정규논문 (Regular Paper)
방송공학회논문지 제20권 제6호, 2015년 11월 (JBE Vol. 20, No. 6, November 2015)
http://dx.doi.org/10.5909/JBE.2015.20.6.921
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

GPB 메카니즘을 활용한 스케일러블 다시점 비디오 부호화를 위한 효율적인 DPB 설계 기법

정 태 준[a)], 고 명 필[a)], 서 광 덕[a)‡]

An Efficient DPB Design Scheme for Scalable Multi-view Video Coding Using GPB Mechanism

Tae-jun Jung[a)], Myung Pil Ko[a)], and Kwang-deok Seo[a)‡]

요 약

본 논문에서는 효율적인 스케일러블 다시점 비디오 부호화를 수행하기 위해 필요한 참조화면 재순서화, 마킹 프로세스, 참조픽처 리스트 구성을 포함하는 DPB (Decoded Picture Buffer) 작동을 위한 새로운 설계 방법을 제안한다. 다양한 실험 결과를 통해서 제안된 방법에 의해 스케일러블 다시점 비디오 부호화에 있어서 개선된 압축효율을 얻을 수 있고 BD-Rate 및 BD-PSNR 측면에서 향상된 비디오 화질을 얻을 수 있음을 확인한다.

Abstract

In this paper, we propose a novel design scheme for the operation of Decoded Picture Buffer (DPB) including reference picture re-ordering, marking process, and reference picture list construction to perform an efficient scalable multi-view video coding. Extensive simulations show that the proposed method can provide improved compression efficiency and improved video quality measured in terms of BD-Rate and BD-PSNR for the scalable multi-view video coding.

Keywords: Scalable video coding, Multi-view video coding, Scalable multi-view video coding, DPB management

a) 연세대학교 컴퓨터정보통신공학부(Division of Computer and Telecommunications Engineering, Yonsei University)
‡ Corresponding Author : 서광덕(Kwang-deok Seo)
E-mail: kdseo@yonsei.ac.kr
Tel: +82-33-760-2788
ORCID: 0000-0001-5823-2857
※ 본 논문은 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 '바이오 제조GMP 기술인력양성사업'의 지원을 받아 수행하였습니다 (N0000961).
· Manuscript received September 4, 2015; Revised November 16, 2015; Accepted November 16, 2015.

Ⅰ. Introduction

User's interests and need for 3D content are rapidly increasing, mainly in the movie industry and in user-created and self-made videos. Thus, 3D displays have been developed or are being developed with the first priority on diverse platforms. Given this tendency, video coding schemes for efficiently transmitting 3D content using various terminals through diverse networks will become necessary. This paper introduces a scalable multi-view video coding scheme for comprehensively processing the 3D content and transmitting it to either 3D displays or existing 2D displays in heterogeneous environments, while supporting one-source multi-use services.

In order to transmit multi-view video in a scalable way over heterogeneous network environments and devices, the multi-view video needs to be encoded in a scalable format [1]. Therefore, it is important to devise an efficient coding structure for scalable multi-view video coding where the multi-layer video coding technology of SVC (scalable video coding) [2] is integrated with the multi-view video coding technology of MVC (multi-view video coding) [3]. Fig. 1 shows the system architecture of the proposed scalable multi-view video coding for dual-views video, where major coding structures and mechanisms of SVC and MVC are integrated. Each view is basically processed by applying the layered coding mechanism of SVC to generate the base and enhancement layers in a scalable way. To achieve inter-view prediction effect, the reconstructed picture of view0 is used as an additional reference frame for coding view1.

To smoothly integrate the coding structures and mechanisms of SVC and MVC, there are some critical problems occurring in the management of reference pictures, which need to be resolved. In this paper, we propose a new design scheme of reference picture list using Generalized P and B-Picture (GPB) mechanism [4], [5] for efficient scalable multi-view video coding.

