

JCT-3V 3차원 비디오 부호화 표준 기술 동향

호요성 광주과학기술원 정보통신공학부 교수



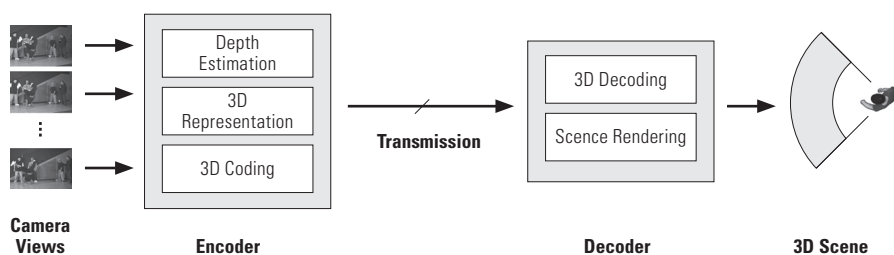
1. 머리말

3차원 비디오는 기존의 2차원 비디오에 입체감을 더한 것이다. 사람은 좌우 눈으로 약간 다른 두 영상을 보고 입체감을 느끼기 때문에, 극장에서 3차원 영화를 감상할 때 입체안경을 사용한다. 2000년대 말부터 3차원 영화가 큰 성공을 거두면서 3차원 TV, 3차원 노트북, 3차원 휴대폰 등의 전자제품이 출시되었으며, 앞으로 다양한 분야에서 3차원 영상을 접할 수 있을 것이다.

3차원 영상 콘텐츠를 만들기 위해서는 다시점의 색상영상과 깊이영상이 필요하다. 그러나 촬영할 수 있는

시점의 수가 제한되어 있어 촬영된 영상으로부터 임의 시점의 영상을 생성해야 한다[1]. 이를 중간영상 합성이라 하며, 인접 시점의 색상영상과 깊이영상을 기반으로 한 투영 및 보간 작업이 수행된다. 3차원 다시점 영상의 경우 시점 수에 비례하여 데이터양이 증가하기 때문에 이를 줄이고자 효율적인 압축 기술이 연구되고 있다. 3차원 비디오를 압축해 발생한 비트열을 전송하고, 이를 수신기에서 복호화 작업을 거친 후 3차원 비디오를 재현할 수 있다. [그림 1]은 3차원 비디오 시스템의 기본 구조를 나타낸 것이다[2].

3차원 비디오 압축 표준을 만들기 위해 국제 표준



[그림 1] 3차원 비디오 시스템의 기본 구조

화 기구인 ISO/IEC 산하의 MPEG(Moving Picture Experts Group)은 3차원 비디오 부호화(3DVC: 3D Video Coding) 표준화 작업을 수행했다[3]. 이후 ITU-T 산하의 VCEG(Video Coding Experts Group)과 JCT-3V(Joint Collaborative Team on 3D Video Extension Development) 그룹을 설립하고, 2012년 7월에 첫 회의를 가졌다.

본 고는 JCT-3V에서 진행한 3차원 비디오 부호화 기술을 소개한다. 세부적으로 AVC(Advanced Video Coding)에 기반을 둔 부호화 기술과 HEVC(High Efficiency Video Coding)에 기반을 둔 부호화 기술로 나눠 설명한다. 또한, 표준화 작업에서 이뤄지는 부호화 기술의 평가 방법과 향후 이어질 활동에 대해 간단히 살펴본다.

2. AVC-기반 3차원 비디오 부호화(3D-AVC)

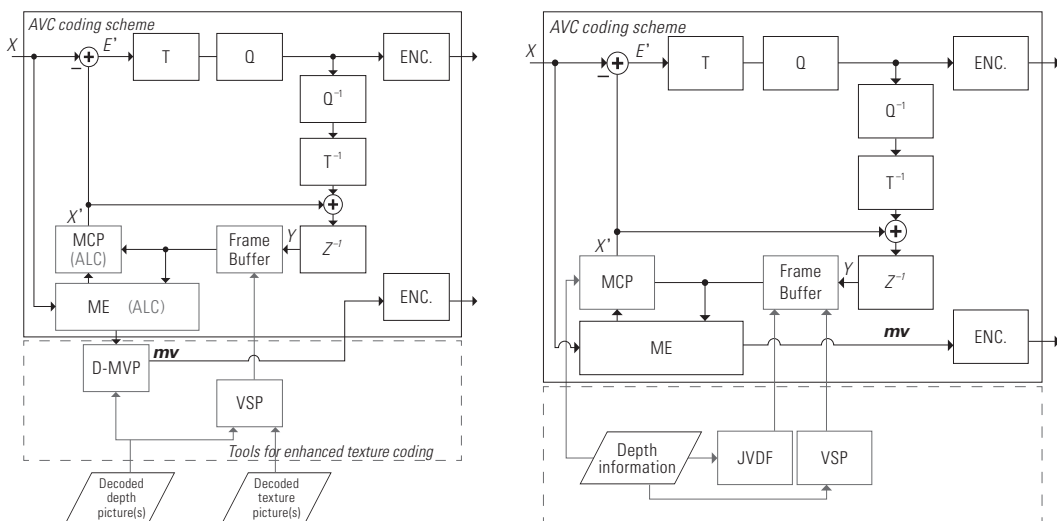
3D-AVC는 기존의 2차원 영상 부호화 표준인 H.264/AVC를 확장해 다시점의 색상영상과 깊이영상의 부호화에 이용하는 것이다. 3D-AVC에서는 색상영

