

비디오 코딩을 위한 다중 잔차신호 플리핑 및 변환 방법

김남옥, *강정원, *임성창, **이영렬
 세종대학교, *한국전자통신연구원, **세종대학교

nukim@sju.ac.kr, jungwon@etri.co.kr, sclim@etri.co.kr, **yllee@sejong.ac.kr

Multiple Transform and Residual Flipping for Video Coding

Nam uk kim, Jungwon kang, Sung-Chang lim **Yung-Lyul Lee
 sejong university *sejong university

요 약

비디오 압축에서, 변환은 데이터를 공간 영역에서 주파수 영역으로 변환 함으로써 에너지 압축에서 중요한 역할을 수행한다. HEVC(High Efficiency Video Coding)에서는 DCT-II(Discrete Cosine Transform type-II)를 사용하여 잔차신호 변환을 수행한다. DCT-II 는 픽셀간 상관도가 높은 신호일수록 높은 에너지 집중도를 보이지만, 픽셀간 상관도가 비교적 낮은 블록일수록 낮은 에너지 집중도를 보인다. 본 논문에서는 DST-VII(Discrete Sine Transform type-VII) 및 잔차신호 플리핑을 사용하여 다양한 변환 방법으로 영상을 부호화 및 복호화 하는 알고리즘에 대해 제안한다. 다양한 변환 방법은 부호화기에서 블록단위로 1 가지를 선택하여 비트스트림으로 선택된 방법에 대한 정보를 전송한다. 제안된 방법은 HEVC 대비 약 2.47%의 BD-rate 감소를 보인다.

1. 서론

HEVC(High Efficiency Video Coding) [1]은 가장 최신의 비디오 부호화 표준 기술로 ITU-T 와 VCEC 에서 공동으로 표준화를 진행하여 2013 년도 표준화가 완료되었다. HEVC 는 영상을 우선 CTU(Coding Tree Unit)라는 64×64 크기의 정사각형 블록으로 분할한다. CTU 들은 다시 쿼드트리 구조로 CU(Coding Unit)로 분할된다. CU 는 화면 내 예측을 사용할 것인지, 화면 간 예측을 사용할 것인지 정하는 최소단위이다. CU 는 다시 TU(Transform Unit)로 분할될 수 있다. TU 는 변환을 수행하는 기본 단위이다. CU 는 또한 PU(Prediction Unit)로 분할될 수 있다. PU 는 예측을 수행하는 기본단위이다. HEVC 의 압축과정은 예측, 변환, 양자화, 엔트로피 부호화 순서로 구성된다. 예측은 주변으로부터 현재 블록을 예측하는 과정으로 화면 내 예측, 화면 간 예측 두 가지 방법이 존재한다. 두 가지 예측 방법 중 하나를 선택해 예측신호를 생성한 후 원본신호로부터 빼주어 잔차신호를 생성한다. 이후 변환 양자화를 통해 주파수도메인에서 압축을 수행하며 변환 양자화과정에서 출력되는 변환계수들은 양자화되어 엔트로피 부호화를 통해 비트 스트림으로 변환된다. HEVC 는 화면 내 예측의 잔차신호에 대해서는 8×8 크기 이상의 TU 인 경우는 DCT-II(Discrete Cosine type-II) [2], 4×4 크기의 TU 인 경우는 DST-VII(Discrete Sine Transform type-VII)를 사용하여 변환을 수행한다. 잔차신호들의 상관관계가 비교적 높은 큰 크기의 TU 에서는 DCT-II 변환이 효율적이며, 잔차신호들의 상관관계가 비교적 낮은 4×4 크기의 TU 에서는 DST-VII 변환이 효율적이다. 하지만 이러한 변환기저 선택은 평균모델을 기반으로 하였으며, 실제 잔차신호들은 임의적인 특성이 높은 편이기 때문에 아닌 경우가 많이 발생한다.

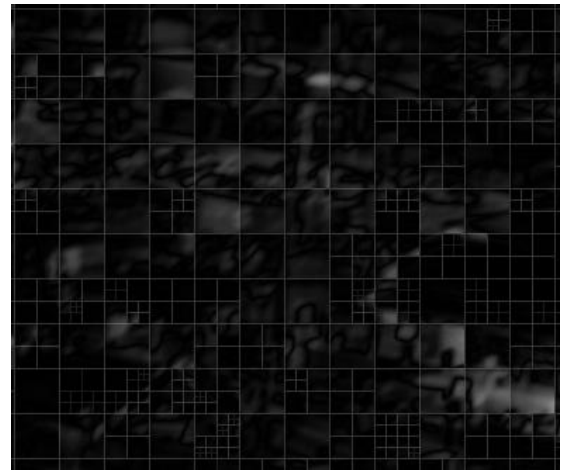


그림 1. RollerCoaster(2160p) 영상의 잔차신호

그림 1 은 3840×2160 해상도인 RollerCoaster 영상의 잔차신호의 일부이다. HEVC 로 부호화 되었으며 흰 선들은 CU 분할 구조를 표현하며 보다 열은 흰 선들은 TU 분할구조를 표현한다. 잔차신호는 절대값이 클수록 밝은색으로 표현되며 절대값이 작을수록 어두운 색으로 표현된다. 그림에서 보면 잔차신호들이 CU, TU 경계의 특정한 모서리나, 꼭지점에 많이 몰려 있는 것을 확인 할 수 있다. 대부분 우측하단의 꼭지점에 많이 몰려 있지만 다른 방향에 몰려있는 경우도 발생하는 것을 확인 할 수 있다. 이러한 경우에는 DCT-II 를 사용하여 변환을 하는 것보다는 DST-VII 와 같은 비대칭 변환 기저를 사용하는

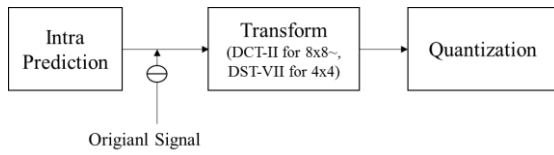


그림 2. HEVC의 예측 부호화 과정

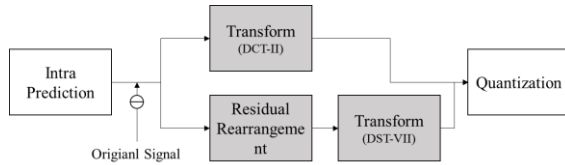


그림 3. 제안하는 예측 부호화 과정

것이 효율적이다. 본 논문에서는 잔차신호들이 한 곳으로 몰려 있는 특성을 고려하여 화면 내 예측된 TU에 대해 변환을 수행할 때 DCT-II와 DST-VII 중 하나를 선택하여 변환을 수행하는 방법을 제안한다. 그리고 잔차신호가 집중되는 위치를 맞추기 위해 DST-VII로 변환을 수행하기 전에 잔차신호를 플리핑 하는 방법을 제안한다. 본 논문의 2 장에서는 제안하는 방법에 대해 설명하고, 3 장에서는 실험 결과 그리고 마지막 4 장에서는 결론으로 이 논문을 마친다.

2. 제안하는 방법

그림 2는 기존의 HEVC 예측 및 부호화 과정을 보여준다. 예측된 신호를 원본 신호로부터 빼 주어 잔차신호들을 생성하고, 그 잔차신호들에 대해 TU 단위로 변환을 수행하는데 8x8 크기 이상의 TU에 대해서는 DCT-II 2차원 변환을 수행하고, 화면 내 예측이며 4x4 크기의 TU에 대해서는 DST-VII 2차원 변환을 수행한다. 그리고 변환된 계수를 양자화와 엔트로피 부호화를 통해 비트 스트림으로 변환한다. 그림 3은 제안하는 HEVC 예측 및 부호화 과정을 보여준다. 제안하는 방법은 음영된 블록으로 표시하였다. 변환을 수행할 때 TU의 크기와 상관없이 DCT-II 이차원 변환과 DST-VII 이차원 변환 중 하나를 선택하여 변환을 수행한다. 그리고 DST-VII 이차원 변환을 수행할 때는 변환 전에 잔차신호들의 재정렬을 수행한다. 잔차신호 재정렬은 잔차신호들을 신호의 세기가 큰 영역을 TU의 오른쪽의 아래의 모서리로 이동시키기 위해 사용된다.

