

MV-HEVC 기반 라이트필드 영상처리 기술 동향

김성빈¹, 정종범¹, 장동민¹, 김인애¹, 이순빈¹, 류일웅², 류은석¹

성균관대학교 컴퓨터교육과¹, 가천대학교 컴퓨터공학과²

beencooke@skku.edu, uof4949@skku.edu, jdm360@skku.edu, inaelk@skku.edu,

soonbinlee@skku.edu, dlfdnd96@gc.gachon.ac.kr, esryu@skku.edu

An Overview of MV-HEVC based Light Field Image Processing

Sungbin Kim¹, Jong-Beom Jeong¹, Dongmin Jang¹, Inae Kim¹, Soonbin Lee¹, Il-Woong

Ryu², Eun-Seok Ryu¹

Department of Computer Education, Sungkyunkwan University¹,

Department of Computer Engineering, Gachon University²

요 약

최근 3차원 영상에 대한 관심이 증가하면서 3차원 영상 정보 취득과 취득된 데이터를 다루는 라이트필드 영상 처리 기술에 대한 다양한 연구가 진행되고 있다. joint photographic experts group(JPEG)에서는 2014년부터 JPEG Pleno라는 이름으로 라이트필드 영상 정보 압축을 위한 표준화 과제를 진행 중이며, moving picture experts group(MPEG)에서도 라이트필드 영상 압축 기술에 대한 표준화 과제를 시작하였다. 본 논문은 최근 발표된 multiview high efficiency video coding(MV-HEVC) 기술을 이용한 라이트필드 영상처리 기술들을 소개한다.

1. 서 론

최근 3차원 영상에 대한 관심이 증가함에 따라 360 카메라나 라이트필드 카메라 등을 이용하여 취득한 3차원 영상을 효율적으로 처리하기 위한 여러 가지 방법들이 제시되었다 [1][2]. 인간이 완전한 입체감을 느끼기 위해서는 기본적으로 양안 시차, 운동 시차, 주시, 초점의 네 가지 물리적 요인을 충족해야 한다. 라이트필드 기술은 운동 시차 표현에 있어 아직 제약이 있지만, 나머지 세 가지 요인을 충족하기 때문에 3차원 영상을 표현하기에 적합한 기술이라고 볼 수 있다.

라이트필드는 여러 시점에서 획득된 빛의 방향과 세기를 이용하여 3차원 영상을 표현하며, 이와 같이 표현된 영상은 한 시점만을 표현하는 전통적인 영상과는 달리 여러 시점에서 취득되고 서로 간 중복성이 높은 영상들로 이루어진다. 이러한 특성으로 인해 라이트필드 영상은 일반적인 영상 압축 방식보다는 다시점 영상 압축 방식으로 성능 연구가 이루어졌다. 본 논문에서는 다시점 비디오 코딩을 이용하여 라이트필드 영상을 보다 효율적으로 압축하는 기술들을 소개한다.

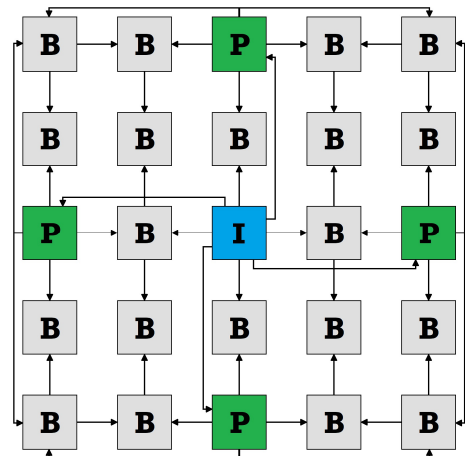
본 논문의 구성은 다음과 같다. 2절은 다시점 비디오 코딩 기술인 MV-HEVC를 소개하고, 3절은 MV-HEVC를 이용한 라이트필드 영상 압축 기술들을 소개한다. 마지막 4절에서는 본 논문의 결론을 맺는다.

