

차세대 비디오 코덱(JVET) 부호화 알고리즘

Coding Tools for Improving Coding Efficiency of Future Video Coding

김신재, 도지훈

지도교수: 김재곤

요 약

차세대 비디오 코덱(JEM: Joint Exploration Model)은 HEVC(High Efficiency Video Coding)대비 약 2배의 부호화 효율을 가지는 것을 목표로 진행하고 있다. 본 논문에서는 현재까지 진행된 JEM7.0 대비 더 높은 부호화 효율을 얻을 수 있는 Reference sample padding(RSP)과 Merge mode 기법을 제안한다. 제안기법은 RSP의 경우 존재하는 참조샘플의 선형성을 판단하여 비어있는 샘플 값을 선형적으로 채운다. Merge mode의 경우 Correlation이 높은 Block을 우선적으로 탐색한다. 실험결과 Padding기법의 경우 JEM 7.0 대비 All Intra 부호화에서 0.02% BD-rate 증가로 큰 변화 폭이 없었다. Merge mode의 경우 Random Access 부호화에서 0.12%의 BD-rate 감소 효과를 얻을 수 있었다.

Keywords : Future Video Coding, JEM software, Reference sample padding, Merge mode

I. 서 론

최근 H.264 대비 2배의 부호화 효율을 갖는 HEVC 표준화가 완료되었다. 하지만 5G의 새로운 모바일 인프라의 등장과 차세대 미디어로서 360도 VR(Virtual Reality)이 주목 받는 등 고화질의 실감 미디어 서비스가 빠르게 확산될 것으로 예상됨에 따라 보다 더 높은 성능의 비디오 압축의 필요성은 더욱 높아지고 있다. 이에 ITU-T VCEG과 ISO/IEC이 JVET(Joint Video Exploration Team)을 구성하여 2020년을 목표로 HEVC보다 2배의 효율을 갖는 차세대 비디오 코덱 표준화를 진행 중이다.

JVET은 차세대 비디오 부호화 기술을 탐색하고 있는 단계이며 이를 위한 참조 SW 코덱으로 JEM(Joint Exploration Model) 7.0을 배포하였다. JEM 7.0은 HEVC의 부호화 틀을 기반으로 이를 확장한 기술과 몇 가지의 새로운 부호화 틀들을 채택하여 현재 약 29%의 부호화 성능 향상을 보이는 반면에 10배에 달하는 복잡도 증가를 보인다[1]. 2017년 10월 MPEG회의에서 본격적으로 표준화 진행을 하겠다고 공표한 상태이다[2].

기존의 HEVC 및 JEM에서는 화면 내 예측 부호화 과정에서 참조샘플(Reference sample)을 이용하여 화면 내 예측블록을 구성하고 있으며, 참조샘플이 가용하지 않을 경우 가장 가까운 샘플로 참조샘플 화소위치를 채우는 패딩(Padding)을 한다[3]. 화면 내 예측 부호화 성

능을 높이기 위해서는 현재블록을 보다 잘 예측할 수 참조샘플 패딩이 필요하다.

기존 HEVC 대비 JEM의 추가적인 기술 중 대표적으로 기존 HEVC의 CU, PU, TU 계층적인 부호화 블록 구조 개념을 없애고 그림 1과 같이 QTBT로 더 유연하고 다양한 CU 구조를 갖는 블록 구조를 채택하였다[1]. 이를 통해 HEVC에선 현재 CU 블록 내의 여러 PU블록에서 Merge mode를 진행하였다면, JEM에선 현재 CU블록에서 Merge mode를 사용해 움직임 벡터를 추정한다.

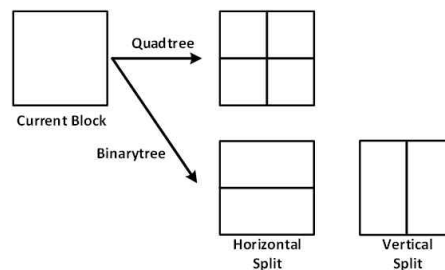


그림 1. JEM의 QTBT 블록 분할 구조

본 논문에서는 JEM에서의 보다 효율적인 예측을 위한 참조샘플 패딩 기법과 Merge mode 기법을 제시한다.

