

## 몰입형 비디오 압축을 위한 화면 내 블록 카피 성능 분석

이순빈<sup>1)</sup>, 정종범<sup>1)</sup>, 류일웅<sup>2)</sup>, 김성빈<sup>1)</sup>, 김인에<sup>1)</sup>, 류은석<sup>1)</sup>

1) 성균관대학교 컴퓨터교육과, 2) 가천대학교 컴퓨터공학과  
soonbinlee@skku.edu, uof4949@skku.edu, dlfdnd96@gc.gachon.ac.kr,

beencoocke@skku.edu, inaelk@skku.edu, esryu@skku.edu

### Intra Block Copy Analysis to Improve Coding Efficiency for Immersive Video

Soonbin Lee, Jong-Beom Jeong, Il-Woong Ryu, Sungbin Kim, Inae Kim  
and Eun-Seok Ryu

1) Department of Computer Education, Sungkyunkwan University,

2) Department of Computer Engineering, Gachon University

#### 요 약

최근 MPEG-I 그룹에서는 표준화가 진행중인 몰입형 미디어(Immersive Media)에 대한 압축 성능 탐색이 이루어지고 있다. 몰입형 비디오는 다수의 시점 영상과 깊이 맵을 통한 깊이 맵 기반 이미지 렌더링(DIBR)을 바탕으로 제한적 6DoF를 제공하고자 하는 기술이다. 현재 MIV(Model for Immersive Video) 기술에서는 바탕 시점(Basic View)과 각 시점의 고유한 영상 정보를 패치 단위로 모아둔 추가 시점(Additional View)으로 처리하는 모델을 채택하고 있다. 그 중에서 추가 시점은 일반적인 영상과는 달리 시간적/공간적 상관성이 떨어지는 분절적인 형태로 이루어져 있어 비디오 인코더에 대해 최적화가 되어 있지 않으며, 처리 방법의 특성에 따라 자기 유사적인 형태를 지니게 된다. 따라서 MIV 에서 스크린 콘텐츠 코딩 성능과 함께 화면 내 블록 카피(IBC: intra block copy) 기술에 대한 성능을 분석 결과를 제시한다. IBC 미적용 대비 최대 7.56%의 Y-PSNR BD-rate 감소가 가능함을 확인하였으며, 영상의 특성에 따라 IBC의 선택 비율을 확인하여 추가 시점의 효율적인 압축 형태를 고찰한다.

#### 1. 서론

현재 표준화 단체 MPEG(moving picture experts group)에서는 여러 위치에 다양한 시점으로 존재하는 영상들을 취득하고 처리하여 사용자에게 화면 전환 이동에 따른 운동 시차(motion parallax)를 제공할 수 있는 몰입형 미디어 기술에 관한 표준화가 진행되고 있다. 이러한 표준화의 일환으로 2019년 3월 MPEG 126차 회의에서 MPEG-I(Immersive) 서브그룹은 6DoF(degrees of freedom)을 제공하는 몰입형 미디어를 지원하는 테스트 모델(Test Model for Immersive Video, TMIV)를 정의하였다[1].

6DoF 몰입형 미디어 기술 표준화는 몰입형 미디어를 위한 다수의 영상 처리를 목적으로 MIV(Model for Immersive Video) 기술에 대한 성능 탐색을 현재도 지속적으로 수행하고 있다. 6DoF 기술은 자유로운 사용자의 시선과 움직임에 따른 시점을 지원하므로, 이를 위해 여러 위치에서 취득된 영상들을 동시에 압축 및 전송하여 사용자 시점의 영상을 취득하고 합성하는 과정에서 높은 대역폭과 연산량이 요구된다.

현재 MIV에서는 깊이 맵 기반 이미지 렌더링(DIBR) 기법을 바탕으로 하여 다수의 영상들을 깊이 맵과 함께 취득한 후, 깊이 맵을 바탕으로 각 영상의 시점과 거리를 계산하여 가상의 영상 시점을 합성하는 방식으로 진행된다. 따라서 사용자의 시선과 위치에 맞도록 시점을 합성해낼 수 있으며, 이로 인해 사용자는 보다 상호작용이 가능하고 몰입감 있는 미디어를 감상할 수 있게 된다.