2. 관련 연구

2.1 MV-HEVC

MV-HEVC는 MPEG과 ITU-T의 video coding experts group(VCEG)이 함께 개발한 영상 부호기인 HEVC를 확장한 다시점 비디오 코딩 기술이다. MV-HEVC는 동시에 획득된 여러 영상들 간의 시간적, 공간적 중복성 뿐 아니라 뷰들 간의 중복성을 제거하기 위해 개발되었다[3].

MV-HEVC는 I-B-P-B-P 형태로 이루어진 IBP-프레임 패턴과 양방향으로부터 움직임을 예측하는 B-프레임을 계층적



(그림 1) I-프레임 중앙 배치 구조

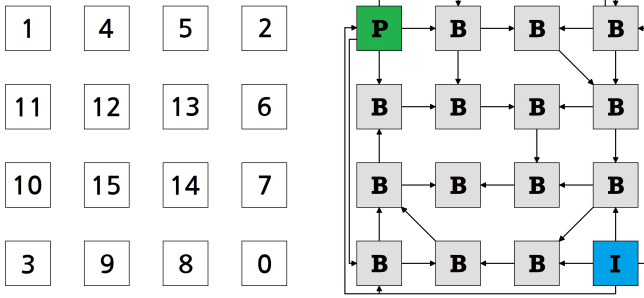
으로 사용하는 계층적 B-프레임 구조를 주로 사용하기 때문에 여러 개의 시점 영상을 가지는 라이트필드 영상을 압축하기에 적합하다. 이러한 특성으로 인해 다시점 비디오 코딩 기술을 이용한 라이트필드 영상처리에 대한 연구가 진행되고 있으며, 3절에서 그 중 3가지 기술을 소개할 것이다.

3. 다시점 비디오 코딩 기반 라이트필드 영상처리

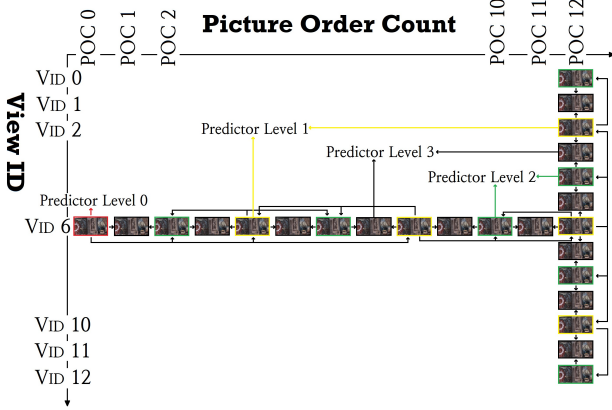
3.1 I-프레임 중앙 배치 기법

I-프레임 중앙 배치 기법은 그림 1과 같이 라이트필드 영상의 정중앙에 있는 시점 영상을 I-프레임으로 부호화하고, 이를 참조하여 나머지 시점 영상들을 부호화하는 기법이다[4].

중심축 위에 있는 시점 영상들은 부호화된 I-프레임을 참조하여 IBP-프레임 패턴으로 부호화되는데, P-프레임은 I-프레임만을 참조하고, B-프레임은 앞뒤의 두 시점 영상을 참조한다. 네 모퉁이의 시점 영상들은 가장 가까운 두 개의 P-프레임을 참조하여 부호화하고, 나머지 시점 영상들은 가장 가까운 두 시점 영상을 참조하여 부호화한다.



(그림 2) 나선형 스캐닝 구조(a), IPBBBB 참조 구조(b)



(그림 3) 예측 레벨 할당 및 참조 구조

I-프레임을 중앙에 배치함으로써 중심부의 시점 영상들 간의 높은 유사성 때문에 발생하는 압축 효율 향상 등의 이점을 취할 수 있지만, 영상의 특성에 따라 최적의 효율을 내지 못할 가능성도 있다.

3.2 IPBBBB 코딩 구조

IPBBBB 코딩 구조는 양방향으로 참조하는 B-프레임의 비중을 늘려 시점 영상 간의 중복성을 최대한 활용하여 압축률을 높이는 구조이다[6].