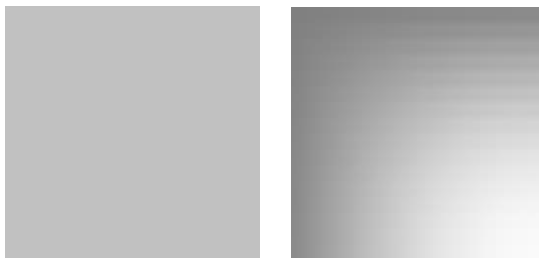


그림 4. DCT-II(왼쪽)와 DST-VII(오른쪽) 이차원 변환기저의 첫번째 기저

그림 4는 DCT-II 이차원 변환기저의 첫번째 기저와 DST-VII 이차원 변환기저의 첫번째 기저를 보여준다. DCT-II는 그림과 같이 평균적인 특성을 지니기 때문에 잔차신호들이 TU 영역에 전체적으로 고르게 분포할 경우 높은 성능을 보일 것이다.

반면에 DST-VII은 잔차신호들이 TU의 좌측과 위쪽에 적게 분포하고, 오른쪽과 아래로 갈수록 많이 분포할 경우 높은 압축 성능을 보일 것이다. HEVC는 화면 내 예측에 참조하는 복원 신호들은 현재 블록의 왼쪽과 위쪽에 있기 때문에 좌측과 위쪽 부분의 예측오차의 확률이 작고 좌측과 위쪽으로부터 멀어질수록 오차의 확률이 점점 증가할 것이다. DST-VII은 이러한 특성과 유사한 기저 모양을 가지기 때문에 이러한 경우 효율적인 변환 기저가 될 수 있다. 하지만 그림 1에서 보인 예시와 같이 잔차신호가 집중되는 위치가 다른 경우도 발생할 수 있다. 이러한 경우 잔차신호가 집중되는 위치를 DST-VII과 같은 우측하단으로 이동해 줌으로써 효율적으로 변환을 수행할 수 있다. 이러한 과정을 잔차신호 재배열이라고 하며 본 논문에서는 4가지 재배열 방법을 제안한다. 4가지 재배열 방법은 다음과 같다. 재배열 없음, 가로플립, 세로플립 그리고 가로세로플립이다. 재배열 없음은 말 그대로 재배열을 수행하지 않는다. 가로 플립은 잔차신호를 가로 방향으로 반전시킨다. 세로 플립은 잔차신호들을 세로 방향으로 반전시킨다. 가로세로플립은 잔차신호들을 가로 및 세로 방향으로 반전시킨다. 각각 4가지 재배열 방법들을 수식으로 표현하면 다음과 같다.

재배열 없음 : $p'(x, y) = p(x, y)$

가로 플립 : $p'(x, y) = p(nT - 1 - x, y)$

세로 플립 : $p'(x, y) = p(x, nT - 1 - y)$

가로세로 플립 : $p'(x, y) = p(nT - 1 - x, nT - 1 - y)$

$p(x, y)$ 는 재배열 이전의 잔차신호들이며 $p'(x, y)$ 는 재배열된 잔차신호들이다. nT 는 TU의 너비 및 높이에 해당한다.

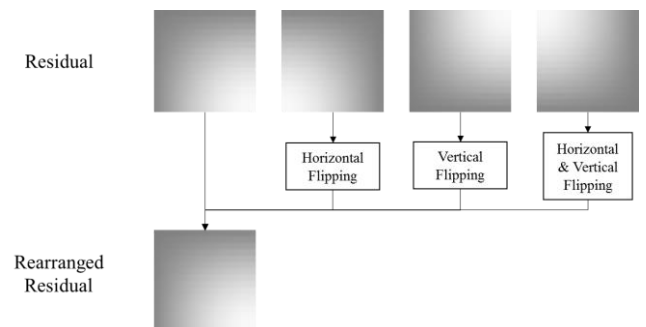


그림 5. 제안하는 4가지 재배열 방법

그림 5는 4가지 재배열 방법을 보여준다. 4가지 잔차신호들은 각각 4가지 재배열 방법들의 이상적인 잔차신호 입력을 의미한다. 각각의 입력된 잔차신호들은 잔차신호 재배열 과정을 통해 우측 하단에 집중되게 된다. 그리고 재배열된 잔차신호들은 DST-VII를 이용하여 변환을 수행한다. 부호화기에서 RD 비용 비교를 통해 4가지 재배열 방법 중 한가지 방법이 선택되며, 선택 정보는 복호화기로 전송된다.

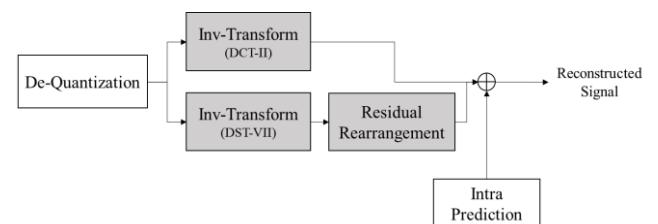


그림 6. 제안하는 복호화 과정

그림 6 는 제안하는 화면 내 예측 복호화 과정에 대한 도식이다. 역양자화 이후 부호화기로부터 변환정보를 획득하여 DCT-II 또는 DST-VII 으로 역변환을 수행한다. DST-VII 으로 역변환을 수행한 경우 잔차신호 재배열을 통해 잔차신호를 복원한다. 본 논문에서 제안한 재배열 방법은 재배열과 역재배열이 동일한 프로세스를 갖는다.

3. 실험 결과

제안하는 방법은 HEVC 참조 소프트웨어 HM16.6[3] 에서 구현되었으며, HEVC 공통 실험 조건[4]의 Intra, high efficiency, 10 bit 조건에서 실험하였다. 테스트에 사용한 영상은 JVET 테스트 영상들이며 표 1 에 클래스 별로 정리하였다. 모든 테스트 영상들은 8 프레임 간격으로 건너뛰며 부호화되었다. 양자화 파라미터는 22, 27, 32, 37 을 사용하였다.

표 1. 테스트 영상

클래스	시퀀스	해상도	비트심도	프레임율
A1	Tango	4096x2160	10	60 fps
A1	Drums100	3840x2160	10	100 fps
A1	CampfireParty	3840x2160	10	30 fps
A1	ToddlerFountain	4096x2160	10	60 fps
A2	CatRobot	3840x2160	10	60 fps
A2	TrafficFlow	3840x2160	10	30 fps
A2	DaylightRoad	3840x2160	10	60 fps
A2	Rollercoaster	4096x2160	10	60 fps
B	Kimono	1920x1080	8	24 fps
B	ParkScene	1920x1080	8	24 fps
B	Cactus	1920x1080	8	50 fps
B	BasketballDrive	1920x1080	8	50 fps
B	BQTerrace	1920x1080	8	60 fps
C	BasketballDrill	832x480	8	50 fps
C	BQMall	832x480	8	60 fps
C	PartyScene	832x480	8	50 fps
C	RaceHorses	832x480	8	30 fps
D	BasketballPass	416x240	8	50 fps
D	BQSquare	416x240	8	60 fps
D	BlowingBubbles	416x240	8	50 fps
D	RaceHorses	416x240	8	30 fps
E	FourPeople	1280x720	8	60 fps
E	Johnny	1280x720	8	60 fps
E	KristenAndSara	1280x720	8	60 fps
F	BasketballDrillText	832x480	8	50 fps
F	ChinaSpeed	1024x768	8	30 fps
F	SlideEditing	1280x720	8	30 fps