II. Reference sample padding (RSP)

1. RSP 기법

JEM에서의 이용 불가능한 참조샘플에 대한 패딩 방법은 HEVC와 마찬가지로 가장 가까운 샘플로 비어있는 샘플을 채운다[4]. 참조샘플의 일부만 이용 가능한 경우는 그림 2와 같이 크게 3가지로 나눌 수 있다. 첫 번째는 위쪽과 왼쪽이 일부만 이용 가능한 경우로, CTU단위로 부호화가 진행되면서 이미 부호화된 블록이 일부 존재하는 경우이다. 두 번째는 위쪽의 샘플들만 이용 가능한 경우이고 세 번째는 왼쪽의 샘플들만 이용 가능한 경우이다. 위의 두 경우는 CTU나 CU블록의 경계에 있는 블록들을 말한다.

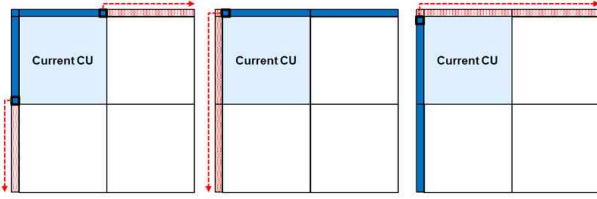


그림 2 참조샘플 패딩 방법

2. 제안하는 기법

식 1은 참조샘플들의 선형성을 판단하는 식이다. 참조샘플이 존재하는 영역 혹은 이미 부호화된 블록에서 세 개의 샘플을 이용하여 선형성을 판단하고 샘플들이 선형적이다 판단되면 계산된 변화량을 패딩하는 영역의 위치에 따라 그림 3과 같이 채운다. 식 1의 x 값은 선형성을 판단하는 계수로서 x 의 값이 커질수록 선형성을 판단하는 정도가 높아지고 최대 10까지 적용할 수 있다. 채워지는 샘플의 화소 값은 식 2에 따라 계산된다. 두 번째와 세 번째의 경우 경계의 특성상 옆 또는 위 블록에 변화량을 계산할 때 쓰이는 샘플이 존재하지 않는다. 이에 대해서는 이미 부호화된 샘플을 참조샘플로 삼는다.

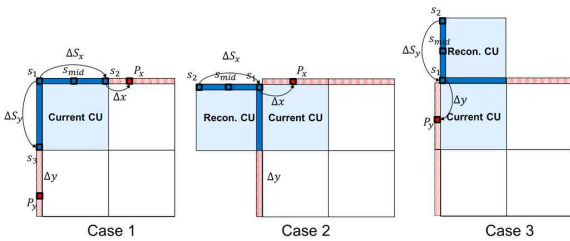


그림 3 제안하는 참조샘플 패딩 방법

$$Abs(s_1 + s_2 - 2s_{mid}) < (1 \ll (BitDepth_Y) - x) \quad (1)$$

$$\begin{aligned} s_1 - s_2 &= \Delta S_x & s_1 - s_3 &= \Delta S_y \\ P_x &= s_2 - \frac{\Delta x \Delta S_x}{width_{CU}} & P_y &= s_3 - \frac{\Delta y \Delta S_y}{height_{CU}} \end{aligned} \quad (2)$$

3. 실험 결과

제안한 방법의 성능 분석을 위해 JEM 7.0에 구현을 하여 비교분석을 하였다. 모든 픽처에 대하여 화면 내 예측을 수행하는 All Intra Configuration으로 부호화를 진행하였다. 실험에 사용된 영상은 JVET CTC (Common Test Condition)에서 사용되는 시퀀스 중 B Class에 해당하는 4개의 시퀀스를 사용하였다. 각 영상마다 22, 27, 32, 37 총 4개의 QP(Quantization Parameter)로 실험하였다[5]. 식 1에서 선형성을 판단하는 계수 x 는 모든 시퀀스에 8로 적용하였다.

표 1은 첫 번째 경우에만 제안방법을 적용했을 때의 실험결과이다. 시퀀스에 따라 Y성분에서의 BD-rate가 감소하거나 변화가 거의 없는 것을 확인할 수 있었다.

표 2는 모든 경우에 제안방법을 적용했을 때의 실험결과이다. 모든 시퀀스의 Y성분의 BD-rate가 조금 증가하거나 변화가 거의 없는 것을 확인할 수 있었다.

