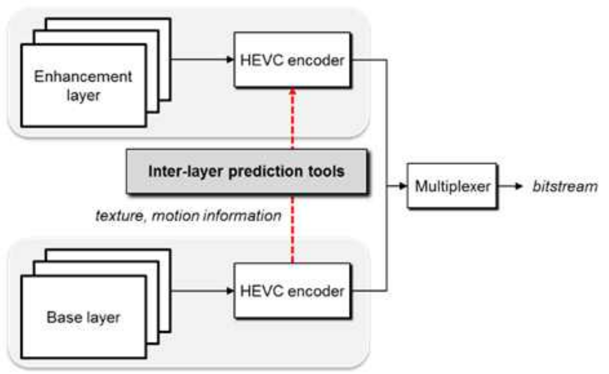

Fig. 1. Block diagram of SHVC.

2. 1. TextureRL

In SHVC, EL can be encoded using texture, motion information of reference layer (RL). In a double-layered video sequence, RL is equivalent to BL as there is only one EL.Inter-layer texture prediction can be conducted with decoded reference layers and TextureRL is one of the inter-layer texture predictiontools which is applied to SHVC.In TextureRL, the EL encoder usestexture block information in RL as a predictor. In spatial scalable encoding cases, texturein RL is up-sampled to form a predictor in EL for resolution consensus.

2. 2 Rate distortion optimization (RDO) process for ELin SHVC

In typical RDO processes, the optimal prediction mode for each block or coding unitis determined by considering not only bitrate but also distortion. Because encoder conducts prediction for all possible block sizes and prediction modes during RDO process, general RDO process occupiesalmost of encoding time up to 90%[5]. The SHVC encoder determines the optimal size of CUs during RDO process[4]. RDO process for EL with TextureRL is presented, as shown in Figure 2. RDO is conducted on all of the permitted CU sizes. The SHVC encoder determines optimal sizes and prediction modes of CUs in terms of rate-distortion optimization (RDO). The SHVC encoder conducts not only single layer video encoding for two or more layers but also inter-layer prediction tools, which make the computational complexity of the SHVC encoder much more complex

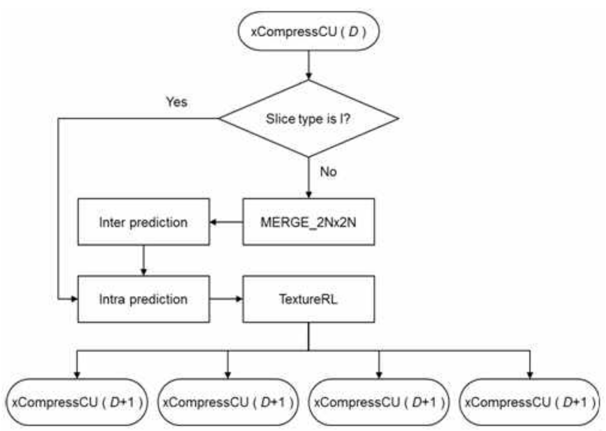

Fig. 2. CU-level RDO process in EL.

than that of HEVC. Research on the fast encoding method for RDO process of SHVC is required for real-time streaming services based on SHVC.

3. Proposed SHVC fast encoding method

For scalable video coding based on HEVC, a CU depth restriction method for fast EL encoding is proposed in this paper. Because an input image on each layer is acquired from the same scene, correlation between adjacent video layers is very high.In Table 1, statistical distribution ofthe depth of CUs in RLwith respect tothedepth of the co-located CU in EL is presented. Common test conditions for SHVC have been used to obtain data in Table 1[6]. DEL indicates the depthof CU in EL and DRL means the depthof co-located CU in RL. In the middle row in Table 1, labeled DRL, probabilities (in percentage) of DEL equals to DRL is shown and the other rows labeled DRL+n show probabilities of DEL equals to DRL+n. Note that about 87% of DEL has a value ranging from DRL−1 to DRL+1 in spatial scalability. In case of quality scalability, about 89% of DEL has a value ranging from DRL to DRL+2.

According to the statistics on similarity of the CU depth between EL and RL, CU-level RDO for EL is

Table 1. Match probability of the CU depth between layers.

D_{EL}	Random access			Low delay P		
	2x	1.5x	Quality	2x	1.5x	Quality
D_{RL}-3	0.12%	0.22%	0.03%	0.09%	0.17%	0.01%
D_{RL}-2	2.33%	2.33%	0.65%	2.30%	2.16%	0.47%
D_{RL}-1	15.87%	15.96%	4.61%	16.27%	15.37%	4.08%
D_{RL}	42.86%	35.07%	26.75%	42.10%	34.43%	23.25%
D_{RL}+1	27.50%	28.76%	35.43%	29.07%	32.09%	37.40%
D_{RL}+2	10.18%	10.92%	26.86%	9.09%	11.94%	28.35%
D_{RL}+3	1.14%	6.75%	5.67%	1.09%	3.84%	6.44%

```
BEGIN
Spatial ratio = Width_EL/Width_BL
 IF Spatial ratio == 2 or 1.5 THEN
Pred_min=D_RL-1
Pred_max=D_RL+1
 ELSE IF Spatial ratio == 1 THEN
Pred_min=D_RL
Pred_max=D_RL+2
 ENDIF
 IF Pred_min<0THENPred_min=0ENDIF
 IF Pred_max>D_maxTHENPred_max=D_maxENDIF
END
```

Fig. 3. Pseudo code of determining the CU depth rangeforRDO process of EL.