Ⅱ. Design Of Reference Picture List Based on GPB Mechanism

In HEVC, a new type of B picture, the generalized P and B (GPB) picture, has been introduced to preserve low delay operations while providing increased coding performance. A GPB picture only allows prediction from past reference pictures. Fig. 2 shows the temporal prediction structure of low-delay mode. The low-delay mode is designed for real-time video communication requiring minimum latency, such as video telephony and video conferencing [5]. Every picture except the first one is predicted from the past pictures. The B pictures used in the low-de

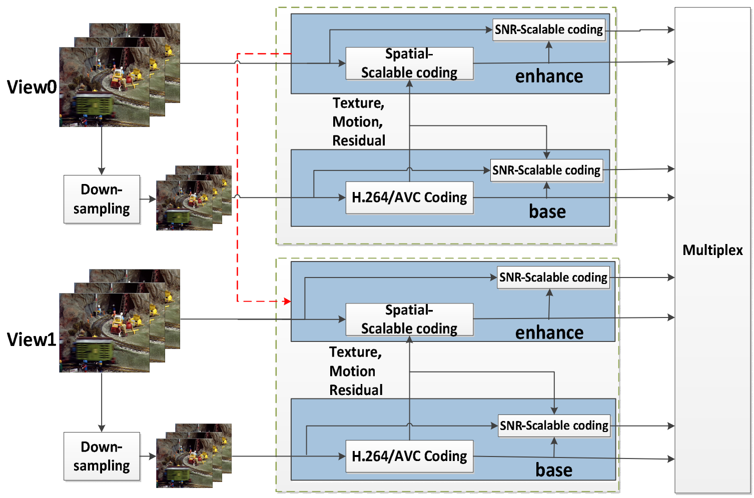

그림 1. 제안한 스케일러블 다시점 비디오 부호화를 위한 시스템 구조도

Fig. 1. System architecture of the proposed scalable multi-view video coding

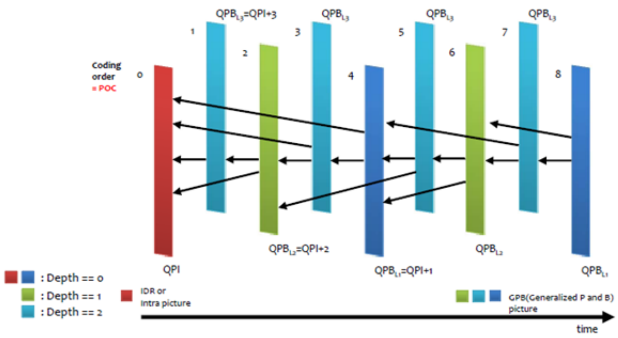

그림 2. 저지연 모드에서 사용되는 시간적 예측 구조.
Fig. 2. Temporal prediction structure used in low-delay mode

lay mode are called GPB pictures. GPB pictures use only temporally previous reference pictures whose picture order counter (POC) is smaller than the current picture.

In order to estimate the motion displacement more accurately in the inter-view prediction process and reduce residual signals for the proposed scalable multi-view video coding, we propose a novel design scheme for constructing a reference picture list by using GPB mechanism.

Fig. 3 shows the management of reference picture list based on GPB mechanism, where dual views (view0 and

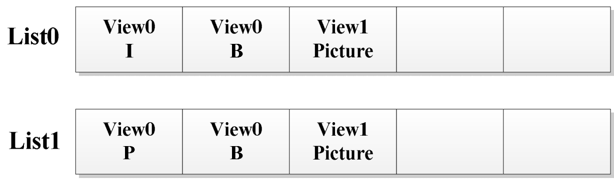

그림 3. GPB 기반의 참조 화면 리스트 관리
Fig. 3. Management of the reference picture list based on GPB

view1) are employed for multi-view video coding. To enhance the prediction accuracy for predictively coding pictures of view0, the reference picture of view1 is copied from list0 to list1 and is used as an additional reference frame for coding view0.

Fig. 4 shows the overall architecture of the designed reference picture list for the proposed scalable multi-view video coding employing inter-view prediction. In this architecture, predictive coding efficiency can be improved by using enhancement layer picture of view0 (EGPB0_2) as an additional reference picture for coding the picture of view1 (EGPB1_2). As shown in Fig. 4, EGPB0_2 of list0 is copied to list1 in the reference picture list for inter-view prediction.