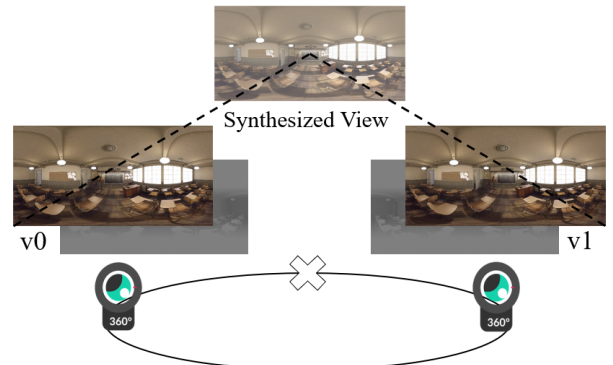


그림 1. 깊이 맵 기반 이미지 렌더링(DIBR)을 통한 중간 시점 생성의 예시

그림 1은 깊이 맵 기반 이미지 렌더링 기술을 이용하여 중간 시점의 영상을 생성해내는 예시를 보여준다. 이 때 몰입형 미디어는 텍스처(texture) 영상과 깊이 맵(depth map)을 포함하는 다수의 영상으로 이루어진다.

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구로 MIV에 대한 기술 소개와 처리 과정에 대해 설명한다. 또한 해당 기술의 특성으로 인해 나타나는 문제점을 서술한다. 3절에서는 앞서 서술한 문제점에 대해 스크린 콘텐츠 코딩 기술에 대한 성능 분석을 진행한다. 마지막으로 4절에서는 본 논문에 대한 결론을 맺는다.

## 2. MIV (Model for Immersive Video)

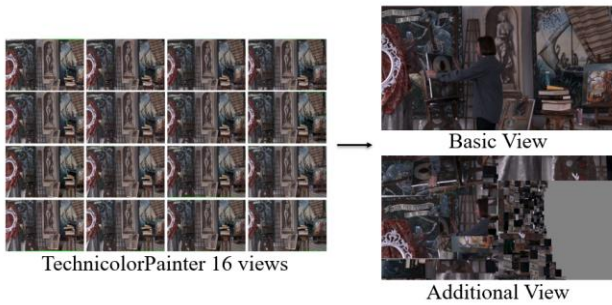


그림 2. *TechnicolorPainter* 영상의 바탕 시점(Basic view)과 추가 시점(Additional View) 예시

MIV 에서 기본이 되는 영상 처리 모델은 여러 입력 영상 중에서 정보를 가장 많이 담고 있는 시점의 영상을 계산해 바탕 시점(Basic view)으로 지정한 후, 나머지 시점의 영상들은 전체 영상이 아닌 바탕 시점과의 차이만을 보존하는 영상들로 따로 분리한다. 이러한 차이만을 갖고 있는 영상들을 추가 시점(Additional view)라고 표현하며, 추가 시점은 영상의 특성에 따라 바탕 시점이 미처 포함할 수 없는 주변 영상, 각 영상의 차이, 바탕 시점에서 볼 때 가려져 보이지 않는 물체 등을 포함한다[2].

이 때 바탕 시점과 추가 시점의 차이를 계산하는 크기 단위를 지정하게 되며, 지정된 단위 크기로 분리된 영상 정보들을 패치(Patch)라고 표현한다. 이 패치들이 모인 이미지들을 아틀라스(Atlas)라고 표현하며, 다시 모인 패치를 활용해 원래의 전체 영상들을 복원할 수 있도록 각 패치의 위치 정보 등을 메타데이터로 함께 전송한다. 이 처리 과정으로 현재 4K 해상도 크기로 제한된 코덱의 개수를 비약적으로 절감할 수 있는 장점을 갖는다.

그러나 이러한 추가 시점은 일반적인 영상과는 달리 공간적 상관성이 매우 떨어지며, 압축 과정에서 각 주변을 예측하기가 쉽지 않은 패턴으로 나타나 상대적으로 압축 효율이 낮아지는 단점이 존재한다. 본 논문에서는 추가 시점의 패치들이 시점이 가까운 다수의 영상에서 추출되어 자기 유사성의 패턴을 지니는 점에 착안하여 현재 스크린 콘텐츠(screen content coding) 기술과 하나인 화면 내 블록 카피(IBC) 기술에 대한 효율 분석을 시행하였다.

