

VVC 화면 내 부호화 복잡도 감소를 위한 부호화 기술 선택 방법

김범윤, 박지윤, 전병우
 성균관대학교 전자전기컴퓨터공학과
 {kbumyoon, jiyoonpark, bjeon}@skku.edu

VVC Intra Coding Tool Selection for Low Complexity Video Coding

Bumyoon Kim, Jeeyoon Park, and Byeungwoo Jeon
 Department of Electrical and Computer Engineering, Sungkyunkwan University

요 약

최신 비디오 압축 표준인 VVC(Versatile Video Coding)의 화면 내 부호화 기술이 사용하는 MIP (Matrix-based Intra Prediction), MRLP (Multiple Reference Line Prediction), ISP (Intra Sub-Partition) 기술의 성능평가를 위하여 VTM (VVC Test Model) 10.0 을 이용한 각 기술들의 비활성화 성능은 보고된 바 있지만, 최근 버전인 VTM 16.0 을 기준으로 하나 또는 복수개의 기술을 비활성화 하였을 때의 성능은 아직 보고된 바 없다. 본 논문에서는 VVC 의 화면 내 부호화 기술들을 활성화 및 비활성화 성능실험을 통하여 부호화 성능 또는 부호화 복잡도를 우선시하는 응용에 따라 VVC 부호화기를 경량화 시킬 수 있는 부호화 기술 선택 방법을 제안한다.

1. 서론

2020 년 7 월에 표준화가 완료된 VVC(Versatile Video Coding) 비디오 압축 기술은 HEVC(High Efficiency Video Coding)에 비해 AI (All Intra) 조건하에 Y 채널에서 -25.06%, Cb 채널 및 Cr 채널에서 각각 -25.37%, -26.85%의 BDBR 이득이 있지만 약 26 배의 부호화 시간 증가도 있음이 보고되었다[1]. 실시간 동영상 서비스 환경을 고려하여 부호화 효율을 크게 떨어뜨리지 않으면서 VVC 의 복잡도 감소에 대한 연구가 필요하다. VVC 의 화면 내 예측 기술에는 MIP (Matrix-based Intra Prediction), MRLP (Multiple Reference Line Prediction), ISP (Intra Sub-Partition)가 있으며 VTM (VVC Test Model) 10.0 에서 각 기술에 대한 tool-off (비활성화) 테스트의 BDBR 변화는 표 1 과 같다[2].

표 1 VTM 10.0 화면 내 기술의 tool-off test 성능 변화

Tool	BDBR (%)			Relative Encoding Time
	Y	Cb	Cr	
MIP	0.63	0.17	0.18	90%
MRLP	0.33	0.12	0.14	99%
ISP	0.48	0.28	0.23	85%

한편, 복수개의 기술을 비활성화 한다고 하더라도 부호화 복잡도는 각 기술에 대한 부호화 복잡도의 합에 해당하는 만큼 선형적으로 감소하지 않으며, 또한 가장 최신 버전인 VTM 16.0 에서 하나 또는 복수개의 기술을 비활성화 했을 경우에 대한 부호화 성능 비교는 아직 보고된 바 없다. VTM 16.0 은 VTM 10.0 대비 AI 조건하에 Y 채널에서 -0.67%, Cb 채널 및 Cr 채널에서 각각 -0.55%, -0.30%의 BDBR 이득이 있고, 복잡도는 91%로 감소하였다.

본 논문에서는 MIP, MRLP, ISP 기술을 비활성화하는 모든 경우에 대한 부호화 성능을 비교함으로써 VVC 화면 내 부호화 복잡도 감소를 위한 화면 내 예측 기술에 대한 최적의 선택을 제안하고자 한다.

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 지역지능화혁신인재 양성 (Grand ICT 연구센터) 사업 (HTP-2022-2015-0-00742) 과, 과학기술정보통신부의 재원으로 한국연구재단의 사업(NRF-2020R1A2C2007673)의 연구결과로 수행되었음.

2. 인트라 부호화 기술 및 선택적 활성화

본 연구에서 고려하는 화면 내 예측 기술인 MIP, MRLP, ISP 각 기술들에 대한 설명은 다음과 같다.

MIP 는 현재 블록의 참조 화소를 필터링한 뒤 사전에 학습된 행렬과의 곱으로 예측자를 생성하는 기술이다. 4x4 크기의 블록의 경우 16 종류의 행렬을 사용하며, 너비가 4 이거나 높이가 4 인 블록 또는 8x8 크기의 블록은 8 종류의 행렬을 사용한다. 그 외 크기의 블록에 대해서는 6 종류의 행렬을 사용한다. 이 때 예측자 생성 시 각 행렬을 전치(transpose) 하여 곱할 수 있다. 이 때 행렬의 전치 여부는 *intra_mip_transposed_flag* 를 통해 신호되며 행렬의 종류는 *intra_mip_mode* 를 통해 신호된다. 현재 블록의 예측 모드가 MIP 인 경우 MRLP, ISP 는 적용될 수 없다[3].