상과 깊이영상의 상관관계를 이용하므로 부호화 순서가 중요하다. 색상영상을 T, 깊이영상을 D, 기준 시점을 0이라 하고, 두 개의 참조 시점을 각각 1과 2라고 하면, 부호화는 T0-D0-D1-D2-T1-T2 순서로 진행되며 이를 깊이 우선 부호화(Depth first coding)라고 한다[4]. [그림 2]는 3D-AVC에서 색상영상과 깊이영상의 부호화 순서도를 나타낸 것이다. 일반적으로 깊이영상은 색상영상에 비해 단순하고 부호화 중요성이 상대적으로 떨어져 깊이영상의 해상도를 다운 샘플링해 부호화한다.

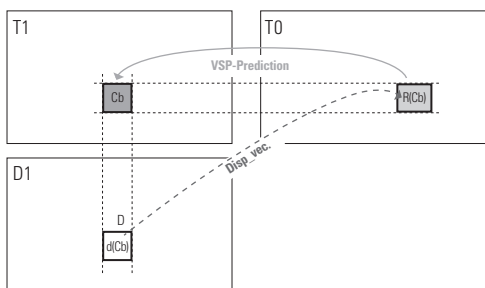
2.1 영상합성 예측(VSP: View synthesis prediction)

영상합성 예측은 복원된 인접 시점의 영상을 이용해 부호화할 시점의 영상을 합성하고, 이를 참조영상으로 이용하는 방법이다. 합성된 영상은 인접 시점의 영상보다 부호화할 영상과 더 비슷하므로 더 효율적인 부호화를 가능하게 한다. 이 기술은 색상영상과 깊이영상 부호화에서 모두 사용된다.

최근 영상합성 예측에서는 후방 투영(Backward warping)을 많이 사용한다. 이는 참조할 합성영상을



[그림 2] 3D-AVC에서 색상영상(좌)과 깊이영상(우) 부호화 순서도



[그림 3] 영상 합성 예측

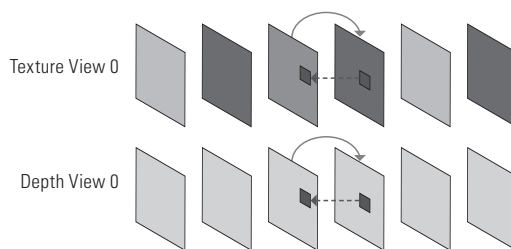
직접 생성하는 것이 아니라, 깊이영상으로부터 얻은 변위 벡터(Disparity vector)를 이용해 이미 부호화한 색상영상의 데이터에서 예측값을 얻는다[5]. [그림 3]에서 보듯 이 과정은 블록별로 수행된다. 이 방법은 색상영상 부호화에만 사용된다.

2.2 색상-깊이 움직임 예측

앞서 설명한 것처럼 기준 시점에서는 색상영상을 먼저 부호화한 뒤 깊이영상을 부호화한다. 이때 색상영상과 깊이영상의 움직임 정보가 매우 비슷하므로, 색상영상에서의 매크로블록(Macroblock) 종류, 움직임 벡터, 참조 인덱스 등의 움직임 정보를 깊이영상 부호화에 그대로 사용한다. [그림 4]는 색상-깊이 움직임 예측 기술을 보여준다.

2.3 적응적 밝기 보상(ALC: Adaptive luminance compensation)

다시점 영상 부호화에는 다시점 영상 촬영 시 발생할 수 있는 영상의 시점 간 밝기 차이를 보상하는 밝기 보상 기술이 포함되어 있다. 이 기술에서는 색상 영상의 밝기 차이를 줄이기 위해 적응적으로 밝기 값을 조정한다. 움직임 벡터를 통해 참조할 영역을 찾은 뒤, 가중치를 이용해 예측한 블록의 색상 밝기 값을 보정한다[6].



[그림 4] 색상-깊이 움직임 예측

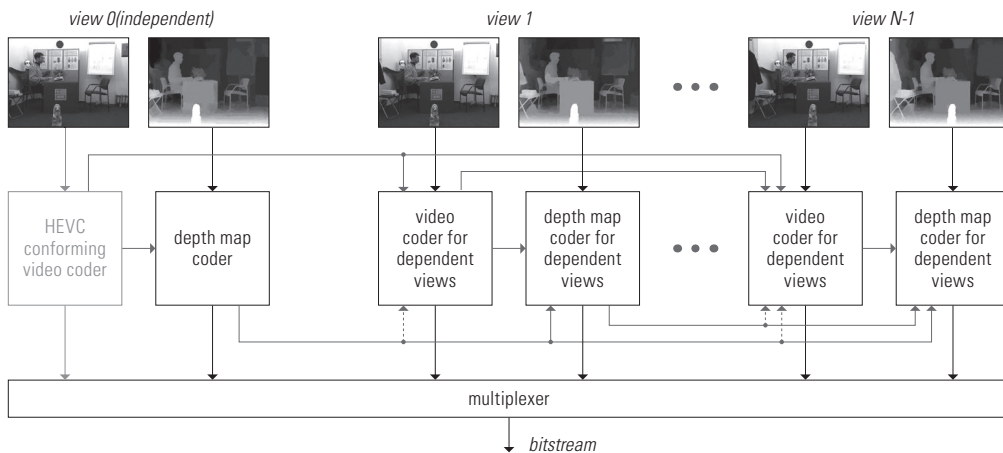
2.4 시점간 깊이 필터링(JVDF: Joint inter-view depth filtering)

시점간 깊이 필터링은 시점간 존재하는 깊이정보의 불규칙함을 보상한다. 깊이영상은 오차에 민감하여 시점간 일관성이 매우 중요하다. 다시점 깊이영상을 투사시켜 깊이값의 차이를 분석하고, 필요한 경우 평균값을 이용해 오차를 보정한다. 이러한 깊이값 보정은 깊이영상의 움직임 예측/보상 과정을 더 효율적으로 수행하기 위해서 사용된다.

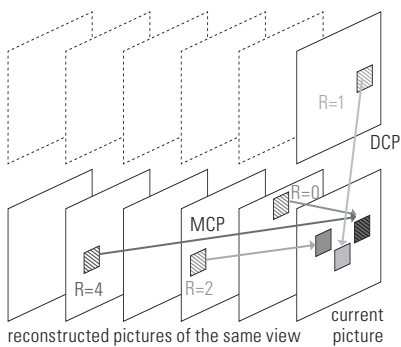
3. HEVC-기반 3차원 비디오 부호화(3D-HEVC)

HEVC는 2013년 1월에 완성된 고해상도 영상 부호화를 위한 새로운 표준으로 기존의 H.264/AVC에 비해 약 2배의 압축 효율을 제공한다. 3D-HEVC는 다시점의 색상영상과 깊이영상의 부호화에 이용할 수 있도록 HEVC를 확장한 것이다.

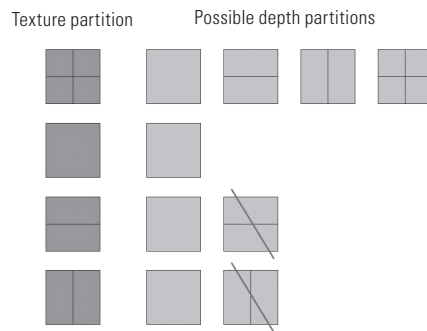
3D-HEVC의 부호화 순서는 2장에서 설명한 3D-AVC의 부호화 순서와 다르다. 3D-HEVC에서는 T0-D0-T1-D1-T2-D2 순으로 부호화가 수행된다[7]. 이를 색상 우선 부호화(Texture first coding)라고 하며, 항상 한 시점의 색상영상과 깊이영상을 부호화한 후에 다른 시점의 색상영상과 깊이영상을 부호화한다. 색상영상이 먼저 부호화되기 때문에 색상영상 부호화에서의 정보가 깊이영상 부호화에서 사용된다. [그



[그림 5] 3D-HEVC의 부호화 및 예측 구조



[그림 6] 변위 보상 예측



[그림 7] 깊이 쿼드트리 분할 제한

림 5는 3D-HEVC의 부호화 및 예측 구조를 나타낸다. 3D-AVC와 또 다른 점은 다운 샘플링된 깊이영상을 사용하지 않고 원래 깊이영상의 해상도를 그대로 사용한다. 이는 깊이영상을 더 정확하게 부호화하기 위함이다.

3.1 변위보상 예측(DCP: Disparity compensated prediction)

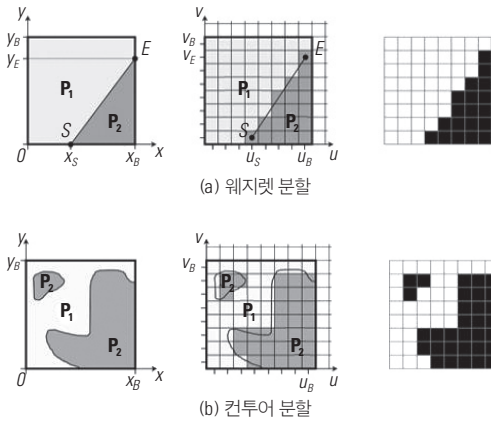
변위보상 예측은 기존의 움직임보상 예측(MCP: Motion compensated prediction)과 비슷한 기술이다. 움직임보상 예측은 한 시점에서 시간 축으로 움직임을 예측하지만, 변위보상 예측은 [그림 6]과 같이 시점 축으로 변위를 통해 예측한다. 이 기술은 참조시점의 색상영상 부호화에만 사용된다.

3.2 깊이 쿼드트리 분할 제한(DQL: Depth quadtree limitation)

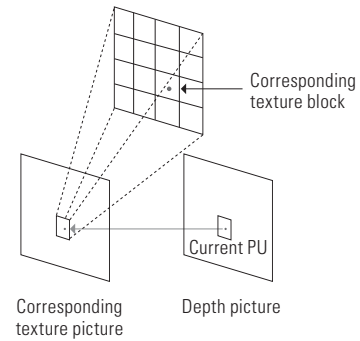
HEVC에서는 영상을 H.264/AVC처럼 매크로블록이 아닌 쿼드트리(Quadtree) 방식으로 분할하는데, 색상영상과 깊이영상의 쿼드트리는 밀접한 관계가 있다[8]. 또한, 깊이영상은 색상영상에 비해 단순하므로 분할도 더 단순하다. 색상영상의 분할 구조에 기반을 두어 깊이영상의 분할 구조를 전송하면 보내는 정보량을 줄일 수 있다.

3.3 깊이 모델링 모드(DMM: Depth modeling modes)

깊이 모델링 모드는 깊이영상 화면 내 부호화(Intra coding)에서 사용된다. HEVC는 35개의 화면 내



[그림 8] 깊이 모델링 모드에서 영역 분할 방법



[그림 9] 색상-깊이 움직임 예측

모드를 사용하는데, 3D-HEVC에서는 깊이값이 급격하게 변하는 경계(Edge) 영역을 더욱 효율적으로 부호화하고자 4개의 깊이 모델링 모드를 추가로 사용한다. 깊이영상 블록의 분할 모양에 따라 웨지렛(Wedgelet) 분할 또는 컨투어(Contour) 분할로 구분된다. [그림 8]에서 볼 수 있듯 웨지렛은 직선으로 분할된 것을 의미하고, 컨투어는 임의의 모양을 포괄하는 형태를 의미한다. 3D-HEVC에서는 3개의 웨지렛 모드와 1개의 컨투어 모드를 사용한다.

3.4 색상-깊이 움직임 예측

3D-HEVC에서는 부호화 순서상 같은 시점에서 항상 색상영상이 깊이영상보다 먼저 부호화된다. 색상영상과 깊이영상의 움직임 정보는 비슷하므로 깊이영상을 부호화할 때는 동일 위치의 색상 블록(Corresponding texture block)이 병합 후보(Merge candidate)로 추가된다[9]. [그림 9]는 색상-깊이 움직임 예측을 보여준다.

4. 기술평가 방법(Common test conditions)과 향후 일정

3차원 비디오 부호화 기술 평가를 위해 3D-AVC

를 위한 3DV-ATM과 3D-HEVC를 위한 3DV-HTM 참조 소프트웨어를 사용한다[10]~[11]. 중간 시점 영상합성을 위해서는 3DV-HTM에 포함된 영상합성 소프트웨어를 사용한다. 성능 평가를 위해 <표 1>과 같이 총 7개의 YUV 4:2:0 형식의 화소당 8비트 실험 영상을 사용한다. 이 중 'Undo_Dancer'와 'GT_Fly'는 컴퓨터 그래픽스로 생성되어 정확한 깊이값을 갖지만, 다른 실험 영상의 깊이영상은 다시점 카메라 시스템으로 촬영한 색상영상으로부터 깊이 예측(Depth estimation)을 통해 구한 것이다. 2013년 4월에 3D-AVC에 대한 개정표준안(DAM: Draft Amendment)이 발행됐으며, 2014년에 표준 제정이 이뤄질 예정이다. 3D-HEVC는 Scalable HEVC 표준화 작업이 완료된 후 마무리될 예정이다.

<표 1> 3차원 비디오 부호화 실험 영상

시퀀스 ID	시퀀스 이름	부호화 시점	제공 기관	해상도
S01	Poznan_Hall2	7-6-5	Poznan	1920×1088
S02	Poznan_Street	5-4-3	Poznan	
S03	Undo_Dancer	1-5-9	Nokia	
S04	GT_Fly	9-5-1	Nokia	
S05	Kendo	1-3-5	Nagoya	1024×768
S06	Balloons	1-3-5	Nagoya	
S08	Newspaper	2-4-6	GIST	

5. 맺음말

본 고에서는 3차원 비디오 부호화 표준을 소개하고, 핵심 부호화 기술을 간략히 살펴보았다. 3차원 비디오 부호화 표준화 작업은 현재 진행 중이며, 3D-AVC는 2014년, 3D-HEVC는 그 후에 완성될 예정이다. 향후 제정된 국제 표준을 기반으로 3차원 입체영상 관련 사업들이 더욱 크게 성장할 것으로 기대된다.

[참고문헌]

- [1] A. Kubota, A. Smolic, M. Magnor, T. Chen, and M. Tanimoto, 'Multi-View Imaging and 3DTV - Special Issue Overview and Introduction,' IEEE Signal Processing Magazine, vol.24(6), 2007. 11.
- [2] 호요성, 이천, '다시점 3차원 영상 처리 및 부호화 기술, 진샘미디아, 2013. 4.
- [3] 호요성, 이천, '3차원 비디오 부호화 표준 기술', TTA 저널, 135호, pp. 79-83, 2011. 6.
- [4] D. Rusanovskyy, F. Chen, L. Zhang, and T. Suzuki, '3D-AVC Test Model 5,' 3rd JCT3V Meeting, JCT3V-C1003, 2013. 1.
- [5] W. Su, D. Rusanovskyy, and M. Hannuksela, '3DV-CE1.a: Block-based View Synthesis Prediction for 3DV-ATM,' 1st JCT3V Meeting, JCT3V-A0107, 2012. 7.
- [6] M. Mishurovskiy, A. Fartukov, I. Kovliga, and J. Lee, '3D-CE2.a Results on Inter-view Coding with Adaptive Luminance Compensation,' 2nd JCT3V Meeting, JCT3V-B0031, 2012. 10.
- [7] G. Tech, K. Wegner, Y. Chen, and S. Yea, '3D-HEVC Test Model 3,' 3rd JCT3V Meeting, JCT3V-C1005, 2013. 1.
- [8] J. Jung and E. Mora, '3D-CE3.h: Depth Quadtree Prediction for 3DHTM 4.1,' 2nd JCT3V Meeting, JCT3V-B0068, 2012. 10.
- [9] Y. Chen, J. Lin, Y. Huang, and S. Lei, '3D-CE3.h results on removal of parsing dependency and picture buffers for motion parameter inheritance,' 3rd JCT3V Meeting, JCT3V-C0137, 2013. 1.
- [10] 3DV-ATM, <http://mpeg3dv.research.nokia.com/svn/mpeg3dv>.
- [11] 3DV-HTM, https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware. 

정보통신 용어해설

인메모리 컴퓨팅 IMC(In Memory Computing) [컴퓨터]

컴퓨터 하드디스크에 저장된 데이터가 아니라 RAM의 데이터로 질의하는 기법. 서버상에서 거의 동시에 계산 작업을 수행하여 대량의 데이터를 다른 곳으로 이동하여 처리하지 않아도 되기 때문에 계산 속도가 획기적으로 개선된다. 메모리에서 고급 통계 분석이 가능하기 때문에 빠른 속도로 쏟아져 나오는 빅데이터를 때맞게 분석해 의사 결정에 반영할 수 있는 것이다.