표 1 첫 번째 경우에만 제안방법 적용 시

HD sequence	BD-rate		
	Y	U	V
ParkScene	0.02%	-0.10%	-0.13%
Cactus	0.03%	-0.32%	0.16%
BasketballDrive	-0.09%	-0.14%	-0.32%
BQTerrace	-0.03%	0.23%	0.27%

표 2 모든 경우에 제안방법 적용 시

HD sequence	BD-rate		
	Y	U	V
ParkScene	0.01%	-0.08%	-0.04%
Cactus	0.02%	0.07%	0.13%
BasketballDrive	0.02%	-0.26%	-0.10%
BQTerrace	0.05%	0.17%	-0.14%

Bitstream Analyzer를 이용하여 제안방법의 결과에 대한 분석을 하였다. 그림 4는 성능증가가 예상되는 부분의 예시를 보여준다. 그림 4를 보면 원본과의 차분 값 범위가 기존방법보다 제안방법일 때 더 작다는 것을 확인할 수 있다. 차분 값의 범위가 작다는 것은 양자화에러를 감소시킬 수가 있고 이는 화질의 증가로 이어진다. 하지만 반대의 경우도 존재하고 이러한 경우에는

오히려 화질의 감소로 이어진다.

그림 5는 성능감소가 예상되는 부분의 예시를 보여 준다. 그림 5를 보면 제안방법 적용 시 화질 증가를 위해 블록이 더 많이 분리 되는 것을 알 수 있다. 이러한 경우에는 bit의 증가를 야기할 수 가있다.

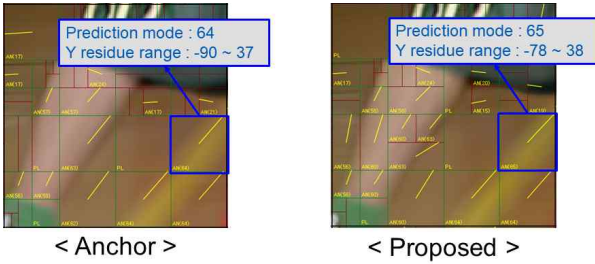


그림 4 성능 증가의 예시

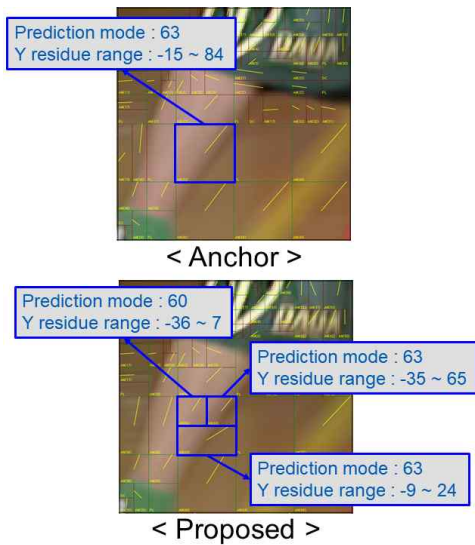


그림 5 성능 감소의 예시

III. Merge mode

1. Spatial Merge 후보 참조 기법

JEM에서의 Merge 후보를 받아오는 순서는 HEVC와 유사하게 Spatial Candidates, Temporal Candidates, Combined bi-predictive Candidates, Zero vector Candidates의 순서를 사용하여 움직임 정보를 추정하고, 인접한 블록과 Correlation이 높은 블록에서 큰 부호화 효율을 얻는다. 위의 4개의 순서 중 Spatial Candidates의 경우 HEVC와 유사하게 주변 블록에선 Left, Above, Above Right, LeftBottom, AboveLeft의 순서대로 다섯 개의 블록을 순차적으로 탐색하여 최대 4개의

Candidates를 추가할 수 있다.