특히 MIV 에서 패치들을 조합할 때, 아틀라스의 크기를 최대한 줄이기 위하여 고안된 PiP(patch-in-patch) 옵션으로 추가 시점의 유효하지 않은 공간에 다시 패치들을 붙일 수 있다. PiP 옵션을 활성화하면 추가 시점 영상에서는 서로 비슷한 모양의 패치들이 한 영상 안에 더 많이 포함 될 수 있고, 이로 인해 화면 내 블록 카피 기술에서 이득을 얻을 수 있다.

그림 3 은 PiP 옵션을 활성화하여 처리된 1 번째 추가 시점 영상을 나타낸다. 여러 영상의 차이에서 도출된 패치들은 한 영상 내에서 자기 유사성의 패턴을 보인다. 이러한 패턴은 특히 바탕 시점만으로 모든 시점을 복원해내지 못해 일부 배경 영상이 추가 시점으로 넘어갈 때 잘 나타나는 경향을 보인다.

처리 결과는 영상 특성에 따라 크게 달라지는데, 예를 들어 영상 간의 시점이나 거리 차이가 큰 영상의 경우 패치들은 각 영상의 시점에서만 나타나는 고유한 정보를 더 많이 가지고 있으며, 바탕 시점보다는 추가 시점의 정보가 많을 수 있다.



그림 3. *InterdigitalHijack* 영상의 추가 시점 영상

## 3. 성능 분석

기존의 비디오 압축 표준 기술로는 현재 수요가 증가하고 있는 애니메이션, 컴퓨터 그래픽스와 같은 스크린 콘텐츠 영상에 대한 압축을 효율적으로 지원하지 못한다는 문제가 제기됨에 따라 HEVC-SCC 에 대한 표준화가 진행되었다[3].

HEVC-SCC 에서는 비교적 색상과 모양이 단조로운 특성을 반영한 색 공간 예측 기법, 팔레트 모드 등 효율적인 스크린 콘텐츠 영상 압축을 위한 기술들을 포함한다. 특히 그 중에서도 화면 내 블록 카피(IBC) 기법은 SCC 에 채택된 부호화 기술 중 가장 높은 성능 향상을 보여주고 있다[4].

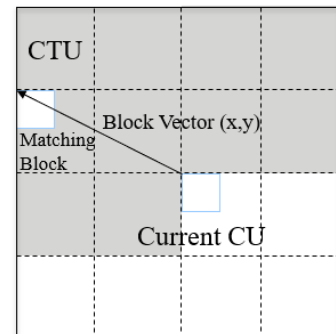


그림 4. 화면 내 블록 카피 기술 개념도[3].

해당 기술은 같은 프레임 내 또는 슬라이스 단위 내의 영역에서, 이미 복원된 샘플들 중 부호화 중인 블록과 가장 유사한 블록을 예측 블록으로 사용하는 기술이다. IBC 기법은 단순한 형태나 그래픽 특성상 텍스처 반복 등이 자주 일어나는 영상들에 대하여 예측성능을 높이기 위한 모드이다. 특히 반복적으로 같은 모양을 나타내는 텍스트와 그래픽이 혼합된 영상에서 IBC의 선택 비율이 상대적으로 높게 나타난다.

본 논문에서는 MIV 영상에서 스크린 콘텐츠 코딩 틀에 대한 성능 분석과, 자기 유사성 패턴을 나타내는 추가 시점에 대한 IBC 기법 성능 분석을 추가로 진행하였다. HM 16.16+SCM8.6, MIV 3.0 으로 현재 MPEG-I 에서 진행중인 표준 CTC (Common Test Condition)에 정의되어 있는 컴퓨터 그래픽스(CG: computer graphics) 영상인 Class A,B,C,N 에 대하여 실험을 진행하였으며, 완전한 프레임(full frame) 조건인 97 프레임으로 진행하였다[5].