F	SlideShow	1280x720	8	30 fps
---	-----------	----------	---	--------

표 2. 실험 결과

	All Intra Main 10 Over HM-16.6			
	Y	U	V	EncT
Class A1	-3.19%	-0.57%	-0.83%	158%
Class A2	-3.17%	-1.45%	-1.46%	155%
Class B	-2.58%	-0.73%	-0.71%	163%
Class C	-2.29%	0.00%	-0.07%	169%
Class D	-1.96%	0.05%	0.30%	173%
Class E	-2.46%	-0.73%	-0.81%	156%
Overall	-2.61%	-0.57%	-0.59%	162%
Class F(optional)	-2.23%	-0.55%	-0.68%	163%

표 2 는 제안하는 방법의 실험 결과를 보인다. 제안하는 방법은 HM16.6 과 성능을 비교하였으며 BD-rate(Bjontegaard Distortion rate) [5]를 통해 성능 차이를 계산하였다. Y,U,V 는 각각 색 성분들의 BD-rate 를 의미하며 EncT 는 HM-16.6 대비 제안하는 알고리즘의 부호화 시간 비율, DecT 는 HM-16.6 대비 제안하는 알고리즘의 복호화 시간 비율을 의미한다. 고해상도 영상일수록 제안하는 알고리즘은 높은 성능을 보이며 해상도가 낮아질수록 제안하는 알고리즘의 성능 향상 폭이 줄어드는 것을 확인할 수 있다. 고해상도 영상에서는 픽셀간의 상관관계가 비교적 높기 때문에 큰 크기의 블록이 많이 발생한다. 그리고 제안하는 방법은 큰 블록일수록 성능 향상이 높게 나타나기 때문에 해상도에 따라 성능차이가 발생하는 것으로 보인다. 부호화 시간은 잔차신호 재배열의 RD 비용 계산량 증가로 인해 HEVC 대비 60% 정도의 계산시간이 증가하였다.

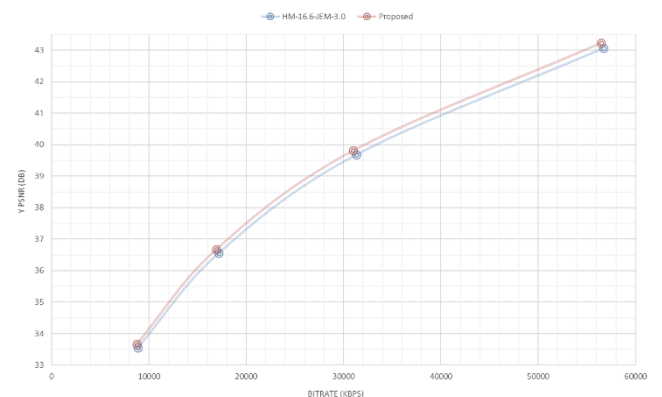


그림 7. ToddlerFountain 의 PSNR 곡선

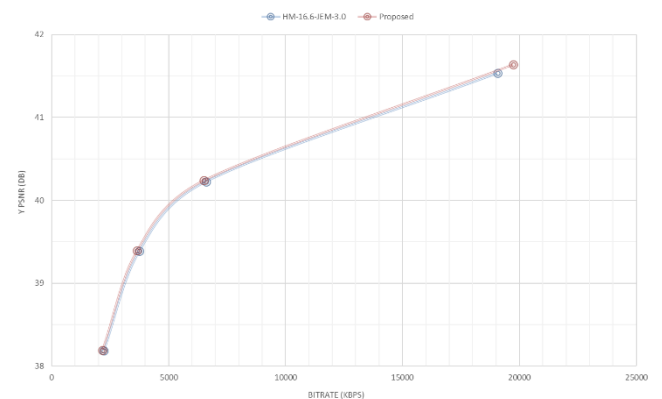


그림 8. Tango 의 PSNR 곡선

그림 7 과 그림 8 은 각각 ToddlerFountain 과 Tango 에서의

PSNR-BitRate 곡선을 보여준다. 제안하는 방법의 PSNR 곡선은 HM16.6 PSNR 곡선보다 항상 위에 있기 때문에 모든 비트율에서 제안하는 방법의 성능이 HEVC 보다 항상 높음을 확인할 수 있다. 그리고 모든 테스트 영상에서 성능 향상을 보이므로 제안하는 방법은 안정적인 성능향상을 얻어낼 수 있다고 볼 수 있다.

표 3. Class A0~E 의 변환 선택 비중 평균

TU	DCT-II	DST-VII (No Flip)	DST-VII (H-Flip)	DST-VII (V-Flip)	DST-VII (HV-Flip)
4x4	24.6%	36.3%	7.8%	27.6%	3.6%
8x8	34.6%	39.1%	8.8%	13.5%	4.0%
16x16	40.2%	37.2%	8.6%	10.8%	3.2%
32x32	62.4%	23.7%	5.8%	6.2%	2.0%

표 4. Class F 의 변환 선택 비중 평균

TU	DCT-II	DST-VII (No Flip)	DST-VII (H-Flip)	DST-VII (V-Flip)	DST-VII (HV-Flip)
4x4	62.0%	11.2%	2.8%	23.0%	1.0%
8x8	64.4%	21.2%	3.4%	10.1%	0.9%
16x16	63.1%	26.1%	3.1%	6.7%	0.9%
32x32	78.6%	14.4%	2.3%	4.1%	0.6%

표 3 은 Class A0 부터 Class E 까지의 영상을 제안하는 방법으로 부호화 하였을 경우 변환 종류가 선택되는 비중의 평균을 나타낸다. DCT-II 는 큰 블록일수록 선택되는 비중이 높아지는 특성을 보인다. 반면에 DST-VII 는 작은 블록일수록 선택되는 비중이 높아지는 것을 확인할 수 있다. 잔차신호를 재배열 방법은 재배열 하지 않음, V-Flip, H-Flip, HV-Flip 순으로 선택되는 확률이 높다. H-Flip 과 HV-Flip 의 선택 비중이 합쳐서 10% 내외로 적은 편이지만 해당 두 재배열 방법은 1% 정도의 BD-rate 성능향상을 가져온다.

표 4 는 스크린 콘텐츠인 Class F 에 대한 변환 종류 선택 비중의 평균을 나타낸다. 다른 클래스들과 다르게 DCT-II 의 선택 비중이 매우 높게 나타났다. 잔차신호 재배열의 선택확률은 다른 클래스들과 비슷하게 나타났다.

4. 결론

제안하는 DST-VII 과 잔차신호 재배열을 이용한 변환방법은 고해상도 영상에 대해 HEVC 대비 평균적으로 3.2% 그리고 전체 테스트 영상에 대해 평균 2.6%의 BD-rate 향상을 보인다.

감사의 글

이 논문은 일부 2017 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. 2016-0-00572, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2 배 압축을 제공하는 5 세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)

참 고 문 헌

- [1] B. Bross, W.-J. Han, J. -R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand, High Efficiency Video Coding (HEVC) Text Specification Draft 10 (for FDIS & Consent), JCT-VC document JCTVC-L1003, vol. 1, 12thmeeting, Geneva, Switzerland, 2013.
- [2] V. Britanak, P. Yip, and K. R. Rao, "Discrete Cosine and Sine Transforms – General Properties, Fast Algorithms, and Integer Approximations," 1st Edition, Academic Press, Inc, San Diego, 2006.
- [3] HM reference software HM-16.6, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/, Jun. 2015.
- [4] JCT-VC document, JCTVC-L1100, Common HM Test conditions and software reference configurations, Geneva, Switzerland, Jan. 2013
- [5] G. Bjøntegaard, Calculation of Average PSNR Differences Between RDCurves, ITU-T Q.6/SG16, Doc. VCEG-M33, 2001

색차 채널을 위한 적응형 다중변환기술 성능분석

박지윤, 전병우

성균관대학교 전기전자컴퓨터공학과

{jiyoonpark, bjeon}@skku.edu

Analysis of Adaptive Multiple Transform in Chroma Channel

Jeeyoon Park and Byeungwoo Jeon

Department of Electrical and Computer Engineering

Sungkyunkwan University

요 약

최근 UHD 콘텐츠에 대한 스트리밍 서비스가 증가함에 따라 보다 높은 압축효율을 갖는 부호화 기술에 대한 필요성이 증가하고 있으며 이에 따라 차세대 비디오 코덱을 위한 국제표준화 노력이 JVET 를 중심으로 진행되고 있다. 본 논문에서는 적응형 다중변환 기술을 휘도 채널에 적용하는 것뿐만 아니라 동일한 방법으로 화면 내/간 블록 모두에서 색차채널에도 적용하도록 하여 그 효율 향상여부를 분석하였다. 실험결과 제안방법이 기존 방법 대비 BDBR 측면에서 $Y(0.03\%)$, $U(-1.40\%)$, $V(-0.96\%)$ 의 결과를 얻을 수 있었다.

