JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 18, No. 1, pp. 141-147, Feb. 2017

스케일러블 HEVC 부호화 효율 개선을 위한 계층 간 적응적 필터 선택 알고리즘

이 종 혁[1] · 김 병 규[2*]
[1]숙명여자대학교 빅데이터활용연구센터
[2]숙명여자대학교 IT공학과

Adaptive Inter-layer Filter Selection Mechanism for Improved Scalable Extensions of High Efficiency Video Coding (SHVC)

Jong-Hyeok Lee[1] · Byung-Gyu Kim[2*]
[1]Big data Utilization Research Center, Sookmyung Woman's University, Seoul 04310, Korea
[2]*Dept. of IT engineering, Sookmyung Woman's University, Seoul 04310, Korea

[요 약]

스케일러블 HEVC에서 상위계층의 계층 간 예측에서 기본계층의 부호화 잔차 영상에 대한 업샘플링 된 결과를 참조하여 예측하게 된다. 본 논문에서는 고효율 영상 압축 기반 스케일러블 부호화 (Scalable Extension of High Efficiency Video Coding)에서 상위계층 잔차 데이터 예측에 대한 개선 기법을 제안한다. 제안하는 적응적 필터 선택 기법은 스무싱 필터와 샤프닝 필터를 사용함으로써 계층 간 예측 방법에서 효율을 향상시킨다. 기존의 업샘플링 필터와 두 개의 필터를 추가하여 율-왜곡 비용함수 기반의 경쟁기법을 통한 계층 간 예측 알고리즘을 SHVC 5.0에 구현함으로써 Y, U, V 컴포넌트에 대한 평균 1.5%, 2.1%, 1.7%의 BD-rate 향상을 보여준다.

[Abstract]

Scalable extension of High Efficiency Video Coding (SHVC) standard uses the up-sampled residual data from the base layer to make a residual data in the enhancement layer. This paper describes an efficient algorithm for improving coding gain by using the filtered residual signal of base layer in the Scalable extension of High Efficiency Video Coding (SHVC). The proposed adaptive filter selection mechanism uses the smoothing and sharpening filters to enhance the quality of inter-layer prediction. Based on two filters and the existing up-sampling filter, a rate-distortion (RD)-cost fuction–based competitive scheme is proposed to get better quality of video. Experimental results showed that average BD-rate gains of 1.5%, 2.1%, and 1.7% for Y, U and V components, respectively, were achieved, compared with SHVC reference software 5.0, which is based on HEVC reference model (HM) 13.

색인어 : 스케일러블 부호화, HEVC 기반 스케일러블 부호화, 계층 간 예측, 적응적 필터링

Key word : SHVC, Scalable Extension of HEVC, Inter-layer prediction, Adaptive filtering

http://dx.doi.org/10.9728/dcs.2017.18.1.141

Received 21 December 2016; **Revised** 22 February 2017
Accepted 25 February 2017

***Corresponding Author; Byung-Gyu Kim**

Tel: +82-02-2077-7293
E-mail: bg.kim@sm.ac.kr

Ⅰ. Introduction

As display technology is grown up, user and market demand high quality content of video with high bit-depth and resolution like 4K and 8K ultra high definition (UHD), and various color format. 4K UHD content has huge data amount as 4x more than full high definition (Full HD), and 3D video content.

A next generation standardization of video coding is material to solve problems of data rate increased and restricted bandwidth in network transmission.

Therefore, A newest video compression standard has been developed by Joint Collaborative Team on Video Coding (JCT-VC) at January 2013, called high efficiency video coding (HEVC) [1].

The HEVC is a state of the art video compression standard that provides a bitrate reduction in the range of 50%, compared with the previous Advanced Video Coding (H.264/AVC) standard with a similar video quality [2, 3]. The HEVC has advanced extensions of scalable, range, 3D video. Range extension has been added to support high bit depths and rich color format [4]. To service 3D video content efficiently, 3D-HEVC has been researched with various stereo matching algorithm [5].

