

심층적 강화학습 기반 적응적 GOP 선택을 통한 HEVC/H.265 인코더 제어

이정경, 김나영, 강제원 *

이화여자대학교 전자전기공학과

jk1931204@gmail.com, 12skdud21@gmail.com, jewonk@ewha.ac.kr

Deep Reinforcement Learning based Adaptive GOP Selection for HEVC/H.265 Encoder

Jung-Kyung Lee^{1,2}, Nayoung Kim¹, Je-Won Kang^{1,2}

¹) Department of Electronic and Electrical Engineering, Ewha Womans University

²) Smart Factory Multidisciplinary Program, Ewha Womans University

요 약

본 논문에서는 심층적 강화학습 기반 GOP (Group of Picture) 크기를 선택하여 HEVC/H.265 의 인코더를 제어하는 방법을 제안한다. 기존 방법에서는 현재 비디오 신호를 부호화 하는 과정에서 이미 부호화한 정보를 사용해야하는 부호화 의존성에 관한 문제가 있었다. 제안 방법은 강화학습 방식을 도입하여 이러한 문제를 극복하고 입력 비디오의 시간적 상관도에 따라 GOP 의 크기를 적응적으로 선택하여 부호화 한다. 본 논문에서는 GOP 선택을 위한 강화학습 환경을 새롭게 정의하고 부호화 성능에 따른 보상을 부여하는 방식으로 학습을 수행한다. 제안된 적응적 GOP 선택에 따라 인코더 제어 시, 부호화 방법의 부호화 효율이 -6.07% BD-rate 향상된 실험 결과를 보이며 본 방법의 우수성을 입증한다.

1. 서론

영상 신호를 포함하는 멀티미디어의 방대한 데이터 사용에 따라 이를 효과적으로 저장 및 전송하기 위해 비디오 압축에 관한 연구 및 표준화 기술개발이 계속 진행되어 왔다. HEVC/H.265[1]을 포함한 비디오 부호화 기술에서는 GOP (Group of Picture) 단위로 영상을 나누어 부호화를 진행한다. 따라서, 현재 비디오 신호를 부호화 하는 과정에서 이미 부호화한 정보를 사용해야하는 부호화 의존성에 관한 문제가 있다. 따라서, 입력 영상의 특성이 각기 달라 최고의 부호화 효율을 내는 최적의 GOP 를 찾고자 하였다.

비디오 부호화 과정에서 GOP 크기는 해당 GOP 내 프레임 간의 시간적 상관도와 연관성이 존재한다. 이는 GOP 의 크기 변화는 GOP 의 참조계층 구조와 현재 부호화 할 프레임의 참조 프레임의 변화를 가져오기 때문이다. 예를 들어 급격한 장면

변화나 큰 움직임 변화가 있는 비디오의 경우 GOP 크기가 크게 되면, 참조 프레임이 텍스처 정보가 현재 부호화를 진행하는 프레임과 상이하게 달라지게 되어 부호화 효율 악화를 초래한다. 따라서 큰 움직임 변화를 보이는 비디오에서는 프레임 간 시간적 상관도가 작기 때문에 GOP 를 작게, 반면 적은 움직임 변화를 보이는 비디오에서는 시간적 상관도가 크기 때문에 GOP 크기를 크게 사용하면 부호화 성능을 높일 수 있게 된다.

기존 연구는 영상 내의 움직임을 파악하여 정도에 따라 GOP 크기를 선택하는 방법을 택하였다 [2, 3, 4]. [2]는 영상의 움직임의 안정도(steadiness)를 정의하여 HEVC 에서의 GOP 선택 방법을 제시하였다. 정의된 안정도 값을 구간 별로 임계값을 설정해 구간 별로 다른 GOP 크기를 선택하였다. [3]는 입력 영상의 공간적, 시간적 복잡도를 부호화 전 미리 계산하여 임계값을 통해 GOP 크기를 결정하였다. 이러한 결정론적인 가정 기반 알고리즘에 벗어나기 위해 Wyner-Ziv 비디오 압축에서

Decision Tree 방법이 고안하였다 [5]. 하지만, 앞선 연구와 비슷하게 영상의 특징을 경험적 탐색 방법으로 정의하여 사용하였다. 제안 방법에서는 선행연구 방법을 탈피해 심층적 강화학습을 사용하여 적응적으로 입력 비디오에 대한 최적의 GOP 크기를 선택하는 인코더 제어 방법을 제안하였다.

본 논문에서는 먼저 GOP 크기에 따른 가능한 코딩 시나리오를 설명하고 심층적 강화학습의 환경 및 학습 방법을 소개한다. 그 후, 3 절에서는 기존의 고정된 GOP 를 사용하는 부호화 방법 대비 제안 방법의 성능 향상을 실험적으로 보인다.

2. 제안 방법

1). 가능한 GOP 크기 선택 시나리오

그림 1 은 POC 0~32 를 기준으로 최소 GOP 크기가 8, 최대 GOP 크기가 32 일 때 선택 가능한 GOP 조합을 보여준다. 고정 GOP8 사용시에는 GOP (8,8,8,8)을 고정 GOP16 사용시에는 (16,16)을 이용하여 동일하게 부호화가 진행된다. 반면 적응적 GOP 시나리오에서는 GOP (16,16), (8,8,16), (16,8,8)의 선택이 가능하다.

GOP32	32			
GOP16,16	16	16		
GOP8,8,8,8	8	8	8	8
GOP8,8,16	8	8	16	
GOP16,8,8	16	8	8	

그림 1. 32 장의 프레임을 기준으로 GOP 크기에 따른 가능한 GOP 선택 시나리오

2). 적응적 GOP 선택을 위한 강화학습

제안 방법에서는 강화학습을 위한 적응적 GOP 선택 환경(Environment) 안에서 GOP 선택이라는 동작(Action)을 취할 수 있는 객체(Agent)를 인코더로, 입력 비디오가 현재 환경의 상태(State)로 정의한다. 제안 방법은 부호화 효율에 따른 보상(Reward)이 최대화하는 방향으로 학습한다.

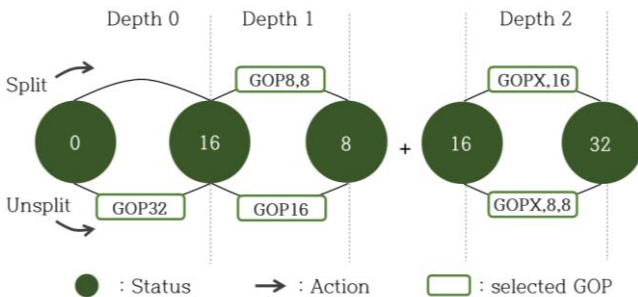


그림 2 강화학습 환경 설정

그림 2 는 강화학습 기반 GOP 선택 환경의 예시를 보여준다. 깊이(Depth)가 0 일 때, 현재 상태인 POC 0 에서 동작(Action)이 분할(Split) 또는 비분할(Unsplit)을 결정하여 다음 부호화를 진행할 프레임으로 POC16 나 POC32 를 선택하게 된다. 이때 비분할을 선택하였다면 GOP32 로 부호화를 진행해 학습이 끝나게 되고, 분할을 선택했다면 현재 상태가 16 이 되어 깊이가 2 로 변경되며 앞서 동일한 방법으로 다음 깊이에서의 동작을 결정하게 된다. 깊이 2 에서는 POC16 과 32 사이에서 GOP8 로 부호화 할지, GOP16 으로 부호화를 진행할지 선택하게 된다. 예컨대, 깊이 (0, 1, 2)에서 각각 (분할, 분할, 비분할)을 선택하였다면 POC 32 일 때 GOP (8,8,16)으로 복호화된다. 제안 방법은 POC 크기에 따라 POC 32 를 기준으로 확장 가능하다. 제안 방법을 학습하기 위해서 DQN (Deep Q-Networks)[6]을 사용하였고, 모든 깊이에 대해서 네트워크의 GOP 선택 시나리오가 가장 좋은 부호화 효율을 보이게 되면 보상은 +1 을 주었다.

3. 실험 결과

본 논문에서 제안하는 강화학습 기반 인코더 제어 방법을 평가하기 위해, 고정 GOP16 을 사용하는 HM 16.9 을 비교실험으로 진행하였다. 테스트 영상은 유튜브(YouTube)에서 장면 전환이 있는 비디오 시퀀스 10 개를 취득하여 사용하였다. 부호화 시, Random Access 코딩 시나리오와 QP 22, 27, 32, 37 에서 BD-rate 를 측정하였다.

실험 결과, 적응적 GOP 을 적용하였을 때, 고정 GOP16 대비 -6.07%의 효율을 보였다. 표 2 에서 적응적 GOP 방법이 -6.07%으로 고정 GOP32 보다 -1.7% 높고 고정 GOP8 보다 -8.3% 높은 부호화 효율을 내는 것을 확인할 수 있다.

Sequence No.	Anchor : HM16.9. GOP(16)			
	고정 GOP32	고정 GOP8	제안 방법	Predicted GOP
1	-2.08	0.21	-2.08	GOP(32)
2	-12.22	7.80	-12.22	GOP(32)
3	-0.20	0.26	-0.20	GOP(32)
4	0.73	-11.99	-11.99	GOP(8,8,8,8)
5	-17.61	-11.06	-17.61	GOP(32)
6	-5.82	3.52	-5.82	GOP(32)
7	0.61	-1.53	-1.53	GOP(8,8,8,8)
8	-8.85	5.22	-8.85	GOP(32)
9	0.45	17.50	0.44	GOP(8,8,16)
10	1.18	12.35	-0.84	GOP(8,8,16)
평균	-4.38	+2.23	-6.07	

표 1. 고정 GOP 방법과 제안 방법의 BD-rate(%) 비교

4. 결론

본 논문에서는 비디오 특성에 따른 강화학습 기반 GOP 크기 결정 방법을 제안하였다. 제안 방법의 실험 결과, 기존 고정 GOP 를 사용하는 HM16.9 대비 -6.07% BD-rate 를 개선하였다. 강화학습을 적용한 입력 비디오의 적응적 GOP 크기 선택 학습 환경 및 방법을 제안하고, 강화학습 네트워크의 출력을 통하여 가변적인 GOP 크기를 결정하는 연구는 본 방식이 최초라는 점에 큰 의미를 갖는다.

참조문헌

- [1] Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." IEEE Transactions on circuits and systems for video technology 22.12 (2012): 1649-1668.
- [2] Poobalasingam, Vigneswaran, and Ebroul Izquierdo. "Steadiness analysis for optimal GOP size selection in HEVC." 2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016.
- [3] Saika, Shintaro, et al. "Improvement of H. 265/HEVC encoding for 8K UHD TV by CU size expansion and inter/intra prediction mode selection." 2018 IEEE International Conference on Consumer Electronics (ICCE). IEEE, 2018.
- [4] Zatt, Bruno, et al. "Gop structure adaptive to the video content for efficient H. 264/AVC encoding." 2010 IEEE International Conference on Image Processing. IEEE, 2010.
- [5] Huong, Thao Nguyen Thi, et al. "Artificial intelligence based adaptive gop size selection for effective wyner-ziv video coding." 2018 International Conference on Advanced Technologies for Communications (ATC). IEEE, 2018.
- [6] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).