1. 서론

2013 년 HEVC 국제표준화를 2013 년 완료 후 [1], 다시 HEVC 대비 2 배의 압축률을 목표로 하는 차세대 비디오 부호화 (Future Video Coding; 이하 FVC) 표준화가 ISO/IEC MPEG (JTC1/SC29/WG11)과 ITU-T VCEG (Q6/16)의 공동 비디오 부호화 팀 (Joint Video Exploration Team; 이하 JVET) 주관으로 진행중이다. 한편 JVET 의 공동 실험을 위하여, 부호화 효율을 증가시키는 주요 비디오 부호화 기술들을 선택하고 이들을 SW 로 집약한 Joint Exploration Model (JEM)이 개발되었는데, 최신버전인 JEM7.2 은 YCbCr 4:2:0 이외의 컬러 포맷과 10bit 이상의 bit depth 를 지원한다. 또한, JEM7.2 에는 적응형 다중변환 (Adaptive multi core transform, AMT) 기술이 추가되었는데, 윌-왜곡(RDO: Rate Distortion Optimization)과정으로 인해 상대적으로 많은 부호화 시간이 요구됨으로, 시간 복잡성 및 계산량을 줄이기 위하여 화면 내/간 예측 모두 휘도 채널에만 적용되고 있다. 즉, 색차 채널에는 HEVC 와 같은 고정된 변환이 사용하고 있다[1]. 이를 개선하기 위해 [2]에서는 JEM3.0 버전에서 기존에 휘도 블록에만 적용되었던 AMT 를 색차신호 변환 블록에 추가 신호 없이 그대로 적용하는 방법을 제안한 바가 있다. 그러나 이 제안방법에서는 화면 내 예측블록의 색차신호에 대한 AMT 는 고려하지 않고 화면 간 예측블록의 경우에만 적용하였다. 본 논문에서는 휘도 성분에 적용된 AMT 변환을 JEM7.2 버전에서 화면 내/간 예측 모두 색차신호에도 적용할 경우 화면 내/간 예측블록에서 색차채널 코딩 효율성 향상여부를 분석한다.

2. 적응형 다중변환

적응형 다중 변환은 차세대 비디오 압축 알고리즘의 표준화를 위해 진행되어온 JVET 회의에서 채택된 방법으로, 휘도성분의 코딩효율성 향상에 상당한 기여를 하고 있다. 1 차 변환과정에서 DCT-II 및 4×4 DST-VII 만 사용하던 HEVC 와는 달리, DST-VII, DCT-VIII, DST-I 및 DCT-V 의 새로운 변환방식이 추가되어, 화면 내/간 예측모드에 따라 특성화된 최적의 변환을 찾아 잔차신호 부호화에 사용된다. AMT 가 적용되는 CU 내의 휘도 채널 코딩블록(Coding Block: CB)의 경우 AMT 신호 플래그가 추가되어 사용할 수평 및 수직변환 조합 집합을 지시한다. 또한, 이 정해진 조합 내에서 윌-왜곡(RDO)과정을 통해 특정한 최적의 수평 및 수직변환을 결정한다. 화면 내 잔차신호 부호화에 있어서, 화면 내 예측 모드별로 상이한 잔차신호 통계로 인해, 예측모드 별로 다른 변환 후보가 사용된다. 그러나 HEVC 표준에서 윌-왜곡(RDO)과정은 좋은 압축 성능을 보이는 반면, 상대적으로 많은 부호화 시간이 요구된다. 따라서 최신버전인 JEM7.2 에서는 화면 내/간 휘도 블록에서만 적용되고 있다.

3. 실험방법 및 분석

JEM 에서 색차 신호는 HEVC 에서 사용되던 전통적인 5 개 화면 내 예측 모드와 채널 간 중복성(Redundancy)을 줄이기 위한 CCLM (cross-component linear model) 예측 모드 6 개, 즉, 총 11 개의 화면 내 예측 모드를 가질 수 있다 [3]. 6 가지의 CCLM 모드에는 CCLM 1 모드, MMLM(MMLM: Multi model LM)모드 그리고 4 종류의 MFLM MFLM(Multi filter LM)으로 구성된다. 휘도 화소로부터 색차 화소값을 예측하기

¹ 이 논문은 2017 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No. 2016-0-00572, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2 배 압축을 제공하는 5 세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)

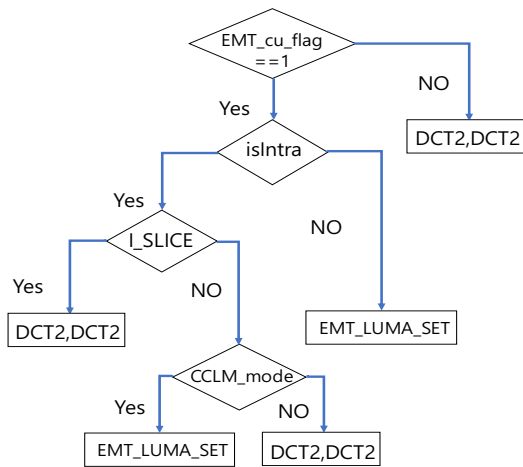


그림 1. 색차신호 변환에 적용한 실험방법.

위해 전체 CU 를 대표하여 하나의 선형 모델을 사용하는 것이 CCLM1 모델이라면, MMLM 에서는 두개의 선형 모델을 사용할 수 있다. 또한 4 종류의 MFLM 모드들은 각각 색차 신호의 하향표본화(down sampling)할 때 사용되는 필터로 구분된다. Cross component 란 휘도 신호와 색차신호 사이의 중복성을 의미하는데, 이러한 연관성(correlation)을 추가로 제거하는 6 개의 CCLM 모드 중 한 개가 색차 예측 모드로 선택된 화면 내 예측 블록에서는 휘도 신호와 색차신호의 연관성이 상대적으로 깊다고 할 수 있다. 따라서 휘도 신호의 다중변환을 사용할 때 다른 화면 내 예측 블록의 경우보다 더 효과적으로 적용된다.

한편, 화면 내 예측의 경우 AMT 는 화면 내 예측 모드에 기반하여 잔차신호의 특성을 통계로 얻은 것 이기 때문에, 채널의 종류와 상관없이 동일하게 효율적으로 작용할 가능성이 크다. 그러나 HEVC 와는 달리 JEM 에서 달라진 기본 조건들이 있다. QTBT 블록 분할 방법이 적용됨과 동시에[3], 기존 HEVC 의 I-slice 와는 다르게, JEM 의 휘도 성분 Y 와 색차 성분 Cb, Cr 의 블록은 독립적으로 구성된다. 따라서 색차 채널의 CU 에서 휘도 채널정보를 그대로 가져오는 것은 알맞지 않다. 그러므로 본 논문에서는 색차신호에 대한 별도의 적응형 다중변환 방법에 대해 몇 가지 조건을 추가하였다.

첫번째로, 휘도 신호와 색차신호의 블록 분할 방법이 다른 I-slice 에서는 기존의 방법과 같이 DCT2 가 색차신호에 적용된다. 휘도 신호와 색차신호의 블록 분할방법이 동일한 B, P-slice 에서는 휘도 신호에서 정해진 다중변환을 동일하게 색차신호에 적용할 수 있도록 하였는데, 화면 내 예측 블록에 있어서는, 색차신호의 화면 내 예측모드가 CCLM mode 6 개에 속할 때에만 휘도 성분의 다중변환을 동일하게 사용하도록 하였다.