Scalable video coding (SVC) scheme supports different terminal and network environments in a single bit stream. SVC decodes only the highest layer while the base layer(BL) is used for single loop decoding [6, 7]. The method for decoding only the minimum information necessary for inter-layer prediction has reduced the calculation load and memory usage, but has increased the complexity of implementation. To remove the redundancy between layers in the SVC, bitrate reduction methods had been proposed [8, 9].

Recently, scalable extension of HEVC (SHVC) [10] has adopted an easily implemented multi-loop coding framework to possibly re-use the existing HEVC codec design. The low level process of each enhancement layer (EL) of the SHVC codec is kept the same as a single layer HEVC codec, and only high level syntax changes are applied at the enhancement layer for inter-layer processing and operation point signaling. The high level syntax architecture adopted by SHVC allows implementation for maximal reuse of the existing HEVC design. Inter-layer processing of reconstructed reference layer pictures allows inter-layer texture prediction as a so-called “reference index” approach [11] for inter-layer motion prediction [12] that is applied to improve the coding efficiency of enhancement layers.

SHVC is the scalable extension of HEVC, which provides traditional scalability options in terms of quality, spatial resolution and temporal frame rate and newer scalability options as well. SHVC can be used to deliver UHD or 4K resolution video content to mix of clients having varying characteristics.

We propose an adaptive filter selection scheme for the inter-layer prediction process to improve image quality. To make the better residual data in the enhancement layer, smoothing filter and sharpening filter are applied for video processing in inter-layer prediction.

The remainder of this paper is organized as follows: In section 2, an inter-layer adaptive filter selection technique is described and the coding performance of the proposed method is presented in section 3. Finally, section 4 presents conclusions.

Ⅱ. Inter-layer adaptive filter selection technique

The objective of the SVC standardization has been enable the encoding of a high-quality video bitstream to support various resolution/frame-rate/quality by single bitstream. It can themselves be decoded with a complexity and reconstruction quality similar to that achieved using the existing H.264/AVC design with the same quantity of data as in the subset bit stream.

The SVC can be reduced computation and memory usage, but complexity of implementation has been increased. It has been used with restriction because the SVC cannot be optimized of design structure. In SHVC, implementation complexity can be reduced and increased about encoding efficiency by contributed multi-loop encoding structure.

(Figure 1) shows transcoding for non-scalable video with individual transcoding for each case. Transcoding for Scalable video is indicated in (Figure 2).

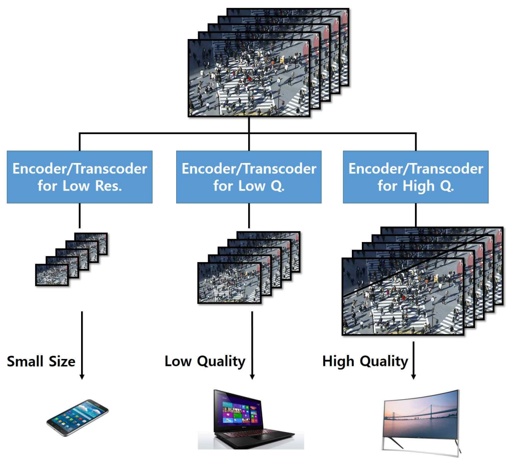

그림 1. 비스케일러블 비디오 데이터의 트랜스코딩
Fig. 1. Transcoding for non-scalable video

(Figure 3) shows a block diagram of inter-layer prediction in the SHVC. First, original video source is input and down-sampling filter is performed to be used in base layer coding. After base layer coding, reconstructed result that up-sampling filter is processed is gone to inter-layer prediction in enhancement layer.