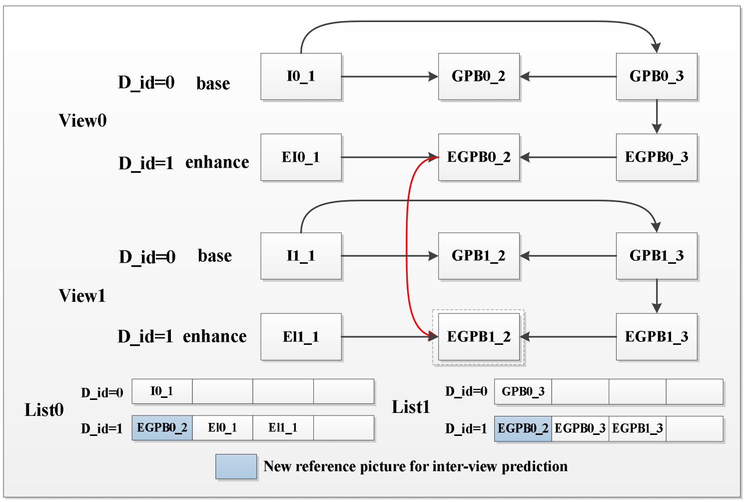

그림 4. 시점 간 예측을 지원하는 참조 화면 리스트 설계의 전체적인 구조
Fig. 4. Overall architecture of the designed reference picture list supporting inter-view prediction

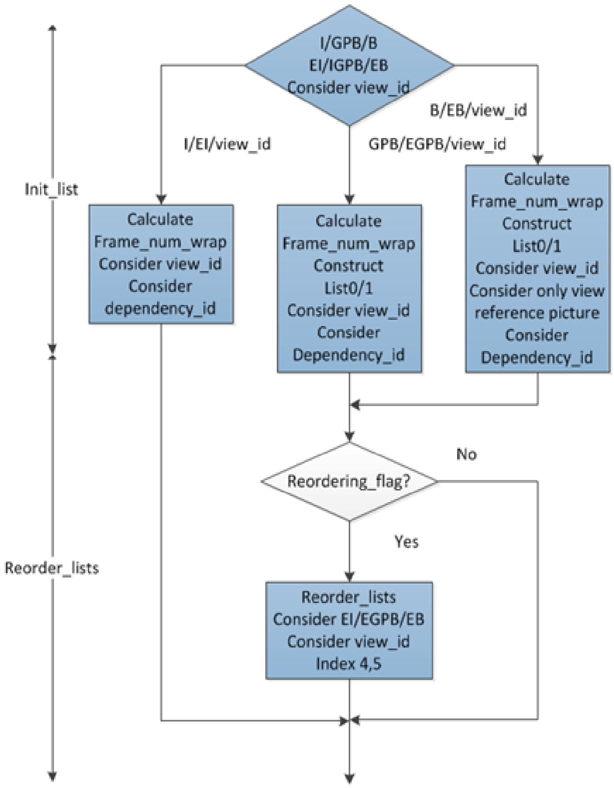

그림 5. 참조 화면 리스트 구성을 위한 흐름도
Fig. 5. Flow chart of reference picture list construction

Fig. 5 shows the proposed flow chart design of reference picture list construction. The proposed design considers not only view_id for MVC reference picture list, but also dependency_id for SVC reference picture list. In Fig. 5, Init_list part is for initializing reference picture list. This initialization process is invoked when decoding I, GPB, B, EI, IGPB or EB slice header. List0 and List1 have intial entries as specified in the initialization process. Reorder_lists part is for reordering reference picture list. If reordering_flag is marked with 0, then reference pictures stored in the reference picture list skip reordering process. Otherwise, reference pictures stored in the reference picture list go through reordering process. Decoded reference pictures are marked as "used for short-term reference" or "used for long-term reference" according to the information specified in the bitstream during the marking process for the decoded reference pictures. Short-term reference pictures are identified by the value of frame_num. Long-term reference pictures are assigned with a long-term frame index according to the information specified in the bitstream during the marking process for the decoded reference pictures.