표 1. 실험조건

실험시퀀스	(Cactus, ParkScene, Kimono)
QP	22,27,32,37
프레임 수	50 frame
부호화 조건	RA(random access)

표 2. 제안한 실험방법 결과

Sequences	EMT_Chroma BDBR (%) (proposed)		
	Y	U	V
Cactus	0.10	-2.00	-1.20
ParkScene	0.00	-0.80	-0.60
Kimono	0.00	-1.40	-1.10
Average	0.03	-1.40	-0.96

4. 실험결과분석

본 논문에서는 JEM7.2 에 있는 휘도 성분에만 적용되고 있는 AMT 를 화면 내/간 예측블록의 색차 신호 채널에도 적용하여 그 성능을 비교하였다. 특별히 화면 내 예측 블록에 있어서는, 색차신호의 화면 내 예측모드가 휘도성분과 색차성분의 연관성이 깊은 CCLM mode 6 개 중 하나의 모드가 선택되었을 때 휘도성분의 다중변환 (수평 및 수직변환 조합 집합)을 동일하게 사용하도록 함으로써 색차 성분 블록의 잔차 블록에서의 변환을 기존의 방법과 다르게 적용하였다. 실험결과 BDBR(%) 측면에서 색차성분인 U, V 에서 각각 1.40%, 0.96% 성능을 확인할 수 있었다. 후에 I-Slice 에서 휘도 성분 Y 와 색차 성분 Cb, Cr 의 블록을 공유한다면 색차 성분 블록에서 더 높은 압축 효율을 예상할 수 있다.

5. 결론

본 논문에서는 휘도 신호 예측에만 적용되던 AMT 를 화면 내/간 예측 부호화 되는 색차신호에도 적용할 경우 색차 신호의 부호화 효율 향상 여부를 분석하였다. 실험결과 휘도성분을 제외한 두개의 색차 성분에서 성능 향상을 확인할 수 있었다.

참고문헌

- [1] G. J. Sullivan, J. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits and Systems for Video Techno.*, vol. 22, no. 12, Dec. 2012.
- [2] T. Ukuba and O. Nakagami, "On Adaptive Multiple Core Transform for Chroma," *JVET-E0036*, Jan. 2017.
- [3] Jian le Chen, "Algorithm Description of Joint Exploration Test Model 7 (JEM 7)," *JVET-G1001*, Jul. 2017.
- [4] J. Chen, V. Seregin, S. Lee, W. Han, J. Kim, B.M. Jeon, "CE6.a: Chroma intra prediction by reconstructed luma samples," *JCTVC-E266*, Geneva, Mar. 2011.
- [5] Frank Bossen, "Common test conditions and software reference configurations," *JCTVC-L1100*, Jan. 2013.
- [6] 박지윤, 전병우, "색차 채널을 위한 적응형 다중변환의 효과분석," 제 30 회 영상처리 및 이해에 관한 워크샵(IPIU), pp.1-2, 2018 02.07-09, 제주..

VVC 를 위한 대체 움직임 정보 병합 후보 생성 기법

박도현, *이진호, *강정원, 김재곤
 한국항공대학교, *한국전자통신연구원
 dhpark@kau.kr, {jinosoul, jungwon}@etri.re.kr, and jgkim@kau.ac.kr

Generation of Alternative Merge Candidates for Versatile Video Coding (VVC)

Dohyeon Park, *Jinho Lee, *Jung Won Kang, and Jae-Gon Kim
 Korea Aerospace University, *ETRI

요 약

최근 JVET(Joint Video Experts Team)은 새로운 비디오 압축 표준인 VVC(Versatile Video Coding)의 표준화를 시작하였다. 기존의 HEVC 및 VVC 의 참조 SW 코덱인 HM 및 VTM(Versatile Test Model)에서는 효율적인 화면간 예측 부호화를 위한 움직임 정보 병합(Merge) 모드를 사용하고 있다. 본 논문에서는 VTM 의 Merge 후보 리스트 구성에서 공간적 주변블록의 움직임 정보가 존재하지 않을 경우, 이를 대체할 수 있는 Merge 후보 리스트 생성 기법을 제시한다. JVET CTC(Common Test Condition)를 이용하여 제안한 기법의 실험을 진행하였고, 실험결과 Y, U, V 성분 각각 0.2%, 0.17%, 0.12%의 BD-rate 감소를 확인하였다.

1. 서론

최근 MPEG(Moving Picture Experts Group)과 VCEG(Video Coding Experts Group)의 협력팀인 JVET(Joint Video Experts Team)에서는 HEVC(High Efficiency Video Coding) 다음의 차세대 비디오 부호화 표준을 VVC(Versatile Video Coding)로 명명하고 본격적인 표준화를 시작하였다. 제 10 차 JVET 회의에서 CfP(Call for Proposal) 응답 부호화 기술들을 비교 검증하기 위한 13 개의 CE(Core Experiment)를 설정하고, VVC 의 시험모델인 S/W 코덱 VTM(Versatile Test Model 1.0)을 배포하였다[1]. 다수의 JVET 의 CfP 응답 기술 제안서에서 화면간 예측 병합(Merge) 모드의 확장 및 수정을 통하여 부호화 성능을 향상할 수 있음을 확인하였다. 이에 따라 관련 기술들을 CE 에 추가하였고 앞으로 성능 비교 및 검증이 진행될 예정이다[2]. 본 논문에서는 VVC 의 화면간 예측 Merge 모드의 대체 움직임 후보 생성 기법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 VVC 의 화면간 예측 Merge 모드에 대해 살펴보고, 3 절에서는 제안하는 Merge 후보 리스트 구성 방법을 설명한다. 4 절에서는 실험을 통해 제안한 기법의 성능을 확인하였고 마지막 5 절에서는 본 논문에 대한 결론을 맺는다.

2. VVC 의 화면간 예측 Merge 모드

화면간 예측은 비디오의 프레임간 유사성을 이용하여 예측하는 방식이며 부호화/복호화가 진행되는 블록마다 움직임 정보를 유도 및 저장할 수 있다. Merge 모드는 현재 블록의 움직임 정보를 표현할 때 공간적/시간적 주변 블록의 움직임

정보를 그대로 사용한다. VVC VTM1.0 의 Merge 모드는 HEVC 와 마찬가지로 총 5 개의 후보 리스트를 구성하고 병합에 선택된 후보 리스트의 인덱스를 전송하여 현재 코딩 블록의 움직임 정보 전송을 생략한다. 후보 리스트는 현재 블록의 공간적 주변 블록, 시간적 주변 블록에 존재하는 움직임 정보, 혼합(Combined) 움직임 정보, 제로(zero) 움직임 정보로 구성된다. 그림 1 은 현재 블록의 시/공간적 주변 블록의 위치를 나타내고 L, A, AR, LB, AL 은 공간적 주변 블록의 위치, T0, T1 블록은 시간적 주변 블록의 위치이다.

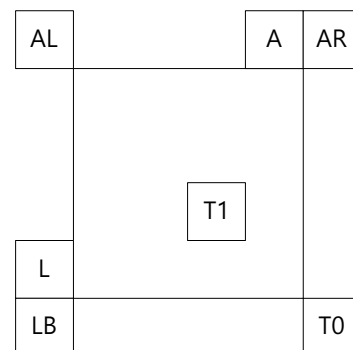


그림 1. 현재 블록의 시/공간적 주변 위치

VVC 의 Merge 후보 리스트는 다음과 같은 과정으로 구성된다. 좌측 블록(L)의 움직임 정보, 상단 블록(A)의 움직임 정보, 우-상단 블록(AR)의 움직임 정보, 좌-하단 블록(LB)의 움직임 정보, 좌-상단(AL) 위치의 움직임 정보, 시간적 주변 블록의 움직임 정보(T0, T1), 혼합 움직임 정보,

제로 움직임 정보 순으로 Merge 후보 리스트가 채워진다. 후보 리스트를 구성하는 과정에서 주변 블록의 움직임 정보를 여러 조건을 기준으로 사용할 수 없다고 판단될 때 해당 움직임 정보는 후보 리스트에서 제외된다.

