

# 초다시점 및 자유시점 비디오 기술의 표준화 동향

## 1. 초다시점 비디오와 자유시점 비디오의 필요성과 목적

일본은 2020년 월드컵을 개최하면 축구장에 카메라를 둘레로 배치하여 3차원 영상으로 축구 시청을 가능하게 하는 것을 목표로 하고 있다. 이에 따라 특히 일본 연구소들이 초다시점 비디오(super multi-view video) 및 자유시점 비디오(free navigation video) 기술 개발 및 표준화 활동에 적극적으로 참여하고 있다 [1].

그림 1은 3차원 영상 캡처와 디스플레이 성능에 대하여 무어의 법칙(Moore's law)를 적용하여 분석한 것이다 [2]. 그래프에서 화소의 수와 시점의 수를 곱한 수치를 분석한 것인데 초다시점 비디오와 자유시점 비디오가 먼 미래가 아님을 알 수 있다. 기대할 수 있는 응용 사례로는 초다시점 비디오를 활용하여 수백 시점의 데이터를 요구하는 3차원 홀로-디스플레이 시연, 자유시점 비디오를 활용하여 축구 TV 중계를 시청할 때 원하는 시점에서의 자유로운 이동 등을 꼽을 수 있다. 본고에서는 초다시점 비디오와 자유시점 비디오에 관련된 기존 기술 및 표준화 동향을 살펴보고자 한다.

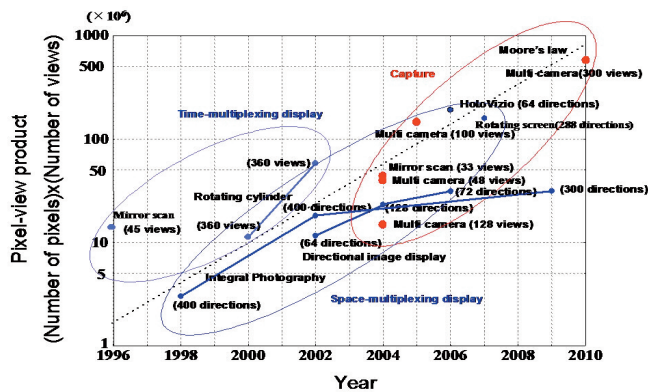


그림 1. 3차원 영상 캡처 및 디스플레이 성능 개발 흐름

## 2. MPEG FTV 그룹의 표준화 활동

ISO/IEC 산하 MPEG(Moving Picture Experts Group)의 FTV(Free-viewpoint Television) 그룹에서는 2000년대부터 다시점 비디오와 3차원 비디오 기술들에 대한 표준화 활동을 적극적으로 진행하였다. FTV 그룹의 활동에 대하여 단계를 분리할 수 있는데 그림 2는 활동 단계 역사를 나타낸다. 1단계에서는 2004년부터 다시점 비디오 부호화(Multi-view video coding, MVC)에 대한 기술을 개발하였다 [3, 4].

다시점 비디오 부호화에서는 시점 간 정보 참조 등의 기술로 중복성을 줄여 데이터를 전송하고 복호기에서 원래 시점의 수만큼 복호화하게 된다. 다시점 비디오 부호화는 2009년에 완료되었고 현재 시장에서는 3D 블루레이 등에서 이용되고 있다. 영상 복호화 이후에 진행되는 임의시점 영상 합성은 다시점 비디오 부호화 표준에 포함되지 않는다. 그림 3은 다시점 비디오 부호화 프레임워크를 설명한다.