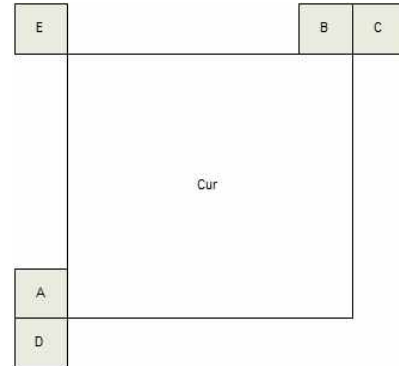


그림 6 Spatial Candidates Mode

2. 제안하는 기법

본 논문에서 제안하는 기법은 현재 CU에서 주변 블록의 상관도를 고려하여 Spatial Merge 후보를 탐색하는 순서를 결정한다. 제안 기법은 Merge 모드의 경우, 인접한 블록과 상관도가 높은 블록에서 더 큰 부호화 효율을 얻을 수 있는 Merge 모드의 특성을 이용하여 현재 CU의 Block Size에 따라 주변 블록과의 상관성이 다르게 될 것이라고 가정하였고 이에 따라 세 가지 방법을 제안한다.

그림 7은 첫 번째 방법을 보여준다. Block Size의 경우에 따라 Correlation이 높다고 생각되는 블록을 먼저 받아올 수 있도록 변경한 예이다. 기존 방법과 다르게 너비와 높이를 고려하여, 높이가 너비보다 Size가 큰 경우에만, Quad Split에서 최종적으로 Vertical로 Split이 되었을 것이라 가정하였고, 이와 같은 경우, 왼쪽 Block과 오른쪽 Block의 Correlation이 낮다는 것을 의미한다. 이와 같은 경우 Merge후보를 받아오는 순서를 변경하도록 했다.(Left, Above만 순서 변경)

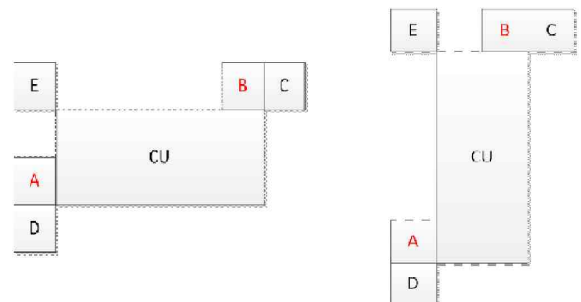


그림 7 제안하는 Spatial Merge 모드 - 방법 1

두 번째 방법은 Block Size에 추가적으로 Binary Split이 이루어질 때, Quad Split이 이루어진 것과 같이 정사각형의 CU Block Size가 발생하는 경우에 대해서 추가적으로 변경한 예이다. 첫 번째 방법에 추가적으로 너비와

높이가 같을 경우에도 최종 Split이 Vertical인 경우, Merge 후보를 받아오는 순서를 변경하였다.(모든 순서 변경)

세 번째 방법은 그림 8과 같이 단순한 Block Size로 결정하는 것이 아니라, 마지막 Split이 무엇인지에 따라 Merge 후보를 받아오도록 결정하였다. 그림 7의 경우, Size만을 고려했기 때문에, Vertical하게 최종적으로 Split이 되었을 경우에도 높이보다 너비가 더 큰 경우가 발생한다. 이를 보완하고자 최종 Split만을 고려하여 Merge후보를 받아오도록 하였다.

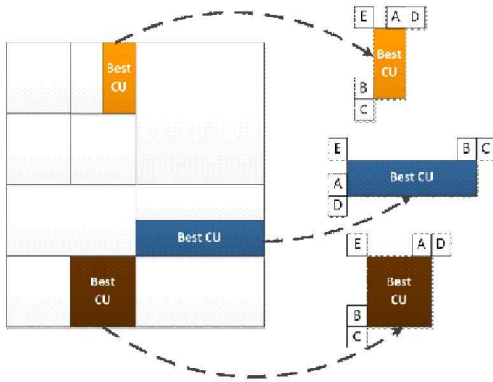


그림 8 제안하는 Spatial Merge 모드- 방법 3

3. 결과

본 논문은 제안한 방법의 성능 분석을 JEM 7.0에 구현을 하여 비교분석을 하였다. 실험에 사용된 영상은 JVET CTC(Common Test Condition)에서 사용되는 B Class에 해당하는 4개의 시퀀스를 사용하였고, 각 영상마다 22, 27, 32, 37의 4개의 QP(Quantization Parameter)로 실험하였다[5].