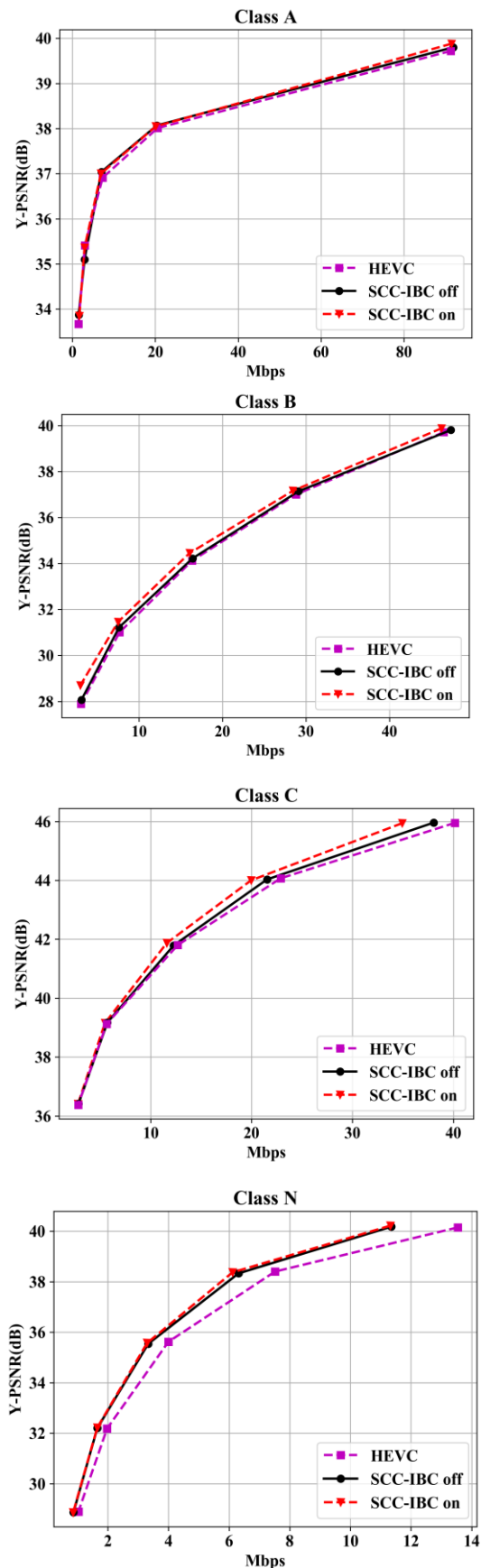


그림 5. 각 영상 별 RD-Curve

각 그래프는 HEVC, SCC+IBC 미적용, SCC+IBC 적용 영상의 RD-Curve 를 나타낸다.

Class A 의 경우 360 영상으로 바탕 시점만으로 대부분의 시점을 포함할 수 있으며, 추가 시점의 정보 또한 상대적으로 적다.

Class B,C 는 A 에 비해 각 영상의 시점이 상대적으로 작다. 그림 3 과 같이 바탕 시점 영상에서 미처 포함되지 못했던 주변 배경 영상이 추가 시점에 추가된 뒤에, PiP 옵션을 통해 유효하지 않은 다른 부분들이 패치로 채워지게 된다. 특히 Class C 는 바탕 시점 영상의 일부와 패치가 적절히 혼합되어 IBC 기법 적용 시 매우 높은 효율을 보여준다.

마지막으로 Class N 의 경우 화면을 가리는 물체(Chess) 가 다수 존재하여 각 영상마다 담고 있는 정보가 다른 특성을 보인다.

Sequence name	BD-rate saving(%)
ClassroomVideo(A)	0.66%
TechnicolorMuseum(B)	1.80%
InterdigitalHijack(C)	7.56%
NokiaChess(N)	0.82%

표 1. IBC 활성화 여부에 따른 Texture BD-rate

표 1 은 IBC 가 활성화되었을 때, 즉 SCC+IBC 미적용과 SCC+IBC 적용 간 텍스처 영상에서의 효율을 나타낸다. 깊이 맵 영상의 경우 높은 비트레이트로 갈수록 전체 비트레이트에서 차지하는 비율은 매우 적어지므로, 텍스처 영상에서의 효율이 높을수록 전체 비트레이트 효율 또한 크게 반영된다고 볼 수 있다.

Sequence name	IBC mode(%)
Class A (Basic)	4.06%
Class A (Additional)	16.95%
Class B (Basic)	9.35%
Class B (Additional)	11.62%
Class C (Basic)	5.63%
Class C (Additional)	19.43%
Class N (Basic)	1.75%
Class N (Additional)	3.97%

표 2. 각 영상의 IBC 선택 비율

표 2 는 IBC 기법을 적용한 각 영상 압축 시에 IBC 모드와 화면 내 예측(intra) 모드 중에서 선택된 비율을 나타낸다. 각 바탕 시점과 추가 시점 영상은 1 개 이상의 영상이 될 수 있으며, IBC 모드 선택 비율은 각 시점 카테고리의 평균으로 도출하였다.