3. 제안하는 대체적 Merge 후보 생성 및 추가

Merge 후보 리스트를 구성하기 위해 현재 블록의 시간적/공간적 주변 블록에 존재하는 움직임 정보를 이용한다. 현재 VVC 의 Merge 후보 유도 과정에서 주변 블록의 움직임 정보가 존재하지 않거나 이미 리스트에 구성된 다른 움직임 정보와 같다면 해당 움직임 정보는 Merge 후보 리스트에서 제외된다. 본 논문에서 제안하는 대체 Merge 후보 유도 방법은 인접 블록의 움직임 정보를 사용할 수 없을 때 해당 인접 블록 주변에 존재하는 움직임 정보를 이용하여 해당 인접 블록의 움직임 정보를 대체하는 움직임 정보를 생성하고 Merge 후보 리스트에 추가하는 것이다. 그림 2 와 같이 좌측 상단에 존재하는 인접 블록의 움직임 정보를 사용할 수 없을 경우 현재 블록 좌측의 움직임 정보와 상단의 움직임 정보를 생성하고 두 움직임 정보를 이용하여 좌측 상단의 움직임 정보를 대체할 수 있는 움직임 정보를 생성한다. L, LB 중 적어도 하나의 움직임 정보가 존재하고 A, AR 움직임 적어도 하나의 움직임 정보가 존재해야 제안하는 방법을 적용할 수 있다.

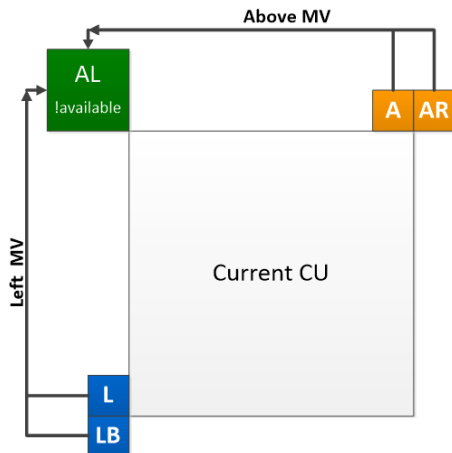


그림 2 대체 움직임 정보 생성 방법

생성된 두 움직임 정보의 가중치 평균으로 AL 위치의 움직임 정보를 대체한다. VVC 의 블록 구조 형태인 QTBT(Quad-tree plus Binary-tree)+TT(Ternary-tree)로 인해 코딩 블록은 정사각형 또는 다양한 직사각형 모양을 가질 수 있다. 제안한 방법에서 좌측과 상단의 움직임 정보를 가중치 평균할 때, 블록 모양에 따른 AL 위치에 대한 움직임 정보의 상관성을 반영하기 위해 (식 1)과 같이 좌측 움직임 정보에 현재 블록의 폭만큼의 가중치를 두고 상단 움직임 정보에는 현재 블록의 높이만큼의 가중치를 두고 평균한다.

$$MV_{AL} = \frac{MV_L \cdot W + MV_A \cdot H}{W \cdot H} \quad (\text{식 1})$$

4. 실험 결과 및 분석

제안하는 대체 Merge 후보 생성 및 유도 방법을 VTM1.0 위에 구현하였고 JVET CTC(Common Test Condition), Random-Access 환경에서 실험을 진행하였다[3]. 표 1 은 제안하는 기법에 대한 BD-rate 결과이다. 실험 결과, 평균 0.2%의 성능향상을 보였고 복호화 시간은 3% 증가하였다.

표 1. 실험결과(VTM1.0, Random access, JVET CTC)

	Y	U	V
Class A1	-0.37%	-0.54%	-0.30%
Class A2	-0.22%	-0.07%	-0.10%
Class B	-0.19%	-0.06%	-0.13%
Class C	-0.05%	-0.10%	0.01%
Overall	-0.20%	-0.17%	-0.12%
Class D	-0.16%	-0.08%	-0.21%
Class F	-0.13%	-0.26%	-0.17%
Dec. time(s)	103%		

5. 결론

본 논문에서는 차세대 비디오 부호화 표준으로 표준화가 진행 중인 VVC 의 화면간 예측 부호화를 위한 추가적인 Merge 후보 유도 방법을 제시하였다. 제안한 방법을 VTM1.0 에 구현하여 0.2% BD-rate 이득을 얻을 수 있음을 확인하였다. 제안 기법은 Merge 후보 리스트를 기존 방법보다 현재 블록의 움직임 정보를 잘 표현할 수 있게 구성하여 부호화 성능이 증가한 것으로 분석된다.

감사의 글

이 논문은 2018 년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(NO. 2016-0-00572, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2 배 압축을 제공하는 5 세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)

참고문헌

- [1] [Online]. Available at: https://jvet.hhi.fraunhofer.de/svn/svn_VVCSoftware_VTM/.
- [2] H. Yang, S. Liu, K. Zhang, "Description of Core Experiment 4 (CE4): Inter prediction and motion vector coding," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG Apr. 2018.
- [3] J. Boyce, K. Suehring, X. Li, V. Seregin, "JVET common test conditions and software reference configurations," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG Apr. 2018.

하드웨어 비디오 인코더를 위한 개선된 병렬 머지 방법

김대연, *임성창, *강정원

(주)칩스앤미디어, *한국전자통신연구원

brad.kim@chipsnmedia.com, sclim@etri.re.kr, jungwon@etri.re.kr

Improved Parallel Merge Scheme for Hardware Video Encoder

Dae-Yeon Kim *Sung-Chang Lim *Jungwon Kang

Chips&Media, Inc. *ETRI(Electronics and Telecommunications Research Institute)

요 약

H.265/HEVC(High Efficiency Video Coding)에서는 하드웨어 비디오 인코더의 처리율(Throughput)을 높이기 위하여 동일 CU(Coding Unit) 내 PU(Prediction Unit)들이 병렬로 머지 후보 리스트를 생성할 수 있는 병렬 머지 방법이 표준 기술로 사용되고 있다. 하지만 이 방법은 동일 CU 내의 PU 간의 의존성만 제거할 수 있고 코딩 순서상의 이전 CU와의 의존성은 제거할 수 없다. 결국 이전 CU의 모드 결정 과정이 완료된 후에 현재 CU 내의 PU가 머지 후보 리스트를 생성할 수 있기 때문에 높은 처리율 향상을 기대할 수 없다. 또한 CU 내의 대부분의 PU들이 가장 인접한 MV(Motion Vector)를 머지 후보로 사용하지 못하여 압축 효율에 대한 손실도 크다. 본 논문에서는 이전 CU와의 의존성을 제거함으로써 높은 처리율을 갖으면서 압축 효율에 대한 손실을 최소화할 수 있는 개선된 병렬 머지 방법을 제안한다. 실험 결과, 제안된 방법은 기존 병렬 머지 방법 대비 동일 화질에서 평균 약 1.8%의 압축률이 향상되는 것으로 나타났다.

