

[멀티미디어응용] MPEG 3차원 비디오 부호화(3DVC) 표준화 작업

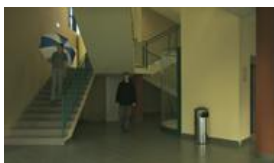
3차원 비디오 압축 표준화

요즘 많은 관심을 가지고 연구되고 있는 3차원 영상 기술은 다시점 카메라로 획득한 다시점 영상으로 3차원 영상을 재현한다. 그러나 모든 시점에 대해 다시점 카메라로 촬영하는 것은 한계가 있고 비효율적이기 때문에 깊이 영상을 이용하여 임의의 시점 영상을 합성하여 보다 많은 시점을 생성하는 방법이 이용되고 있다. 하지만 이 방법은 스테레오 영상에 비해 데이터 양이 크게 증가하므로 표준화된 영상 포맷과 효과적인 부호화 기술이 필요하다. 최근 국제 표준화 기구인 MPEG(Moving Picture Experts Group)의 3차원 비디오 부호화(3DVC: 3D Video Coding) 그룹에서는 이러한 영상들의 데이터 형식(data format)을 정의하고 효과적인 부호화 기술들을 표준화하는 활동을 진행하고 있다.

2011년 3월에 열린 제96차 MPEG 회의에서는 다시점 비디오 영상을 효율적으로 부호화하고 고품질의 중간 영상을 만들기 위한 3차원 비디오 부호화 표준을 만들기 위해 제안요청서(MPEG문서 N12036)가 만들어졌다. 본 고에서는 이 제안요청서에 정의된 내용들을 다루고자 한다.

3차원 비디오 테스트 시퀀스의 분류

표준화 작업을 진행하기 위해서는 실험 환경이 정의되어야 한다. 입력영상이 되는 다시점 색상 영상과 깊이 영상은 YUV 4:2:0 형식을 이용하고, 각 화소당 8비트로 표현한다. 이번 실험에 사용할 테스트 영상은 아래의 표와 같다. Nokia가 제공한 'Undo_Dancer'와 'GT_Fly'는 다시점 카메라로 획득한 것이 아닌 컴퓨터 그래픽스로 만들어진 시퀀스들로 이를 통한 정확한 깊이 영상이 배포되었다. 이는 깊이 측정 소프트웨어인 DERS(Depth Estimation Reference Software)를 이용해 깊이를 구한 다른 시퀀스들과 차별된다.



(a) Poznan_Hall2



(b) Poznan_Street



(c) Undo_Dancer



(d) GT_Fly



(e) Kendo



(f) Balloons



(g) Lovebird1



(h) Newspaper

<그림 1> 8 개의 테스트 시퀀스

각 테스트 시퀀스는 2-시점과 3-시점 두 개의 실험 조건 모두를 이용해야 한다. 여기서 2-시점은 양안식 디스플레이, 3-시점은 다시점 디스플레이 시스템을 고려한 시나리오이다. 예측 구조로는 2-시점은 I-P 구조, 3-시점은 I-B-P 구조를 사용한다. 부호화 방법의 성능을 비교하기 위하여 두 가지의 서로 다른 부호화 방법을 이용할 수 있다.

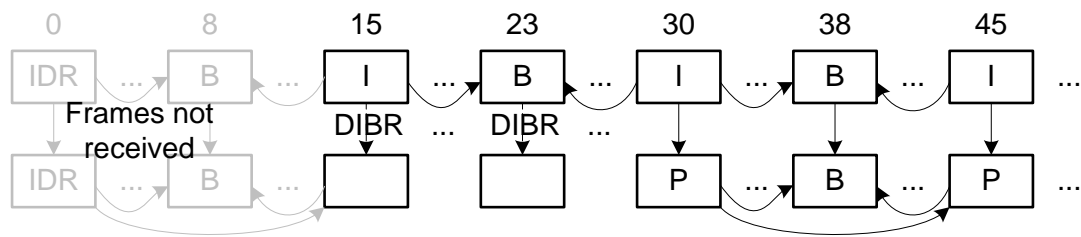
- 1) AVC-Compatible: 기존의 H.264/AVC 부호화 표준과의 호환성을 유지하기 위해 다시점 비디오 부호화(multi-view video coding) 방법을 이용하여 성능을 평가한다. 부호화기는 JMVC 8.3.1을 이용한다.
- 2) HEVC-Compatible & Unconstrained: HEVC와의 호환성을 유지하기 위해 HM2.0을 이용하여 성능을 평가한다.

2011년 4월 15일에 최종 제안요청서(CfP)가 배포되어 기술 등록이 시작되었다. 10월 1일에 주관적 화질 평가를 거쳐 2011년 11월 26일~12월 2일에 열린 제98차 회의에서는 제안된 기술 평가가 이루어졌다. 2012년 2월에 열릴 제99차 회의에서는 AVC 카테고리의 working draft가 진행될 예정이다.

제98차 MPEG 3DV 회의 내용

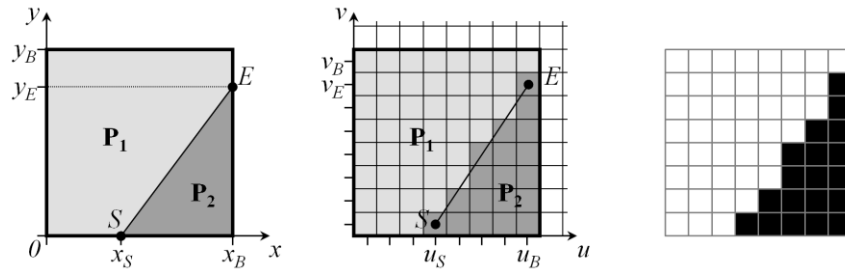
제98차 MPEG 회의에서 기술 평가에서 높은 평가를 받은 기술들을 바탕으로 두 개의 카테고리 AVC, HEVC를 기반으로 각각 3DV-ATM, 3DV-HTM으로 불리는 테스트 모델(TMuC: Test Model under Consideration)을 만들어 배포할 예정이다.

AVC-호환에서는 Nokia가 평가에서 가장 좋은 결과를 보였는데 반영된 기술들을 보자면 영상 합성 예측(view synthesis prediction), 단계적 화면 전환(gradual view refresh) 등이 있다. 영상 합성 예측은 영상을 합성할 때 인접 시점에서 복원한 깊이 영상을 이용하는 방법이다. 단계적 화면 전환은 미리 선택된 시점들을 정밀하게 복호화하고 나머지 시점들은 합성 기술을 이용하여 재건하는 기술이다.



<그림 2> 단계적 화면 전환에서의 복호화 작업

또한, HEVC-호환에서는 HHI의 결과가 가장 우수했다. HHI가 제안한 대표적인 기술로는 윤곽-기반 깊이 모델링(depth modeling modes with wedgelet/contour), 변이 보상 예측(disparity compensated prediction) 등이 있다. 윤곽-기반 깊이 모델링은 깊이 영상에서 깊이값이 경계에서 급격하게 변하는 영역과 천천히 변하는 영역으로 구분하여 wedgelet과 contour를 이용하여 분할하는 방법이다. 변이 보상 예측은 기존의 움직임 보상 예측(motion compensated prediction)을 대신하여 시점 방향으로 변위값을 예측하는 방법이다.



<그림 3> 윤곽-기반 깊이 모델링

이번 기술 평가에서 AVC와 HEVC에서 각각 Nokia와 HHI의 기술들이 가장 높은 평가를 받아 이들 기술들을 바탕으로 테스트 모델(TMuC: Test Model under Consideration)을 배포되고 계속 업데이트될 예정이다. TMuC은 3DV-ATM과 3DV-HTM으로 두 개가 있고 전자는 JM, 후자는 HM을 기반으로 만들어졌다. 앞으로 열릴 회의들에서는 연구기관들이 이 테스트 모델들에 기술을 구현해 기고할 전망이다.

현재 MPEG에서 작업하고 있는 3차원 비디오 부호화 표준은 향후 3차원 실감방송과 가상현실 환경을 구현하는데 핵심적인 영역이므로, 국내 여러 기관들의 큰 관심과 적극적인 참여가 필요하다.

호요성 (광주과학기술원 실감방송연구센터 교수, hoyo@gist.ac.kr)