표 3은 첫 번째 방법을 적용했을 때의 실험결과이다. 이 경우 A Class도 추가적으로 진행하였고, B Class에선 성능의 향상이 평균 0.12%정도 발생하지만, A Class에선 평균 0.02%의 성능저하가 발생한다. 또한 해상도가 커질수록 성능저하가 발생하는 것으로 보인다.

표 3 방법 1적용 시

	BD-rate (piecewise cubic)			BD-rate (cubic)		
	Y	U	V	Y	U	V
FoodMarket	0.06%	0.36%	0.35%	0.08%	0.37%	0.36%
Tango	0.01%	-1.23%	0.03%	0.01%	-1.25%	0.01%
CatRobot	0.07%	-0.03%	0.37%	0.05%	-0.05%	0.32%
DaylightRoad	-0.02%	-0.25%	0.08%	-0.01%	-0.25%	0.09%
BuildingHall	-0.01%	-0.17%	-0.13%	-0.01%	-0.16%	-0.13%
UHD Total	0.02%	-0.26%	0.14%	0.02%	-0.27%	0.13%
BQTerrace	-0.06%	-0.33%	0.38%	-0.07%	-0.32%	0.37%
RitualDance	-0.15%	-0.04%	-0.21%	-0.15%	-0.07%	-0.35%
Timelapse	-0.12%	0.19%	-0.45%	-0.12%	0.12%	-0.55%
BasketballDrive	-0.16%	0.47%	0.57%	-0.18%	0.45%	0.56%
Cactus	-0.10%	0.10%	0.01%	-0.10%	0.10%	0.02%
HD Total	-0.12%	0.08%	0.06%	-0.12%	0.05%	0.01%
Total	-0.05%	-0.09%	0.10%	-0.05%	-0.11%	0.07%

표 4는 두 번째 방법을 적용했을 때의 실험결과이다. B Class에선 평균 0.05%의 성능이 향상된다. 첫 번째 방법에 비해 0.07%정도의 성능이 저하되는 것을 확인할 수 있고, 너비와 높이가 같을 때 Merge후보의 변경을 하지 않는 것이 더 성능이 좋고 Left와 Above만의 순서 변경이 더 적합한 것을 확인할 수 있다.

표 4 방법 2 적용 시

	BD-rate (piecewise cubic)			BD-rate (cubic)		
	Y	U	V	Y	U	V
BQTerrace	-0.03%	-0.37%	-0.13%	-0.05%	-0.38%	-0.13%
RitualDance	-0.03%	-0.18%	0.00%	-0.03%	-0.23%	0.10%
Timelapse	-0.14%	-0.13%	-0.43%	-0.15%	-0.14%	-0.42%
BasketballDrive	-0.06%	0.49%	-0.07%	-0.05%	0.50%	-0.05%
Cactus	0.03%	0.03%	-0.26%	0.04%	0.03%	-0.26%
HD Total	-0.05%	-0.03%	-0.18%	-0.05%	-0.04%	-0.15%

표 5은 세 번째 방법을 적용했을 때의 실험결과이다. 이 경우 B Class에선 성능의 향상이 평균 0.05%정도로 방법[2]와 성능이 유사하다. 이를 토대로 Block Size를 고려했을 경우가 최종 Split을 고려했을 때와 유사하다는 것을 확인할 수 있다.

표 5 방법 3 적용 시

	BD-rate (piecewise cubic)			BD-rate (cubic)		
	Y	U	V	Y	U	V
BQTerrace	-0.11%	-0.27%	-0.06%	-0.11%	-0.28%	-0.06%
RitualDance	0.00%	-0.10%	-0.07%	0.00%	-0.04%	-0.03%
Timelapse	-0.05%	0.41%	0.05%	-0.05%	0.39%	0.08%
BasketballDrive	-0.04%	0.38%	-0.20%	-0.02%	0.34%	-0.20%
Cactus	-0.06%	-0.09%	-0.24%	-0.06%	-0.09%	-0.23%
HD Total	-0.05%	0.06%	-0.11%	-0.05%	0.06%	-0.09%

그림 9는 Bitstream Analyzer를 이용하여 확인한 것으로 파란색의 블록과 같은 경우, Merge 모드의 탐색순서 변경이 이루어지는 것을 확인할 수 있다. 움직임 벡터의 변화량을 확인하여 제안방법의 결과에 대한 분석을 하였다. 원본에 비해 평균적으로 대략 1~2%의 변화가 생겨났고, Correlation이 높은 Above 블록을 우선적으로 탐색한다.