1. 서론

H.265/HEVC(High Efficiency Video Coding)는 ITU-T VCEG (Video Coding Experts Group)과 ISO/IEC MPEG (Moving Picture Experts Group)이 공동으로 개발하여 2013년 4월 버전 1에 대한 표준화가 완료된 최신의 비디오 압축 표준이다. 이 표준은 확장된 블록 크기와 다양한 모양의 블록 크기, 확장된 변환 크기, 고도화된 화면 내 예측, 고정밀 보간 필터, SAO (Sample Adaptive Offset) 등의 기술로 인해 H.264/AVC(Advanced Video Coding) 대비 동일 화질에서 약 2 배의 압축률을 향상시켰다. 또한 연산 복잡도 최소화 및 하드웨어 구현 시의 처리율을 높이기 위한 위해 병렬 다블록 필터, 고처리율 CABAC(Context-Adaptive Binary Arithmetic Coding), WPP (Wave-front Parallel Processing), Tile, 그리고 병렬 머지 방법이 포함되었다[1-3]. 병렬 머지 방법은 특정 영역 내 여러 PU들이 동시에 머지 후보 리스트를 생성할 수 있도록 하는 방법이다. 하지만 이 방법은 압축 효율의 손실이 크며 각 CU 크기를 처리하는 모듈이 독립적으로 구현되고 동시에 처리되는 일반적인 실시간 하드웨어 인코더에서는 코딩 순서상 바로 이전 블록과의 의존성이 여전히 존재하기 때문에 단위 시간 당 처리율을 높이기 어렵다. 이러한 문제점을 해결하기 위해 본 논문에서는 코딩 순서를 고려하여 이전 블록과의 의존성을 제거한 개선된 병렬 머지 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 기존 H.265/HEVC의 병렬 머지 방법을 설명하고 그 문제점을 알아 보고, 3 절에서는 본 논문에서 제안하고 자하는 개선된 병렬

머지 방법을 설명한다. 4 절에서는 제안한 방법의 성능을 평가하기 위한 실험 조건과 실험 결과를 보이고, 5 절에서 본 논문에 대한 결론을 맺는다.

2. H.265/HEVC의 병렬 머지 방법

H.265/HEVC의 화면 간 예측 프레임에서는 동일 객체에 포함된 블록들은 서로 움직임 정보가 동일할 확률이 높다는 특성을 이용하여 공간/시간적으로 인접한 블록 중 어떠한 블록의 움직임 정보를 현재 블록의 움직임 정보로 사용할 것인지에 대한 인덱스만 전송하여 정보량을 줄일 수 있는 머지 기술이 사용되고 있다. 그림 1의 C는 현재 인코딩할 32x32 PU이며 이 때 공간적 머지 후보들은 A0, A1, 그리고 B0, B1, B2이다. 하지만 단일 시간내 처리율을 높이기 위해 인접한 PU들에 대하여 병렬로 ME (Motion Estimation)를 수행하거나 머지 모드 결정해야 하는 실시간 인코더에서는 머지 기술은 공간적으로 인접한 블록과의 의존성 때문에 처리율에 감소될 수 있는 문제가 있다.

이러한 문제를 해결하기 위해 인코더에서 특정 영역 내 여러 PU들이 동시에 머지 후보 리스트를 생성할 수 있는 방법을 제공하여 처리율을 높일 수 있는 병렬 머지 방법이 표준에 채택되었다. 특정 영역의 크기는 PPS (Picture Parameter Set)에 정의되며 MER(Merge Estimation Region)이라고 부른다. 병렬 머지 방법이 사용되는 경우 한 MER 내의 PU들은 MER 외부에 위치한 블록의 움직임 정보만을 머지 후보로 사용할 수 있다. 예를 들어 그림 1에서 64x64 MER이 사용되는 경우, 동일 MER 내부에 위치한 B0와 B1

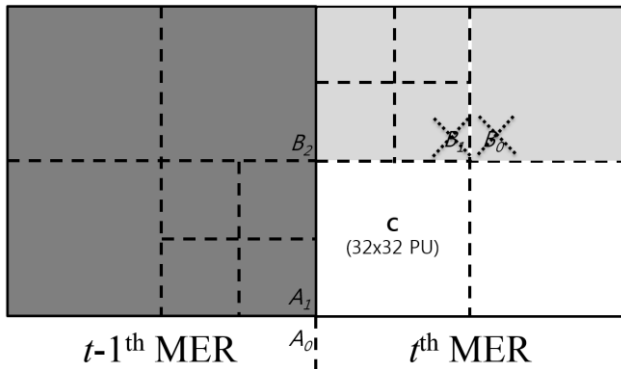


그림 1. H.265/HEVC 의 공간적 머지 후보 블록들과 병렬 머지 사용에 대한 예

은 머지 후보 리스트에서 제외되어 A0, A1, B2 혹은 시간적 머지 후보들만 최종 머지 모드로 선택될 수 있다.

하지만 이 방법을 사용하더라도 코딩 순서상 인접한 블록과의 의존성이 여전히 존재한다. 일반적인 실시간 하드웨어 인코더에서는 후보 CU 크기들의 인코딩을 담당하는 모듈들이 서로 독립적이면서 파이프라인 형태로 구현되기 때문에 결국 코딩 순서상 이전 블록의 최종 모드 결정이 완료될 때까지 현재 블록의 머지 후보 리스트를 생성할 수 없어서 처리율을 높이기 어렵다.

3. 개선된 병렬 머지 방법

인코더는 화질과 비트량을 모두 고려하여 다양한 CU 크기에 대하여 최적의 CU 크기를 결정해야 하며 한 CU 에 대해서도 다양한 예측 모드들(화면 내 예측, 화면 간 예측, 머지 및 스킵) 중 최적의 예측 모드도 결정해야 한다. 일반적인 실시간 하드웨어 인코더는 단위 시간 당 처리할 픽셀의 개수의 양을 고려하여 각 CU 크기들이 동시에 독립적으로 인코딩을 진행한다. 예를 들어 2Nx2N 크기의 CU 가 처리되는 동안 4 개의 NxN CU 및 16 개의 (N/2)x(N/2) CU 가 처리 될 수 있다고 가정하고 각 CU 에 대한 예측 모드 결정 과정이 동시에 진행된다.

하지만 H.265/HEVC 의 머지 기술은 주변 블록의 MV 를 후보로 사용하며 이를 선택한 경우 MV 의 변형없이 그대로 사용해야하기 때문에 주변 블록과 현재 블록간의 의존성이 존재한다. 즉, 주변 블록의 CU 크기와 예측 모드에 대한 결정이 완료되어야 현재 블록의 머지 후보 리스트를 생성할 수 있고 움직임 보상 과정을 통해 최적의 머지 MV 를 결정할 수 있다. 본 논문에서는 이러한 의존성을 제거하기 위하여 개선된 병렬 머지 방법을 제안하였다. 제안된 방법을 위해 PMU (Parallel Merge Unit)라는 블록 단위를 정의하였다. PMU 는 인코더에서 사용하고자하는 가장 큰 CU 크기이다. 만약 현재 인코딩할 CU 가 PMU 내의 첫번째 CU 라면 코딩 순서상 이전 PMU 는 아직 CU 크기와 예측 모드 및 MV 결정이 되지 않은 상태로 가정하여 이전 PMU 에 포함된 머지 후보는 머지 후보 리스트에 추가하지 않는다. 뿐만아니라 머지 후보가 동일 PMU 에 존재하더라도 코딩 순서 상 바로 이전 동일 크기 CU 영역 내 포함된다면 해당 영역 또한 CU 크기와 예측 모드 및 MV 가 결정이 되지 않은 상태로 가정하여 머지 후보 리스트에

추가하지 않는다.

그림 2 와 3 은 제안된 방법으로 인하여 머지 후보 리스트에 포함되지 않은 머지 후보의 예를 나타낸 그림이다. 이 중 그림 2 는 인코더가 64x64 CU 크기를 사용하여 머지 후보 리스트를 생성하려는 순간을 나타낸 그림이다. 이 때, 인코딩 순서상 이전 PMU(t-1 번째)는 64x64 CU 에 대한 예측 모드 결정은 완료되었으나 블록 분할 결정이 완료되지 않은 상태이다. 따라서 현재 블록 C 는 이전 PMU 에 포함된 머지 후보 A1 을 머지 후보 리스트에 추가하지 않는다. 그림 3 은 인코더가 PMU 내 3 번째 32x32 영역에 대해 32x32 CU 크기를 사용하여 머지 후보 리스트를 생성하려는 순간을 나타낸 그림이다. 이 때 코딩 순서상 바로 이전 32x32 영역에 대해 블록 분할 결정이 완료되지 않은 상태이기 때문에 해당 영역에 포함된 B0 는 머지 후보 리스트에 추가하지 않는다.

따라서 제안된 병렬 머지 방법을 사용하는 경우 블록 분할 및 예측 모드가 완료된 영역에 존재하는 머지 후보만 사용할 수 있기 때문에 이전 영역의 인코딩 완료를 기다릴 필요 없이 머지 후보를 이용하여 움직임 보상과 변환 및 양자화등의 일련의 과정을 수행하여 처리율을 높일 수 있다. 또한 기존 H.265/HEVC 의 병렬 머지 방법 보다 머지 후보 리스트에서 제외되는 머지 후보의 수가 적기 때문에 압축 효율의 손실을 최소화할 수 있다.

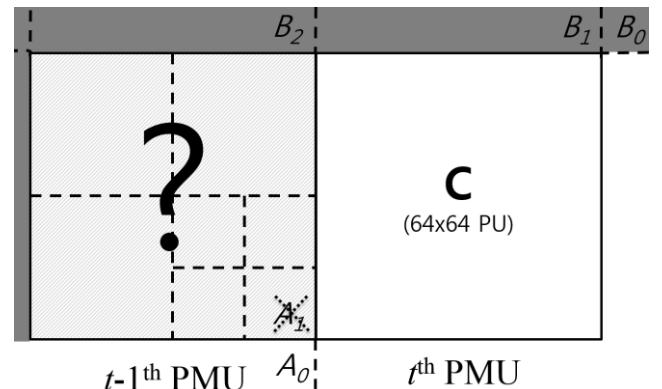


그림 2. PMU 크기가 64x64 이며 64x64 PU(C)를 인코딩하는 경우에 제안된 병렬 머지 방법으로 인해 제외되는 머지 후보

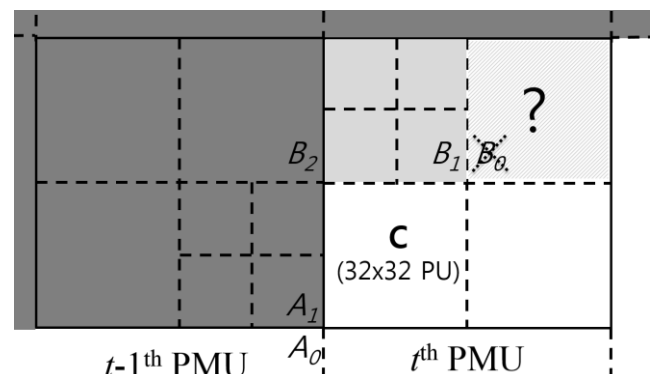


그림 3. PMU 크기가 64x64 이며 세번째 32x32 PU(C)를 인코딩하는 경우에 제안된 병렬 머지 방법으로 인해 제외되는 머지 후보

4. 실험 조건 및 결과

제안된 방법의 평가를 위하여 JCT-VC 의 공통 실험 조건 [4]에서 권장하는 실험 조건과 환경을 사용하였고 이는 표 1 과 같다. 또한 압축 성능 평가 방법은 2 개의 코덱에 대하여 동일 화질 대비 압축 성능의 차이를 나타내는 척도인 BD-rate [5]를 사용하였다.

먼저 기존 병렬 머지 방법의 성능을 평가하기 위하여 병렬 머지를 사용하지 않은 결과를 anchor 로 정하고 병렬 머지를 사용한 경우의 압축 효율 손실율을 알아보았다. 이 때, MER 의 크기에 따라 압축 효율이 달라질 수 있기 때문에 MER 크기 64 와 32 에 대하여 각각 실험하였다. 표 2 는 이에 대한 실험 결과이며 MER 크기가 64x64 인 경우, BD-rate(Y)에 대하여 평균적으로 약 7.6%의 압축률이 하락되며 MER 크기가 32x32 인 경우는 약 6.5%의 압축률이 하락되었다.

제안된 병렬 머지 방법의 성능을 평가하기 위하여 기존 병렬 머지의 MER 의 개념과 유사한 PMU 크기를 64 와 32 에 대하여 각각 실험을 하였다. 표 3 은 이에 대한 실험 결과이며, PMU 64x64 인 경우, BD-rate(Y) 기준으로 약 5.8%의 압축률이 하락되었고 PMU 32x32 인 경우에는 약 5.5%의 압축률이 하락되었다.

따라서 기존 H.265/HEVC 의 병렬 머지 방법 대비 제안된 병렬 머지 방법은 병렬 처리 단위 64x64 인 경우, 평균적으로 약 1.8% 압축률이 향상되며 병렬 처리 단위 32x32 인 경우에는 약 1.0% 압축률이 향상되는 것을 알 수 있다.

표 1. 실험 환경 및 조건

Codebase	HM16.16
Profile	H.265/HEVC Main
Coding structure	Random access
GOP size	16
QP	22, 27, 32, 37
Number of merge candidates	5

표 2. H.265/HEVC 의 병렬 머지 방법에 대한 실험 결과

	MER64			MER32		
	Y	U	V	Y	U	V
Class B	8.6%	6.1%	6.8%	7.1%	4.8%	5.5%
Class C	7.0%	5.9%	6.5%	6.0%	4.8%	5.3%
Class D	7.2%	6.6%	7.1%	6.5%	5.7%	6.1%
Overall	7.6%	6.2%	6.8%	6.5%	5.1%	5.6%

표 3. 제안된 병렬 머지 방법에 대한 실험 결과

	PMU64			PMU32		
	Y	U	V	Y	U	V
Class B	6.8%	4.7%	5.3%	6.3%	4.1%	4.8%
Class C	4.9%	3.3%	3.9%	4.8%	3.2%	3.7%
Class D	5.7%	4.7%	5.0%	5.5%	4.5%	4.7%
Overall	5.8%	4.2%	4.8%	5.5%	4.0%	4.4%

5. 결론

본 논문에서는 H.265/HEVC 의 병렬 머지 방법의 문제점을 제기하였고 이를 해결하기 위한 개선된 병렬 머지 방법을 제안하였다. 개선된 병렬 머지 방법은 머지 후보 리스트 생성 시 코딩 순서 상 바로 이전 블록을 머지 후보에서 제외하기 때문에 이전 블록의 인코딩 완료를 기다릴 필요가 없어 처리율을 높일 수 있으며 실험 결과 기존 병렬 머지 방법보다 압축률의 손실이 적은 것으로 나타났다.

H.265/HEVC 의 다음 표준인 VVC(Versatile Video Codec) 표준은 연산 복잡도가 높은 기술들이 채택될 것으로 판단되고 기존 표준보다 많은 수의 머지 후보들이 고려되어 최적의 머지 모드를 선택하는 것이 중요한 이슈가 될 것이다. 제안된 병렬 머지 방법은 VVC 표준에도 응용 가능하며 처리율을 높이면서 압축률 손실을 최소화할 수 있을 것으로 예상된다.

참고문헌

- [1] G. J. Sullivan, J. Ohm, W.-J. Han, 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.
- [2] Yuan, Y.; Kim, I.-K.; Zheng, X.; Liu, L.; Cao, X.; Lee, S.; Cheon, M.-S.; Lee, T.; He, Y.; Park, J.-H.; , "Quadtree Based Nonsquare Block Structure for Inter Frame Coding in High Efficiency Video Coding," Circuits and Systems for Video Technology, IEEE Transactions on , vol.22, no.12, pp.1707–1719, Dec. 2012
- [3] Minhua Zhou, "Parallelized merge/skip mode for HEVC", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Document JCTVC-F069, Torino, July, 2011
- [4] Bossen F, "Common test conditions and software reference configurations", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Document JCTVC-L1100, Geneva, Jan. 2013
- [5] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T VCEG Meeting, Document VCEG-M33, Austin, Texas, USA, Tech. Rep. 2001.

감사의 글

이 논문은 2018 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임. (No. 2016-0-00572, 초고실감 미디어 서비스 실현을 위해 HEVC/3DA 대비 2 배 압축을 제공하는 5 세대 비디오/오디오 표준 핵심 기술 개발 및 표준화)