효율이 가장 높았던 Class C 의 추가 시점 영상에서 다소 선택 비율이 높으며, 상대적으로 영상 간의 중복성이 적은 Class N 과 같은 경우 IBC 모드에서도 잘 선택되지 않는 경향을 보였다. Class A 의 추가 시점 영상의 경우 선택 비율은 높으나, 추가 시점 자체의 정보량이 크지 않기 때문에 IBC 기법으로도 효율을 크게 볼 수 없었던 것으로 관찰된다.

#### 4. 결론 및 향후 연구

MIV 의 추가 시점 영상은 처리 과정에서 자연적인 영상과는 다른 패턴이 나타나며, IBC 와 같은 스크린 콘텐츠 코딩 틀 적용에 따라 영상 특성에 따른 효율을 볼 수 있었다. 특히 주변 영상과 차분 패치가 한 영상 안에 적절히 포함될 경우, 텍스처 영상에서 최대 7.56%의 이득을 얻을 수 있음이 관찰되었다.

추가 시점 영상을 만드는 과정에서 패치들을 조합하는 방식에 따라 IBC 효율 또한 달라질 수 있다. 추가 시점 영상은 예측이 어려운 형태인만큼 전체 비트레이트에서 많은 비율을 차지하며, 추가 시점 영상을 인코더 친화적으로 압축하는 시도가 계속해서 이루어지고 있다.

IBC 는 현재 HEVC/VVC(Versatile Video Coding) 모두에서 화면의 회전을 고려한 Affine estimation 과 같은 예측 틀을 사용하지 않으므로, 추후 IBC 의 예측 성능을 높이기 위해 패치 회전 비활성화, 패치 순서 정렬 등 성능 탐색을 진행중에 있다.

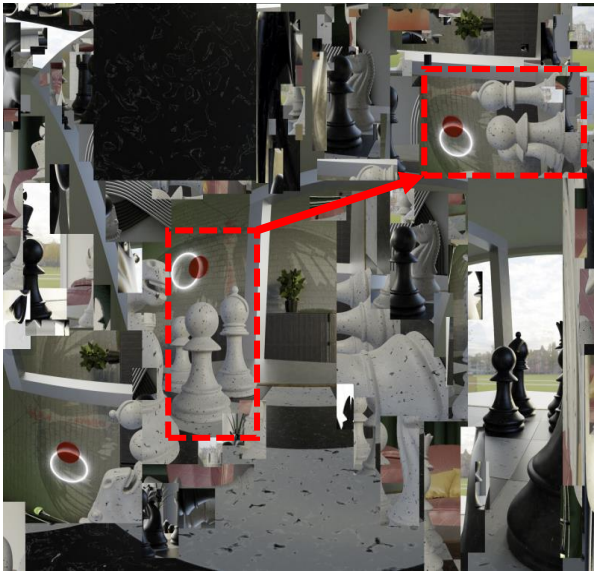


그림 6. Class N 추가 시점 영상에서의 패치 회전에 따른 IBC 예측 실패 예시

50-62, Jan. 2016,

[4] 마중현, 안용조, 심동규, “HEVC 스크린 콘텐츠 코딩 성능 향상을 위한 화면 내 블록 카피 기술 분석”, 2015, 방송미디어공학회 제 20 권 1 호.

[5] “Test Model 3 for Immersive Video”, 128th meeting of ISO/IECJTC1/SC29/WG11, MPEG 128/w18795.

#### Acknowledgement

이 논문은 2020 년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2018-0-00765, 6DoF 지원 초고화질 몰입형 비디오의 압축 및 전송 핵심 기술 개발). 이 논문은 또한 2020 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

#### 참고 문헌

- [1] “Call for Proposals on 3DoF+ Visual.” 2019. 125th MPEG meeting of ISO/IECJTC1/ SC29/WG11, MPEG125/n18145.
- [2] “Working Draft 3 of Immersive Video”. 2019. 128th MPEG meeting of ISO/IECJTC1/ SC29/WG11, MPEG128/w18794.
- [3] "Overview of the Emerging HEVC Screen Content Coding Extension," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp.