

고속 HEVC 부호화기 설계를 위한 화면내 예측 모드 결정 방법

이선영, 노경기, 김형덕, 류성걸, 신재섭
픽스트리

sunnykr12@pixtree.com, hypdm@pixtree.com, donald@pixtree.com,
sgryoo@pixtree.com, jsshin@pixtree.com

Fast Intra Mode Decision Method in HEVC

Sunyoung Lee Gyeonggi Noh Hyeongduck Kim Sungul Ryoo Jae-Seob Shin
PIXTREE, Inc.

요 약

동영상 부호화 표준, HEVC(High Efficiency Video Coding)는 부호화 성능을 극대화하기 위해 총 35 개의 화면내 예측 모드를 사용한다. 화면내 예측 모드는 각도를 가진 모드와 각도가 없는 모드로 구성된다. 부호화 성능을 높이기 위해 사용한 다수의 화면내 예측 모드 방법은 HEVC 부호화기의 복잡도를 증대 시키는데 큰 역할을 하게 된다. 본 논문은 총 35 개의 화면내 예측 모드 중 현재 블록의 주변 블록 정보로부터 얻을 수 있는 예측 모드들 및 각도를 대표하는 예측 모드들을 선별적으로 추려서 후보 예측 모드를 결정하고, 평가 과정을 거쳐 해당 후보 모드 중에서 최종 화면내 예측 모드를 결정한다. 본 제안 방법은 35 개의 전체 화면내 예측 모드 중 소수의 후보 모드만을 평가함으로써 HEVC 표준의 화면내 예측 및 부호화 과정의 복잡도를 감소시키려 한다. 제안 방법을 다양한 테스트 시퀀스에 적용한 결과, 35 개 화면내 예측 모드를 전부 사용한 경우와 비교하여 1.1%의 BD-rates 이 증가하면서 18.7%의 부호화기 복잡도를 감소시킬 수 있었다.

키워드: 고속 모드 결정, 화면내 예측 모드, Fast mode decision, Intra prediction mode, HEVC, Video compression

1. 서론

고화질, 고해상도 동영상에 대한 시장의 요구에 부응하기 위해 두 개의 국제 표준화 기관인 ITU-T VCEG(Video Coding Experts Group)와 ISO/IEC MPEG(Moving Picture Experts Group)이 함께 비디오 부호화 표준인 HEVC(High Efficiency Video Coding)을 제정하였다[1]. HEVC 표준은 기존 MPEG-4 Part 10 AVC/H.264 표준과 비교하여 동일한 화질을 기준으로 2 배의 부호화 성능을 제공한다[2].

부호화 성능을 향상 시키기 위해 HEVC 는 다양한 부호화 기술들을 채택하였다. UHD(Ultra High Definition)에 상응하는 동영상의 해상도를 처리하기 위해 기존 표준들에서 사용하던 부호화 기본 단위인 고정 크기 16x16 매크로 블록을 16x16 부터 64x64 에 이르기까지 가변 크기로 부호화 단위를 지정할 수 있는 CTU(coding tree unit)로 대체하였다. 또한, 화면내 부호화를 위해 총 35 개의 예측 모드를 사용하고 화면간 예측 시에는 Merge, AMVP(advanced motion vector prediction), 및 AMP(asymmetric motion partition) 등의 부호화 방법들을 사용한다. 화질 개선을 위한 DBF(deblocking filter) 및 SAO(sample Adaptive Offset), 병렬 처리를 위한 Tile 및 WPP(wave-front parallel procession) 등의 새로운 툴도 포함하고 있다[3].

HEVC 표준의 부호화 성능을 향상시키고 있는 상기 언급한 다양한 부호화 방법들이 HEVC 표준의 복잡도도 증가시키고 있다. 한가지 예로, 화면내 예측 및 부호화를 위해 각도를 가진 32 개의 모드와 각도가 없는 2 개의 모드를 포함한 총 35 개의 예측 모드를 평가하고 이들 중 최종 한 개의 모드를 선정하는데, 이런 일련의 과정들이 HEVC 부호화기의 복잡도를 증가시키게 된다. 결과적으로, 고속 HEVC 부호화기 설계를 위해서 이런 복잡도를 감소시키려는 시도가 중요한 화제가 되고 있다.