Fig. 6 shows the proposed flow chart design of DPB

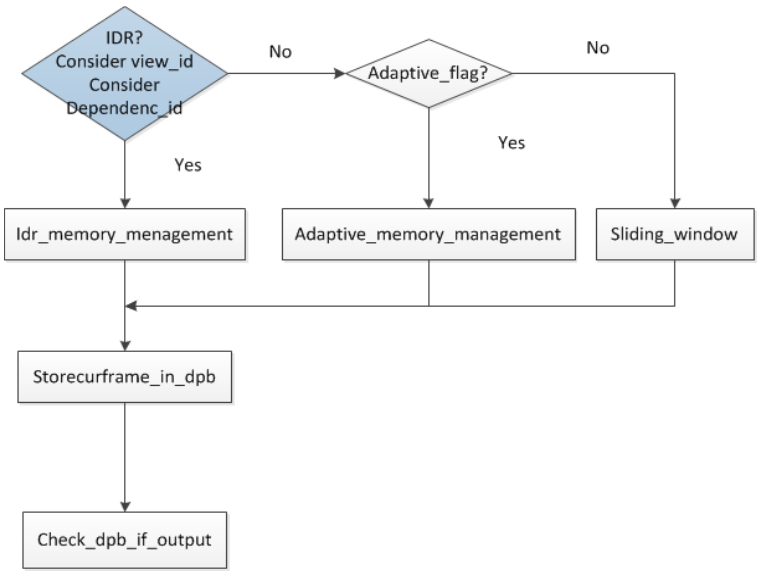

그림 6. DPB 관리를 위한 흐름도
Fig. 6. Flow chart of DPB management

management process. It considers both view_id of MVC coding and dependency_id of SVC coding for DPB management. Marking process for the decoded reference pictures takes the following steps. Depending on whether the current picture is an IDR picture or not, pictures stored in the DPB are determined to be used as a reference picture. If the current picture is an IDR picture, it is determined not to be used for reference picture. For the case of non-IDR picture, DPB management process invokes the sliding window method if adaptive_flag is equal to 0. If adaptive_flag is equal to 1, DPB management process invokes the adaptive memory control method, as shown in Fig. 6.

Ⅲ. Experimental Results

In order to evaluate the compression efficiency and video quality after applying the proposed mechanism in actual coding process, we developed a scalable multi-view video coding system which integrates JSVM (Joint Scalable Video Model) 9.19.14 and JMVC (Joint Multi-view Video Coding) 8.5 together. Test condition for the performance evaluation was set to GOP_size=9, Intra_period=16, and frame_rate=30 frames per second. We used Tunnel and BMX sequences which are composed of dual views, one for the left-view and the other for the right-view. The maximum number of reference frames for motion estimation was set to 6. We compare the performance in terms of compression efficiency and average PSNR between the bitstreams generated by the SVC's conventional predictive coding mechanism and by the proposed predictive coding mechanism. The test is performed using four different values of QP (quantization parameter), 27, 30, 32, and 35.

Table 1 shows the test results for coding the left-view video of the Tunnel and BMX sequences with CIF resolution. “SVC prediction” in Table 1 denotes the results of coding for the left-view video by using SVC’s predictive coding architecture. “Proposed prediction” denotes the results obtained by using the proposed predictive coding architecture. The experimental results in Table 1 show that we could achieve 0.1802 dB increase in BD-PSNR and 1.7157% decrease in BD-Bitrate at the same time by employing the proposed prediction mechanism for the Tunnel sequence. Table 1 also shows that 0.1251 dB increase in BD-PSNR and 2.2188% decrease in BD-Bitrate could be achieved by the proposed prediction mechanism for the BMX sequence. Fig. 7 shows the comparison of PSNR performance by plotting RD (rate-distortion) curve for the BMX test sequences.

표 1. CIF 해상도의 Tunnel 및 BMX 테스트 영상에 대한 다양한 QP 값에 따른 BD-PSNR 및 BD-Bitrate 비교
Table 1. Comparison of BD-PSNR and BD-Bitrate for various QP values for the Tunnel and BMX video sequences of CIF resolution

		Encoding test sequence data					
		SVC prediction		Proposed prediction		BD-PSNR (dB)	BD-Bitrate (%)
		(bits)	(PSNR)	(bits)	(PSNR)		
Tunnel	QP=27	1018984	37.20	1007184	37.22	0.0182	-1.7157
	QP=30	744872	35.35	732328	33.36		
	QP=32	609200	34.13	600864	34.15		
	QP=35	494432	32.69	487520	32.72		
BMX	QP=27	1770816	37.04	1739848	37.04	0.1251	-2.2188
	QP=30	1260248	35.15	1235840	35.16		
	QP=32	1019680	33.97	1000008	33.99		
	QP=35	801520	32.60	787936	32.65		

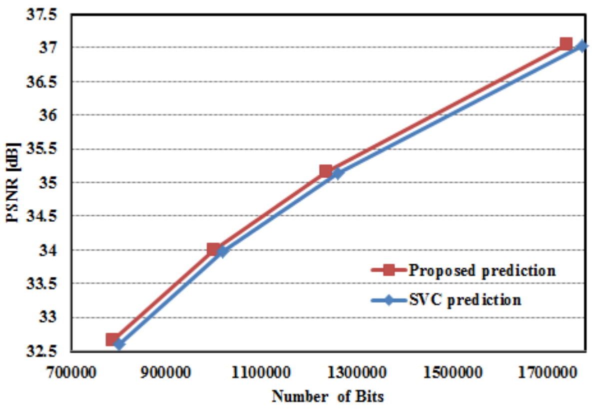

그림 7. BMX 시퀀스에 대한 PSNR 성능 비교
Fig. 7. Comparison of PSNR performance for the BMX sequence

Overall, the application of the proposed prediction architecture results in not only decrease in compressed data size measured in bit-rate, but also quality improvement measured in average PSNR, when compared to the SVC's conventional prediction architecture, and this performance improvement is proportional to the size of QP values. The reason for this behavior is that when the size of QP values increases, quality of the base-layer video becomes deteriorated accordingly during quantization process. Therefore, it is more desirable to perform inter-view prediction for more exact motion estimation rather than to perform inter-layer prediction including up-sampling process referring to the deteriorated base-layer video. With the adoption of inter-view prediction, we can estimate the motion displacement more accurately and gain performance improvements in terms of compression efficiency as well as video quality.

Ⅳ. Conclusions

In this paper, we proposed an efficient design scheme for scalable multi-view video coding based on GPB mechanism. The proposed method results in 1-2% decrease in compressed bitrate and a little increase in PSNR at the same time, when compared to the SVC's conventional coding scheme.

References

[1] N. Ozbek, and M. Tekalp, “Scalable multi-view video coding for interactive 3DTV,” in Proc. Int. Conf. Multimedia and Expo, pp. 213-216, 2006.

[2] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the scalable video coding extension of the H.264/AVC standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 9, pp. 1103-1120, Sep. 2007.

[3] A. Vetro. T. Wiegand, G. Sullivan, “Overview of the Stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard,” Proc. IEEE, vol. 99, no. 4, pp. 694-707, Apr. 2011.

[4] M. Flierl and B. Girod, “Generalized B pictures and the draft H.264/AVC video compression standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 587-597, Jul. 2003.

[5] K. Kim, H. Kim, J. Choi, and G. Park, “MC complexity reduction for generalized P and B pictures in HEVC,” IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 10, pp. 1723-1728, Oct. 2014.

저 자 소 개

정 태 준

- 2010년 2월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2012년 8월 : 연세대학교 전산학과 석사
- 2012년 9월 ~ 현재 : 연세대학교 전산학과 박사과정
- Orcid : 0000-0002-8397-7188
- 주관심분야 : 영상부호화, 영상통신, 멀티미디어 통신 프로토콜

고 명 필

- 2013년 8월 : 연세대학교 컴퓨터정보통신공학부 학사
- 2013년 9월 ~ 현재 : 연세대학교 전산학과 석사과정
- Orcid : 0000-0001-9525-5476
- 주관심분야 : 모바일 소프트웨어, 프로그래밍언어, 컴파일러, 프로그램 분석, 소프트웨어 검증

서 광 덕

- 1996년 2월 : KAIST 전기 및 전자공학과 학사
- 1998년 2월 : KAIST 전기 및 전자공학과 석사
- 2002년 8월 : KAIST 전기 및 전자공학과 박사
- 2002년 8월 ~ 2005년 2월 : LG전자 단말연구소 선임연구원
- 2012년 9월 ~ 2013년 8월 : Courtesy Professor, Univ. of Florida, USA
- 2005년 3월 ~ 현재 : 연세대학교 컴퓨터정보통신공학부 정교수
- Orcid : 0000-0001-5823-2857
- 주관심분야 : 영상부호화, 영상통신, 디지털 방송, 멀티미디어 통신시스템