MRLP 는 화면 내 예측 시 현재 블록과 바로 인접한 참조 화소 대신 한 화소 또는 두 화소 거리만큼 떨어진 위치의 화소를 참조하여 예측자를 생성하는 기술이다[4]. *intra_luma_ref_idx* 를 통해 현재 블록과 참조 화소까지의 거리를 전송한다. *intra_luma_ref_idx* 가 0 이면 바로 인접한 참조 화소를 사용하는 것을 의미하며 1 이면 한 화소 거리, 2 이면 두 화소 거리만큼 떨어진 참조 화소를 사용하는 것을 의미한다. MRLP 는 예측 모드가 MPM (Most Probable Mode) 중 하나 일 때만 적용 가능하며 planar 모드일 때는 적용이 불가능 하다. 또한 현재 블록이 현재 CTU(Coding Tree Unit)의 상단 경계에 인접에 있는 경우에도 MRLP 의 적용이 불가능하다.

ISP 는 현재 블록을 블록의 크기에 따라 2 개 또는 4 개의 서브 블록으로 분할하여 각각의 서브 블록에 대해 동일한 예측 모드로 예측자를 생성 후 서브 블록별로 변환 및 양자화 과정을 수행하는 기술이다. 서브 블록의 분할은 수평 또는 수직 방향으로 가능하며 ISP 적용 여부는 *intra_subpartitions_mode_flag* 를 통해 신호하고 분할 방향은 *intra_subpartitions_split_flag* 를 통해 신호한다. *intra_luma_ref_idx* 가 0 이 아닌 경우 ISP 는 적용이 불가능하고, 블록의 크기가 4x4 인 경우 또한 ISP 의 적용이 불가능하다[5].

본 연구에서는 VVC 의 화면 내 부호화 기술인 MIP, MRLP, ISP 중 두 가지 또는 세 가지 기술을 비활성화하는 각 경우에 대하여 모두 활성화하는 경우 대비 BDBR 변화와 부호화 시간 변화 비교를 통해 VVC 화면 내 부호화의 복잡도를 감소시키는 최적의 기술 선택 방법을 살펴본다.

3. 실험 및 결과

본 논문에서 제안하는 복수개의 기술 비활성화에 대한 성능 평가는 VTM 공통 실험 조건의 AI 구성을 따라 4 가지 QP (Quantization Parameter) 22, 27, 32, 37 을 사용하였으며, 가장

표 2 복수개의 기술 비활성화에 대한 성능 변화

Tool on/off			BDBR (%)			Relative Encoding Time
MRLP	ISP	MIP	Y	Cb	Cr	
X	X	X	1.50	0.73	0.71	69%
X	X	O	0.84	0.56	0.56	83%
X	O	X	0.96	0.36	0.35	83%
O	X	X	1.13	0.53	0.46	75%
O	O	X	0.62	0.20	0.25	89%
O	X	O	0.50	0.32	0.34	85%
X	O	O	0.30	0.14	0.21	100%

최신 버전인 VTM 16.0 을 사용하여 영상의 모든 프레임에 대하여 이루어졌다[6]. VVC 에서 3 가지 화면 내 부호화 기술들을 비활성화하는 모든 경우에 대한 실험 결과는 표 2 와 같으며 VTM 16.0 에서 3 가지 화면 내 기술을 모두 활성화했을 때를 앵커로 했을 때의 BDBR 성능 변화와 부호화 시간 변화를 나타낸다.

4. 결론

본 연구를 통하여 실제 사용환경에서는 표 2 를 참고하여 다양한 선택을 할 수 있으며, 특히 복잡도 감소를 우선시하는 경우 세 기술을 모두 비활성화하는 방법을 선택하고, 부호화 성능 유지를 우선시하는 경우 MRLP 만 비활성화 하는 것이 BDBR 손실이 가장 적으나 부호화 복잡도 감소가 없기 때문에 ISP 만 비활성화하는 방법을 선택하는 것이 우수한 것으로 관찰되었다.

참고 문헌

- [1] F. Bossen, X. Li and K. Suehring, "AHG report: Test model software development (AHG3)," Joint Video Experts Team (JVET), 20th Meeting, teleconference, Tech. Rep. JVET-T0003, Oct. 2020.
- [2] W. Chen *et al.*, "JVET AHG report: Tool reporting procedure (AHG13)," Joint Video Experts Team (JVET), 20th Meeting, teleconference, Tech. Rep. JVET-T0013, Oct. 2020.
- [3] B. Bross, J. Chen, S. Liu, and Y. Wang, "Versatile Video Coding Editorial Refinements on Draft 10," Joint Video Experts Team (JVET), 20th Meeting, teleconference, Tech. Rep. JVET-T2001, Oct. 2020.
- [4] Y. Chang *et al.*, "Multiple Reference Line Coding for Most Probable Modes in Intra Prediction," in *Proc. Data Compression Conference (DCC)*, 2019, pp. 559-559
- [5] D. Santiago *et al.*, "Design of the intra subpartition mode in VVC and its optimized encoder search in VTM," in *Proc. SPIE 11510, Application of Digital Image Processing XLIII*, 2020.
- [6] F. Bossen, J. Boyce, K. Suehring, X. Li, and V. Seregin, "VTM common test conditions and software reference configurations for SDR video," Joint Video Experts Team (JVET), 20th Meeting, teleconference, Tech. Rep. JVET-T2010, Oct. 2020.