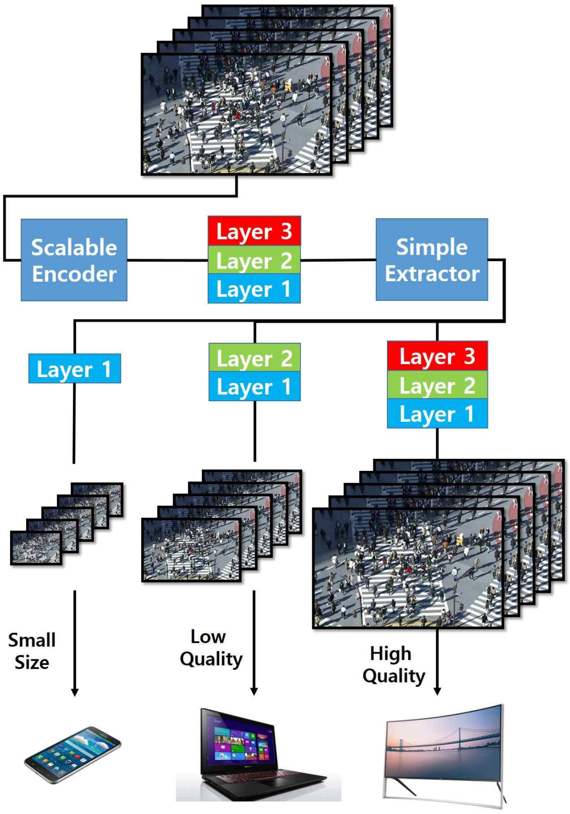

그림 2. 스케일러블 비디오에 대한 트랜스코딩
Fig. 2. Transcoding for scalable video

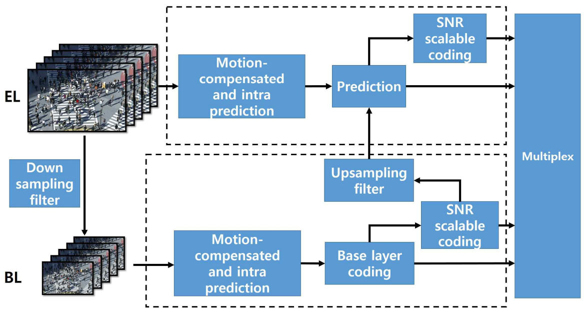

그림 3. SHVC에서의 계층 간 예측 블럭도
Fig. 3. Block diagram of inter-layer prediction in the SHVC

By using the down-sampled image with loss in base layer coding, the prediction error is generated in inter-layer prediction of enhancement layer. To improve this error, various filters are attempted as smoothing and sharpening filters with up-sampling filter.

2-1 Smoothing filter

Similar to scalable extension of H.264/AVC, the residual signal of reference layer can be used to refine the residual signal of the enhancement layer. The prediction error that is presented in the enhancement layer (EL) can be reduced by adding the residual signal of the reference layer. To reduce these error, smoothed inter-layer residual prediction (SILRP) mode has been proposed [13].

In the up-sampling of the residual of reference layer, it is applied transform unit boundaries and can disturb the signal components. By using smoothing filter, error can be reduced in the enhancement layer.

For inter-layer prediction in the EL, prediction of the reference picture uses an up-sampling filter. However, the proposed method selects a better value after comparison of the RD-cost of the up-sampling filter and a smoothing filter. A selected filter image is used as a reference image in the EL. Smoothing filter coefficients are defined as 1, 2, and 1.

2-2 Sharpening filter

The picture with low resolution has been up-sampled from the base-layer. It does not make efficient of inter-layer prediction. To increase the amount of high frequency details near the edges, a sharpening filter has been proposed [14].

Quantization in video coding using a quantization parameter (QP) value is the most important factor for compression control, which controls the compression efficiency. However, quantization is lost with use of compression methods. A high QP value in video coding results in quantization distortion generated at object edges, so a preprocessing sharpening filter is used in order to minimize quantization distortion. We used the Prewitt filter for sharpening. The Prewitt Edge filter is use for detection of edges based on application of horizontal and vertical filters in sequence. Both filters are applied to the image and summed to form a final result. The two filters are basic convolution filters of the form:

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} * I, G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix} * I \quad (1)$$

A smoothing filter, an up-sampling filter, and a sharpening filter were designed for filter selection based on comparison of the RD-cost function.

2-3 Adaptive filter selection scheme

A combination of the three filters which are described as up-sampling, smoothing, and sharpening filter, results in the adaptive filter selection method. (Figure 4) shows the overall flow chart of the proposed adaptive filter selection scheme in the SHVC.