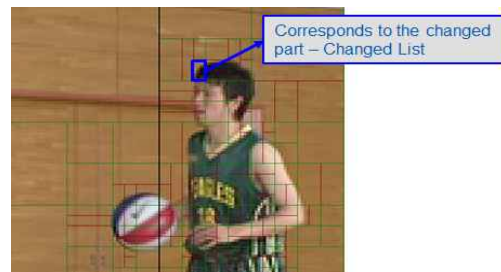


그림 9 제안한 방법이 적용되는 예

IV. 결 론

본 논문은 JEM 7.0대비 더 높은 부호화 효율을 얻기 위해서 향상된 Reference sample padding(RSP)와 Merge mode 기법을 제시하였다. RSP의 경우, 실험결과 기존 JEM7.0 대비 All Intra 부호화 모드에서 0.02%의 BD-rate 증가로 변화폭에 큰 차이가 없었다. 실험결과를 통하여 비어있는 참조샘플을 선형성을 토대로 선형적으로 채워 넣는 것이 기존의 방법을 사용하여 일정한 값으로 모두 채우는 것과 비교하여 부호화 효율을 높이는 것에 크게 관여하지 않음을 확인하였다. Merge mode의 경우, 실험결과 기존 JEM7.0 대비 Random Access 부호화 모드에서 0.12%의 BD-rate 감소를 보였다. 실험 결과를 통해 기존 방법에 비해 Correlation을 고려한 Merge List 탐색 순서의 변경이 더 효과적인 것을 확인 할 수 있었다. 또한 이 방법의 경우 해상도가 낮을수록 더 좋은 효율을 보이는 것으로 확인되었다. 현재 본 논문의 제안 기법을 동시에 적용하고 추가적인 방안을 적용하여 더 높은 효율을 가질 수 있는 기법을 연구할 예정이다.

참고문헌

- [1] J. Chen, E. Alshina, G. J. Sullivan, J. -R. Ohm, and J. Boyce, "Algorithm Description of Joint Exploration Test Model 7," JVET document, JVET-G1001-v1, Jul. 2017.
- [2] A. Segall, V. Baroncini, J. Boyce, J. Chen, T. Suzuki, "Joint Call for Proposals on Video Compression with Capability beyond HEVC," JVET document, JVET-H1002-v6, Oct. 2017.
- [3] G. J. Sullivan, J. Ohm, W. J. Han, T. Wiegand "Overview of the high efficiency video coding (HEVC) standard." IEEE Transactions on circuits and systems for video technology, vol. 22, no. 12, pp. 1649 - 1668, Dec. 2012.
- [4] J. Lainema, F. Bossen, W. J. Han, J. H. Min, K. Ugur, "Intra Coding of the HEVC Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1792 - 1801, Dec. 2012.
- [5] K. Suehring, X. Li, "JVET common test conditions and software reference configurations," JVET document, JVET-B1010, Feb, 2016.

저자약력

김 신 재 (Kim-Sin Jae)

2012년 3월 : 한국항공대학교
항공전자 및 정보통신공학부 입학
2018년 2월 : 한국항공대학교
항공전자정보공학부 졸업 예정
<관심분야> 전자공학, 통신공학,
멀티미디어공학

도 지 훈 (Do-Ji Hoon)

2013년 3월 : 한국항공대학교
항공전자 및 정보통신공학부 입학
2018년 2월 : 한국항공대학교
정보통신학과 졸업 예정
<관심분야> 전자공학, 통신공학,
멀티미디어공학

김 재 곤 (Jae-Gon Kim)



1990년 2월 : 경북대학교
전자 공학과 졸업
1992년 2월 : KAIST
전기 및 전자공학과 석사
2005년 2월 : KAIST
전자전산학과 박사

1992년 3월 ~ 2007년 3월: 한국전자통신연구원
(ETRI)선임연구원/팀장

2001년 8월 ~ 2002년 11월: Columbia 대학, 연구원

2007년 9월 ~ 현재: 한국항공대학교

항공전자정보공학부 교수

<관심분야> 비디오 신호처리, 비디오 부호화, 영상통신,
디지털 방송 미디어