코딩 순서 또한 압축률을 높이기 위해 그림 2의 a와 같이 가장자리부터 중심까지 나선형으로 회전하며 매겨지는데, 네모퉁이를 가장 먼저 코딩한다[7]. I-프레임과 P-프레임은 압축률 향상을 위해 그림 2의 b와 같이 각각 하나씩만 존재한다. B-프레임은 가장 가까운 두 개의 시점 영상을 참조하는데, 거리가 같을 경우 인간의 시각 체계를 고려하여 수평 방향을 우선으로 참조한다.

이러한 코딩 구조는 라이트필드 영상의 압축률을 상승시키는 반면, 임의 접근(random access)을 위한 복잡도가 매우 증가한다. 또한, 임의의 시간과 임의의 시점에 대한 영상을 보려면 모든 시점을 한 번에 디코딩해야하는 제약이 있다.

3.3 예측 레벨 할당 기법

예측 레벨 할당 기법은 라이트필드 영상을 이루는 시점 영상 각각에 2가지 예측 레벨(predictor level)을 부여하고, 예측 레벨에 따라 적절한 양자화 파라미터 값을 할당하여 압축 효율을 높이는 기법이다[5].

예측 레벨은 그림 3과 같이 View ID와 picture order count(POC)에 의해 정해지며, 각각 0에서 3까지 4단계의 예측 레벨로 나뉜다. 영상마다 부여된 2가지 예측 레벨을 이용하면 총 16가지 가중치 조합이 만들어지는데, 가중치는 각 시점 영상의 양자화 파라미터 값을 산출하는 과정에 사용한다.

그림 3은 예측 레벨 할당과 시점 영상 간 참조 구조를 나타내며, 15×15의 시점 영상으로 이루어진 라이트필드 영상의

중앙 13×13에 해당하는 시점 영상만을 압축에 사용하는 과정을 보여준다.

4. 결론

소개한 세 가지 기술은 MV-HEVC를 이용하여 라이트필드 영상을 처리하는 과정 중 예측 구조에 변화를 주어 성능을 개선하는 공통점을 가지며, 이외에도 여러 가지 방법들이 제시되고 있어 다시점 비디오 코딩을 이용한 라이트필드 영상 처리 기술은 더욱 개선될 것으로 보인다.

앞으로도 라이트필드 영상을 더 효율적으로 처리하는 방법이 제시되었지만, 3차원 영상을 완벽하게 표현하기 위해선 영상을 취득할 수 있는 범위가 좁아 발생하는 운동 시차의 제약을 해결해야 할 것이다.

ACKNOWLEDGMENT

본 연구는 한국전력공사의 2016년 선정 기초연구개발과제 연구비에 의해 지원되었음 (과제번호 : R17XA05-68), 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

참고 문헌

- [1] S. Lee, D. Jang, J. Jeong, and E. S. Ryu, "Motion-constrained tile set based 360-degree video streaming using saliency map prediction", ACM Network and Operating System Support for Digital Audio and Video 2019 (NOSSDAV2019), 2019.
- [2] J. Jeong, S. Lee, D. Jang, and E. S. Ryu, "Towards 3DoF+ 360 video streaming system for immersive media", IEEE Access, Vol 7, pp.136399 - 136408, 2019.
- [3] 정종범, 이순빈, 장동민, 류일웅, Tuan Thanh Le, 유재성, 류은석, "실감형 미디어를 위한 3DoF+ 360 비디오 압축 시스템 구현", 한국방송미디어공학회 방송공학회논문지, 제24권 제5호, 2019.
- [4] J. Khoury, "A New prediction structure for efficient MV-HEVC based light field video compression", International Conference on Computing, Networking and Communications (ICNC), Vol.1, pp.588-591, 2019.
- [5] W. Ahmad, "Interpreting plenoptic images as multi-view sequences for improved compression", IEEE International Conference on Image Processing (ICIP), 2017.
- [6] V. Avramelos, "Random access prediction structures for light field video coding with MV-HEVC", Multimedia Tools and Applications, 2020.
- [7] S. Shi, "Efficient compression method for integral images using multi-view video coding", Proceedings of the 18th IEEE International Conference on Image Processing (ICIP), pp.137 - 140, 2011.