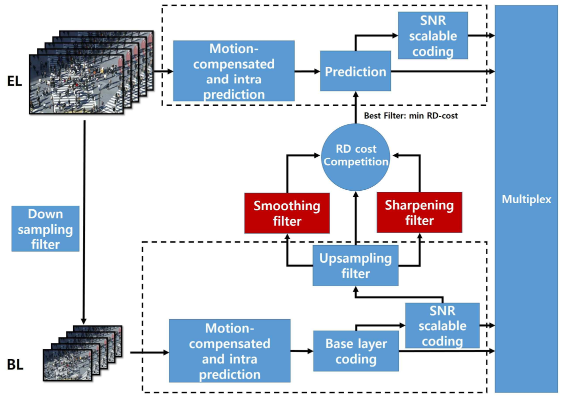

그림 4. SHVC에서 적응적 필터 선택 기법의 전체 흐름도
Fig. 4. Overall flow chart of adaptive filter selection scheme in the SHVC

According to the complexity of texture in each frame or coding block, the characteristics for encoding is usually changed. To cope with this characteristics, we propose an adaptive filter selection technique with competition mode.

Using the up-sampled data of the given coding block, each filter is used to calculate the RD-cost in terms of bitrate and distortion values. Then the best filter which has the lowest RD-cost, is selected to encode the given block by using RD-cost competition process.

Ⅲ. Experimental Results

The proposed method is implemented based on SHVC reference model (SHM) 5.0, which is based on HEVC reference model (HM) 13. Evaluation occurred under a comprehensive set of simulations as defined by SHVC common test conditions [15] of low delay-P and random access. Standard sequences with 1 sequence in class A (Resolution : 2650x1600) and 3 sequences in class B (Resolution : 1920x1080) with spatial 2x scalability were used. Four base layer QP values of 22, 26, 30, and 34 were used with 100 total frames in each test sequence.

<Table 1> shows results of applying only the smoothing filter. Test average BD-rate [16] saving values of 0.9%, 1.2%, and 1.4% were achieved for Y, U, and V, respectively. The BD-rate savings of the Class-B sequences was greater than for the Class-A sequence. Use of Class-B confirmed that different results were based on differences in sequences. A low-texture sequence was confirmed for good performance.

표 1. 계층 간 예측을 위한 스무딩 필터에 의한 성능 결과
Table. 1. Coding performance result of a smoothing filter for inter-layer prediction

Low Delay-P 2X (BD-rate)				
Sequences		Y	U	V
Class A	Traffic	-0.2%	-0.5%	-0.7%
Class B	Kimono	-1.2%	-1.4%	-2.3%
	ParkScene	-0.9%	-1.3%	-1.8%
	BQTerrace	-1.3%	-1.1%	-1.5%
Average		-0.9%	-1.2%	-1.4%

표 2. 계층 간 예측을 위한 샤프닝 필터에 의한 성능 결과
Table. 2. Coding performance result of a sharpening filter for inter-layer prediction

Low Delay-P 2X (BD-rate)				
Sequences		Y	U	V
Class A	Traffic	-1.9%	-2.1%	-1.7%
Class B	Kimono	-0.9%	-1.8%	-2.3%
	ParkScene	-0.7%	-0.8%	-1.2%
	BQTerrace	-0.5%	-1.6%	-1.4%
Average		-1.0%	-1.5%	-1.6%

표 3. 계층 간 예측을 위한 적응적 필터 선택 기법에 의한 성능 결과
Table. 3. Coding performance result of the adaptive filter selection for inter-layer prediction

Low Delay-P 2X (BD-rate)				
Sequences		Y	U	V
Class A	Traffic	-2.1%	-2.8%	-2.7%
Class B	Kimono	-1.4%	-2.2%	-1.9%
	ParkScene	-1.2%	-1.9%	-1.5%
	BQTerrace	-1.3%	-1.7%	-1.1%
Average		-1.5%	-2.1%	-1.7%