본 논문에서는 고속 HEVC 부호화기 설계를 위하여 화면내 예측 및 부호화 과정의 복잡도를 감소시키는 방법을 제안한다. 제안 방법은 총 35 개의 화면내 예측 모드 중에 소수의 후보 예측 모드를 선정하여 전체 35 개 모드에 대한 평가 과정을 생략하게 된다. 후보 모드는 현재 블록의 주변 블록 정보로부터 얻을 수 있는 예측 모드들 및 각도를 대표하는 예측 모드들로 구성된다.

본 논문의 구성은 다음과 같다. 2 절에서는 HEVC 표준의 화면내 예측 및 부호화 과정을 살펴본 후, 3 절에서는 제안하는 후보 화면내 예측 모드를 결정하는 방법을 설명한다. 4 절에서는 제안 방법의 성능을 실험을 통해서 확인하고 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

2. HEVC 화면내 예측 및 부호화 과정

2.1 화면내 예측 모드

HEVC 표준은 화면내 부호화 효율을 높이기 위해 총 35 개 예측 모드를 사용한다. HEVC 의 화면내 예측 모드는 표 1 처럼 0 번부터 34 번까지 총 35 개로 정의된다[1]. 0 번과 1 번은 각도가 없는 Planar 및 DC 모드를 표시하고, 2 번부터 34 번까지의 angular 예측 모드는 다양한 예측 방향성을 지원하는 각도를 가진 모드를 표시한다. 그림 1 은 33 개의 각도를 가진 예측 모드의 방향을 보여준다.

HEVC 화면내 예측 및 부호화 과정은 상기 언급한 총 35 개 모드에 대한 평가 과정을 거쳐 최종 예측 모드를 결정하게 된다.

표 1. 화면내 예측 모드

Intra prediction mode	Associated names
0	Intra_Planar
1	Intra_DC
Otherwise (2..34)	Intra_Angular

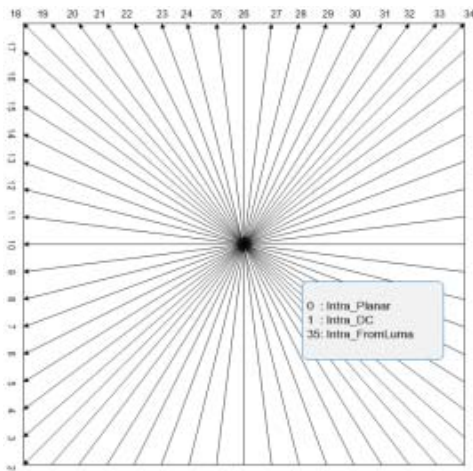


그림 1. 각도를 가진 예측 모드 및 예측 방향

2.2 최종 예측 모드 결정 방법

다수의 화면내 예측 모드를 평가하여 최종 하나의 화면내 예측 모드를 결정하기 위해 RDO(rate distortion optimization) 평가 방법을 사용한다. 최종 예측 모드를 결정하는 과정은 RDO 평가 방법 및 평가할 예측 모드의 개수를 기준으로 두 단계로 나눌 수 있다[4].

첫 번째 단계에서는 총 35 개 예측 모드에 대한 간단한 RDO (SRDO) 평가를 실시한다. 이때 사용하는 SRDO 비용은 식(1)와 같다. 식(1)을 살펴보면 왜곡 측정 방법으로 SATD(sum of absolute Hadamard transformed difference)를, 비트 측정 방법으로 예측 모드를 부호화한 비트를 사용한다.

$$J_{SRDO} = SATD + \lambda_{pred} \times R_{pred} \quad (1)$$

첫 번째 단계를 통해 35 개 모드 중 PU(prediction unit) 크기에 따라 미리 정해진 수만큼의 후보 모드를 추린다. 표 2 는 PU 크기에 따른 후보 모드 개수를 나타낸다.

표 2. PU 크기에 따른 후보 화면내 예측 모드 개수

PU 크기	후보 모드 개수
4x4	8
8x8	
16x16	3
32x32	
64x64	

두 번째 단계는 첫 단계 결과물인 후보 모드들에 대해서 정칙 RDO (FRDO) 평가를 실시한다. 이때 사용하는 FRDO 비용은 식(2)와 같다. 식(2)를 살펴보면 왜곡 측정 방법으로 SSD(sum of squared difference)를, 비트 측정 방법으로는 예측 모드로 해당 블록을 부호화한 비트를 사용한다. 결과적으로, 두 번째 단계에서는 첫 번째 단계보다 정확한 RDO 비용을 계산하여 현재 PU 에서의 최종 화면내 예측 모드를 결정하게 된다.

$$J_{FRDO} = SSD + \lambda_{mod} \times R_{mod} \quad (2)$$

2.3 화면내 예측 모드 부호화 방법

현재 PU 의 최종 화면내 예측 모드를 부호화 하기 위하여 총 35 개의 예측 모드 중에서 주변 블록의 예측 모드와 확률적으로 자주 사용되는 예측 모드를 기반으로 현재 PU 를 위한 3 개의 MPM(most probable mode)을 설정한다. 현재 PU 의 화면내 예측 모드가 설정된 3 개의 MPM 중 하나와 같은 경우 1-2 비트를 사용하여 MPM 인덱스 값을 부호화하고, 3 개의 MPMs 가 같지 않은 경우 5 비트의 고정길이 부호화로 MPM 과 일치하지 않는 나머지 32 개의 모드 중 해당하는 하나의 화면내 예측모드를 부호화 한다.

MPM 을 설정하는 과정은 현재 PU 의 상측 및 좌측에 인접한 블록의 화면내 예측 모드를 가져와 MPM 으로 설정하고, 두 개의 MPM 이 동일하지 않은 경우 세 번째 MPM 으로 자주 사용되는 Planar, DC, Vertical 모드 순으로 중복되지 않은 모드를 설정한다. 반면에 상측 및 좌측의 블록으로부터 정해진 MPM 두 개가 동일한 경우, 동일한 MPM 이 각도를 가진 모드이면 해당 각도와 바로 인접한 각도에 대한 예측 모드들로 나머지 두 개의 MPMs 을 설정하고, 동일한 MPM 이 각도를 가지지 않은 모드이면, Planar, DC, Vertical 모드로 세 개의 MPMs 을 설정한다.

3. 제안하는 화면내 예측 모드 결정 방법

본 논문은 최대 35 개의 화면내 예측 모드 중 현재 PU 의 주변 블록 정보로부터 얻을 수 있는 예측 모드들 및 각도를 대표하는 예측 모드들을 선별적으로 추려서 후보 예측 모드를 결정하고, 평가 과정을 거쳐 해당 후보 모드 중에서 최종적으로 현재 PU 에 대한 화면내 예측 모드를 결정한다.

제안 방법은 그림 2 와 같이 각도를 대표하는 5 개의 예측 모드들 (모드 번호: 2, 10, 18, 26, 34)과 각도가 없는 2 개의 예측 모드들 (모드 번호: 0, 1)로 7 개의 일차 후보 예측 모드를 결정한다. 일차 후보 모드들에 대한 SRDO 평가 과정 후 선정된 최선의 모드가 각도 모드 중 하나이면 해당 각도와 인접한 각도들에 대한 평가를 3 번 반복하여 처리한다. 일례로 26 번 모드가 최선의 모드로 선정이 된다면, 그림 3 과 같이 붉은색 26 번 모드로부터 4 만큼 거리가 떨어진 파란색 22 번 모드와 30 번 모드를 SRDO 로 평가하여 둘 중에 SRDO 비용이 적은 하나의 모드 (22 번)을 일차 후보 모드로 추가한다. 다음으로 22 번 모드로부터 2 만큼 거리가 떨어진 주황색 20 번 모드와 24 번 모드를 SRDO 로 평가하여 하나의 모드 (24 번)을 일차 후보 모드로 추가한다. 마지막으로 24 번 모드로부터 1 만큼 거리가 떨어진 녹색 23 번 모드와 25 번 모드 중에 SRDO 비용이 적은 하나의 모드를 추가함으로써 최대 3 개의 모드를 일차 후보 모드로 등록한다.

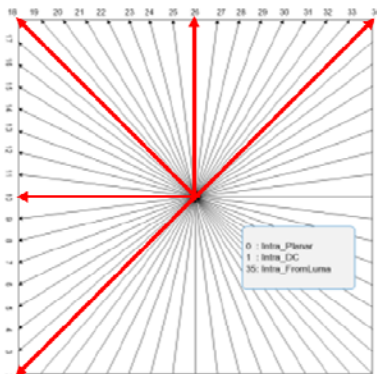


그림 2. 각도를 대표하는 예측 모드

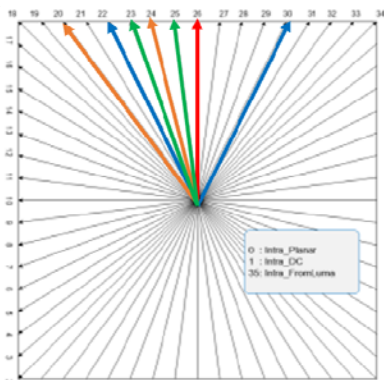


그림 3. 인접한 각도 모드

또한, 현재 PU 의 주변 블록 정보로부터 얻을 수 있는 예측 모드 (즉, 현재 PU 의 화면내 예측 모드를 부호화 하기 위하여 설정한 세 개의 MPMs)들을 일차 후보 모드로 추가한다. 이때, 기존 일차 후보 모드들과 중복되지 않는 모드만을 등록한다. 이로써 최종 일차 후보 모드 선택이 마무리된다.

지금까지 등록된 일차 후보들 중에서 이차 후보 모드를 결정하게 된다. 이차 후보 모드의 개수는 표 2 의 PU 크기에 따른 후보 모드 개수에 따른다. 일차 후보 모드들에 대한 SRDO 비용을 기준으로 이차 후보 모드들이 결정된다. 선택된 이차 후보 모드들은 FRDO 평가를 거쳐 현재 PU 의 화면내 예측 모드로 최종 하나의 모드가 결정된다. 제안 방법의 전체 흐름도는 그림 4 와 같다.

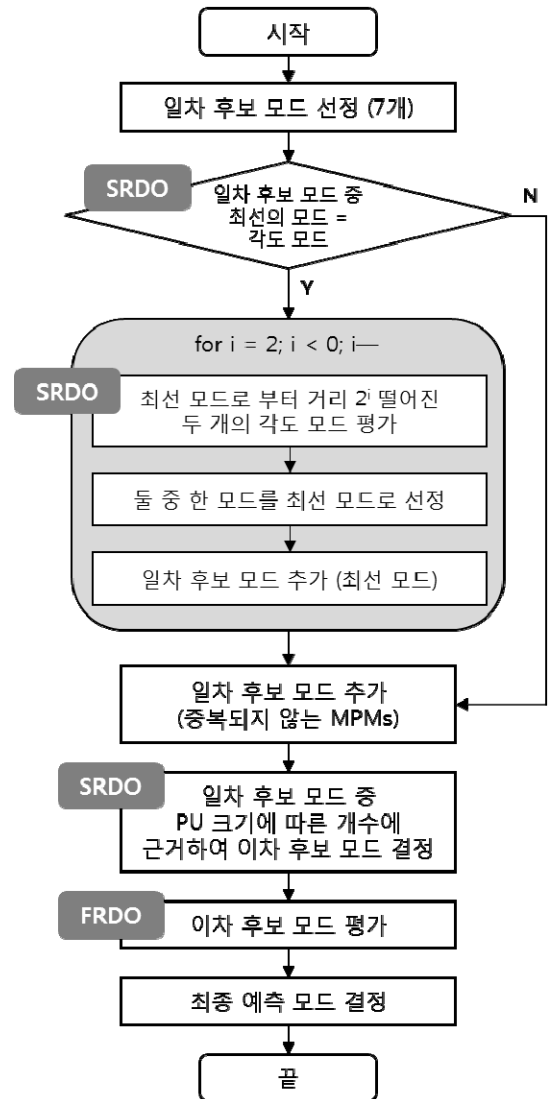


그림 4. 제안 방법의 흐름도

4. 실험 결과

제안하는 화면내 예측 모드 결정 방법은 본사에서 설계한 HEVC main profile 부호화기에 구현하였다. 제안 방법의 평가를 위해, 본사 부호화기의 화면내 예측 및 부호화 시에 총 35 개의 화면내 예측 모드를 모두 사용한 경우와 제안한 후보 예측 모드만을 사용한 경우를 실험하였다. 두 방법의 정확한 성능 비교를 위해 화면내 예측 모드의 개수 이외에는 동일한 실험 조건을 유지하였다.

실험에 사용한 시퀀스는 HEVC 표준의 FHD (1920x1080) 실험 영상인 Kimono, ParkScene, BQTerrace, BasketballDrive, Cactus이며, All-Intra 구조에서 실험하였고, Qp 값은 22, 27, 32, 37 사용하여 전체 구간에 대한 BD-rates 및 부호화기 복잡도를 측정하였다.

표 3 은 기존 방법 대비 제안 방법의 BD-rates 와 부호화기 복잡도를 보인다. 기존 방법은 35 개의 모드에 대해서 SRDO 평가 후, 일부 후보 모드에 대한 FRDO 비용을 계산한다. 즉, 기존 방법이 총 35 개의 모드에 대한 FRDO 평가를 하지 않으므로, 기존 방법 대비 제안 방법의 BD-rate 가 더 좋아지는 예외적인 상황이 발생했음을 Kimono 와 ParkScene 에서 볼 수 있다.

전반적인 실험 결과를 통해, 제안 방법이 All-Intra 부호화 구조에서 기존 방법 대비 1.1%의 BD-rate 증가와 함께 18.7%의 부호화기 복잡도를 감소 시켰음을 알 수 있다. 또한, 제안 방법이 BasketballDrive 및 BQTerrace 시퀀스와 같이 움직임이 크거나 복잡한 화면을 포함한 동영상에 상대적으로 더 큰 효과를 보이는 것으로 확인했다.

표 3. BD-rates 및 부호화기 복잡도 (Enc. Time)

Sequence	BD-Rates			Enc. Time
	Y	U	V	
BasketBallDrive	2.0%	2.6%	2.5%	-29.0%
BQTerrace	3.5%	4.2%	4.6%	-32.5%
Cactus	1.1%	1.1%	1.4%	-31.3%
Kimono	-0.4%	0.3%	0.6%	-0.3%
ParkScene	-0.5%	-0.2%	0.0%	-0.4%
All	1.1%	1.6%	1.8%	-18.7%

5. 결론

본 논문에서는 HEVC 표준의 화면내 부호화 과정의 복잡도를 감소 시키기 위해 총 35 개의 화면내 예측 모드 중 유력한 후보 예측 모드들을 선별하였다. 제안 방법은 현재 PU 의 주변 블록 정보로부터 얻을 수 있는 예측 모드들과 각도를 대표하는 예측 모드들을 선별적으로 추려서 후보 예측 모드를 구성하였다. 총 35 개 모드를 모두 평가하는 기존 방법에 비해 유력 후보 모드만을 평가하는 제안 방법은 실험 결과를 통해 약 1.1%의 BD-rate 를 증가시키면서 약 18.7%의 부호화기 수행 시간을 감소 시켰음을 증명하였다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 방송통신미디어산업원천기술개발사업의 일환으로 수행하였음. [10043450, 8K UHD 및 4K S3D(stereoscopic 3D) 콘텐츠의 획득/저장/Ingest 및 전송용 비디오 서버 기술 개발]

참고 문헌

- [1] B. Bross, W. Han, J. Ohm, G. Sullivan, Y. Wang, and T. Wiegand, "High efficiency video coding (HEVC) text specification draft 10(for FDIS & Consent)," JCTVC-L1003, Jan. 2013.
- [2] "Call for Evidence on High-Performance Video Coding (HVC)," N10553, ITU-T Q6./16 Visual Coding and ISO/IEC JCT1/SC29/WG11 Coding of Moving Pictures and Audio, 2009(4).
- [3] G. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649-1668, Dec. 2012.
- [4] Yinji Piao, Junghye Min, and Jianle Chen, "Encoder improvement of unified intra prediction," JCTVC-C207, Oct. 2010.