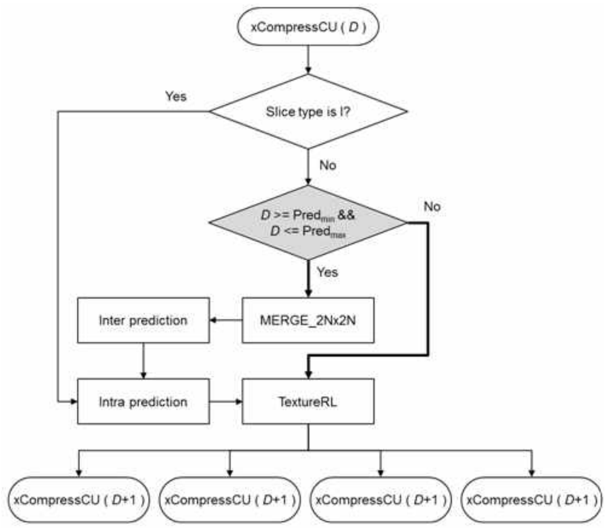

Fig. 4. Proposed fast RDO process.

proposed to reduce computational complexity in this paper. The proposed method restricts the CU depth in EL by referring to the CU depth of co-locatedCU in RL. Figure 3 describes a pseudo code for theCU depth range decision process in EL based on thedepth of co-located CU in RL.Note that Predmin is the minimum CU depth and Predmax is the maximum CU depth for RDO process. The maximum and the minimum CU depth of 2xand 1.5x spatial scalable encoding are DBL-1and DBL+1, respectively. In addition, those of quality scalable encoding are DBLand DBL+2. Figure 4 presents the proposed fast encoding process with restricted CU depth.To prevent degradation of the coding performance, the proposed method is not applied to I-slices and only TextureRL prediction is conducted for CU depths outside of permitted CU depths, instead of none.

4. Experiment results

To evaluate performance of the proposed algorithm, averaging time saving of the proposed method against SHM-1.0 reference software was computed with SHVC common test conditions[6]. SHM-1.0 encoder contains TextureRL mode as an inter-layer prediction tool.

Table 2. Test conditions for 2x and 1.5x spatial scalability.

Seq.	Scalability	Resolution		QP	
		BL	EL	QP_{BL}	QP_{EL}
Class A	2x spatial scalability	1280×800	2560×1600	22 26 30 34	QP_{BL}+0 QP_{BL}+2
Class B	2x and 1.5x spatial scalability	960×540 1280×720	1920×1080		

Table 3. est conditions for Quality scalability.

Seq.	Scalability	Resolution		QP	
		BL	EL	QP_{BL}	QP_{EL}
Class A	Quality scalability	2560×1600	2560×1600	26 30 34 38	QP_{BL}-6 QP_{BL}-4
Class B		1920×1080	1920×1080		

Table 4. Performance of the proposed method for random access.

	2x			1.5x			Quality		
	Y	U	V	Y	U	V	Y	U	V
Class A	0.3%	0.1%	0.2%				0.3%	0.3%	0.3%
Class B	0.9%	0.6%	0.6%	0.3%	0.1%	0.1%	0.3%	0.2%	0.2%
Overall	0.7%	0.5%	0.5%	0.3%	0.1%	0.1%	0.3%	0.2%	0.2%
Encoding Time	**68.5%**			**71.0%**			**80.6%**		

Table 2 and 3 present test conditions for spatial and quality scalable encoding tests, respectively. Encoding configurations contain random access (RA) and low delay P (LD-P) profiles. Performance of the proposed for RA profile is presented in Table 4. Encoding time is a ratio calculated on the basis of encoding time of SHM-1.0.Encoding time is reduced by 26.6% on average with 0.4% average BD-rate increase. Performance of the proposed for LD-P profile is shown in Table 5. On average, encoding time saving is 25.4% with 0.6% BD-rate increase.

5. Conclusion

Fast encoding algorithm for scalable video coding based on HEVC is proposed in this paper. The proposed method reduces EL encoding time by restricting CU depth range for RDO with depth of co-locate CU in RL.Average time saving of the proposed method is up to 31.5% with 0.5% BD-rate increase.

Table 5. Performance of the proposed methodfor low-delay P.

	2x			1.5x			Quality		
	Y	U	V	Y	U	V	Y	U	V
Class A	0.4%	0.4%	0.4%				0.4%	0.6%	0.6%
Class B	0.9%	1.0%	0.8%	0.5%	0.2%	0.2%	0.7%	0.2%	0.0%
Overall	0.8%	0.8%	0.6%	0.5%	0.2%	0.2%	0.6%	0.3%	0.2%
Encoding Time	**71.2%**			**74.6%**			**81.5%**		

REFERENCES

[1] G. J. Sullivan, J. Ohm, Woo-Jin Han, and T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Transactions On Circuits and Systems for Video Technology 22, 12 (2012)

[2] B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand, High efficiency video coding (HEVC) test specification draft 10 (for FIDS & consent), 12thJCT-VC meeting,

(2013)Jan. 14-23; Geneva, CH
[3] J. Chen, J. Boyce, Y. Ye, and M. Hannuksela, SHVC draft text 1, 12^{th} JCT-VC meeting, (2013)Jan.14-23; Geneva, CH
[4] Il-Koo Kim, Junghye Min, Tammy Lee, Woo-Jin Han, and JeongHoon Park, Block partitioning structure in the HEVC standard, IEEE Transactions On Circuits and Systems for Video Technology 22, 12 (2012)
[5] Yongjo Ahn, Taejin Hwang, Sungeun Yoo, Woo-Jin Han, and Donggyu Sim, Statistical characteristics and complexity analysis of HEVC encoder software, Journal of Broadcasting & Electronic Midea 17, 6 (2012)
[6] X. Li, J. Boyce, P. Onno, and Y. Ye, Common test conditions and software reference configurations for the scalable test model, 12^{th}JCT-VC meeting, (2013)Jan.14-23; Geneva, CH