표 4. Random Access 2X 환경에서 계층 간 예측을 위한 적응적 필터 선택 기법에 의한 성능 결과
Table. 4. Coding performance result of the adaptive filter selection for inter-layer prediction in Random Access 2X

Random Access 2X (BD-rate)				
Sequences		Y	U	V
Class A	Traffic	-1.4%	-2.1%	-1.7%
Class B	Kimono	-0.7%	-1.1%	-2.1%
	ParkScene	-0.4%	-1.5%	-1.4%
	BQTerrace	-0.6%	-1.4%	-1.2%
Average		-0.7%	-1.5%	-1.6%

<Table 2> shows results of applying only a sharpening filter. In contrast to use of smoothing filters in high-resolution sequences, the BD-rate saving was high. The difference between classes was 1.0%. Test average BD-rate saving values of 1.0%, 1.5%, and 1.6% were achieved for Y, U, and V, respectively.

<Tables 3> and <Table 4> show results for application of a smoothing filter and a sharpening filter. The BD-rate savings value of low delay-P was greater than for Random Access. For low delay-P, BD-rate saving values of 1.5%, 2.1%, and 1.7% were achieved, with Random Access BD-rate saving values of 0.7%, 1.5%, 1.6%. There was a 0.8% BD-rate gap between low delay-P and random access. Results shown in <Table 3> exhibited differences in the BD-savings rate depending on the resolution shown in <Tables 1> and <Table 2>. However, results shown in <Table 3> indicated that selective application of the filter reduced differences in BD-rate saving values based on resolution.

To check the performance of the proposed algorithm, we compared with [13] which was based on similar approach. This method employed smoothing filter on the up-sampled residual data from the base layer. <Table 5> shows the result in low delay-P. The Class A and B sequences were tested in same condition. From the result, the proposed scheme achieved 0.6% of gain in Y component. In chroma components, up to 0.9% of gain was verified. This means that the proposed scheme is more efficient to make a scalable video content while keeping the same visual quality.

표 5. Low delay-P 2X 환경에서 계층 간 예측을 위한 적응적 필터 선택 기법에 의한 성능 비교
Table. 5. Performance comparison for inter-layer prediction in low delay-P 2X

Low Delay-P 2X (BD-rate)			
Methods	Average values		
	Y	U	V
Method [13]	-0.9%	-1.2%	-1.3%
Proposed method	-1.5%	-2.1%	-1.7%

표 6. Low delay-P 2X 환경에서 개별 방법의 계산 복잡도 비교
Table. 6. The complexity comparison of each method in low delay-P 2X

Low Delay-P 2X (BD-rate)			
Methods	Average values		
	Y	U	V
Method [13]	-0.9%	-1.2%	-1.3%
Proposed method	-1.5%	-2.1%	-1.7%

<Table 6> illustrates the complexity comparison of each method when comparing to SHVC reference model (SHM) 5.0. That is to say, the complexity of SHVC reference model (SHM) 5.0 is the reference (100% of complexity). The proposed method added just 1% of complexity to Method [13] in terms of total encoding time. It means that the computational burden is negligible when compared with Method [13]. Also, about 10% of the computational complexity was observed comparing to the original SHVC reference model.

Ⅳ. Conclusions

An adaptive filter selection algorithm has been proposed based on an RD-cost competition for scalable extension HEVC (SHVC) technology. To improve video quality, the proposed scheme reduced sampling error by adding smoothing and sharpening filters on residual image in inter-layer prediction of EL. Experimental results showed that the proposed method achieved a 1.5% BD-rate savings, on average, for Low Delay-P in the Main profile configuration with SHM 5.0 reference software.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2016R1D1A1B04934750).

References

[1] G. J. Sullivan, J. R. Ohm, W.-J. Han, T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 22, No. 12, pp. 1649-1668, Dec. 2012.

[2] 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 of video coding (HEVC),” *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 22, No. 12, pp. 1669-1684, Dec. 2012.

[3] T. Wiegand and G. J. Sullivan, “The H.264/AVC Video Coding Standard,” *IEEE Signal Processing Magazine,* Vol. II, pp. 148-153, 2007.

[4] David Flynn, Detlev Marpe, Matteo Naccari, Tung Nguyen, Chris Rosewarne, Karl Sharman, Jel Sole, Jizheng Xu, “Overview of the Range Extensions for the HEVC Standard: Tools, Profiles, and Performance,” *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 26, No. 1, pp. 4-19, Jan., 2016.

[5] Gwang-Soo Hong, Byung-Gyu Kim, “Stereo Matching Algorithm Based on Fast Guided Image Filtering for 3-Dimensional Video Service,” *Journal of Digital Contents Society*, Vol. 17, No. 6, pp. 523-529, Dec. 2016.

[6] H. Schwarz, D. Marpe, T. Wiegand, “Constrained inter-layer prediction for single-loop decoding in spatial scalability,” *IEEE International Conference on Image Processing*, Vol. 2, pp. 870-873, Sep. 2005.

[7] H. Schwarz, D. Marpe, T. Wiegand, Further Results on Constrained Inter-layer prediction, Joint Video Team, Busan, KR, Doc. JVT-O074, Apr. 2005.

[8] W. Zhang, A. Men, and P. Chen, “Adaptive Inter-layer Intra Prediction in Scalable Video Coding,” in *IEEE International Symposium on Circuits and Systems (ISCAS)*, Taiwan, pp. 876-879, May, 2009.

[9] Chang-Ming Lee, Yu-Ciao Yang, Jui-Chiu Chiang, “New Inter-layer Intra Prediction for Spatial Scalable Video Coding,” in *19th European Signal Processing Conference (EUSIPCO)*, Barcelona, pp. 2176-2179, Aug. 2011.

[10] ISO/IEC JTC-1/SC29/WG11, Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC), Joint Video Team, Stockholm, Sweden, w12957, July, 2012.

[11] J. Dong, Y. He, Y. He, G. McClellan, E.-S. Ryu, X. Xiu and Y. Ye, Description of scalable video coding technology proposed by InterDigital, Joint Collaborative Team on Video Coding (JCT-VC), Shanghai, China, Doc. JCTVC-K0034, Oct., 2012.

[12] X. Xiu, Y. He, Y. He and Y. Ye, TE C5: Motion field mapping, Joint Collaborative Team on Video Coding (JCT-VC), Geneva, Switzerland, Doc. JCTVC-L0052, Jan., 2013.

[13] H. Lee, J. Lee, J. W. Kang, and J. S. Choi, “Smoothed inter-layer residual prediction for SHVC,” in *IEEE International Conference on ICT Convergence*, Seoul, KR, pp. 604-605, Oct., 2011.

[14] M. Sychev, S. Ikonin, and V. Anisimovski, “Sharpening filter for interlayer prediction,” in *IEEE Visual Communications and Image Processing Conference*, Valletta, Malta, pp. 470-473, Dec., 2014.

[15] X. Li et al, Common SHM test conditionds and software reference configuration, Joint Collaborative Team on Video Coding (JCT-VC), San Jose, US, Doc. JCTVC-P1009, Jan. 2014.

[16] G. Bjontegaard, Calculation of average PSNR differences between RD curves, ITU-T SGI6/Q6, Texas, US, Doc. VCEG-M33, April, 2001.

이종혁(Jong-Hyeok Lee)

2012년 : 선문대학교 컴퓨터공학과 석사 (공학석사)
2013년~현 재 : 선문대학교 컴퓨터공학과 박사과정

2016년～현　재: 숙명여자대학교 빅데이터활용연구소 연구원
※관심분야 : 비디오 신호처리 (Video signal processing), 차세대 비디오 압축 (Future video coding), 딥러닝 (Deep learning) 등

김병규(Byung-Gyu Kim)

1998년 : 한국과학기술원 전기및전자공학 석사 (공학석사)
2004년 : 한국과학기술원 전기및전자공학 박사 (공학박사)

2009년～2015년: 선문대학교 컴퓨터공학과 부교수
2016년～현　재: 숙명여자대학교 IT공학과 부교수
※관심분야 : 비디오 신호처리 (Video signal processing), 딥러닝 (Deep learning) 등