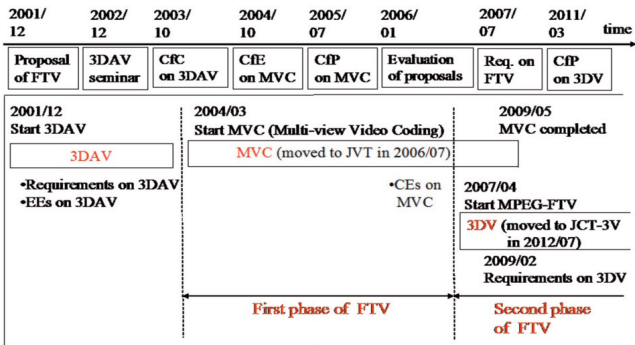


그림 2. MPEG FTV 그룹의 표준화 활동 역사

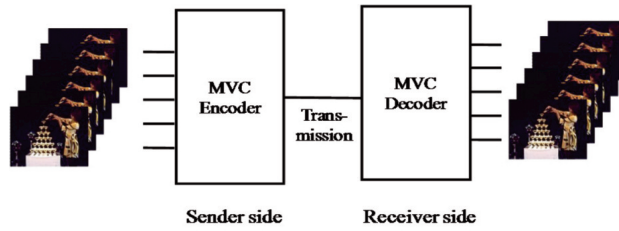


그림 3. 다시점 비디오 부호화 프레임워크

2단계 활동에서는 3차원 비디오 부호화(3D video coding, 3DVC) 기술을 개발하였다. 3차원 비디오 부호화 기술 개발은 다시점 비디오 부호화 표준화 활동이 종료되기 전인 2007년부터 시작되었다. 3차원 비디오 부호화는 다시점 비디오 부호화의 확장인데 가장 큰 차이는 깊이 비디오 부호화의 추가이다. 색상 비디오에서의 객체가 카메라로부터 가까울수록 깊이가 비디오에서는 밝은 값이 생성되고 멀수록 어두운 값이 생성된다.

원본 데이터에서 적은 숫자의 색상 비디오와 깊이 비디오를 부호화하여 전송을 하면 복호기에서 영상 합성을 통해 많은 수의 시점 데이터를 생성한다. 영상 합성은 깊이 정보가 존재하기 때문에 가능하다. 그림 4는 3차원 비디오 부호화의 프레임워크를 나타낸다. MPEG과 ITU-T 산하 VCEG(Video Coding Experts Group)은 함께 JCT3V(Joint Collaborative Team on 3D Video Coding Extension Development) 그룹을 만들어 3차원 비디오 부호화 기술 표준화를 진행하고 있다.

3차원 비디오 부호화에서는 선형적인 카메라 구조에서 수평 방향으로만 영상 합성이 가능하다. 따라서 30시점 이하의 디스플레이를 고려했을 때 큰 문제가 되지는 않지만 앞으로 개발될 차세대 디스플레이에 적용하기에는 3차원 비디오 부호화 기술들이 한계가 있다. 이러한 제약들을 해결할 수 있도록 기존 기술의 확장 및 신규 기술들이 개발되어 초다시점 비디오와 자유시점 비디오에 쓰일 수 있도록 3단계 활동을 2014년부터 진행하고 있다 [5-7].

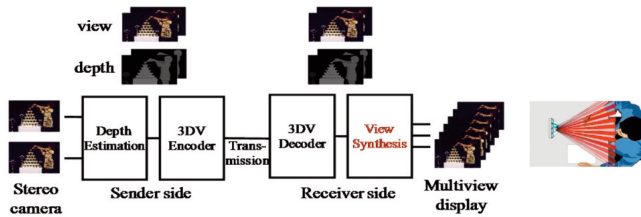


그림 4. 3차원 비디오 부호화 프레임워크

## 3. 실험 시퀀스

기술 평가 및 표준화 활동을 위해서는 실험 시퀀스가 필요한데 초다시점 비디오와 자유시점 비디오 실험에 이용될 수 있는데 데이터를 만들 수 있는 기관은 아직 세계적으로 많지 않다. 제작되는 시퀀스 또한 예기치 못한 에러가 계속 발생하기 때문에 MPEG 회의에서 시퀀스를 논의하고 오류 보완 등을 하는데 여러 미팅 사이클 등의 시간이 소요된다. 아직도 표준화 평가에 이용될 실험 시퀀스가 완전히 확정된 것은 아니지만 대략적인 윤곽은 잡혔고 이 장에서는 이러한 실험 시퀀스들에 대해 소개한다. 표 1은 2015년 5월까지 공개된 실험 시퀀스들이다.

표 1. 초다시점 비디오와 자유시점 비디오 실험 시퀀

| 순번 | 제작기관               | 시퀀스 이름            | 시점 수  | 해상도       |
|----|--------------------|-------------------|-------|-----------|
| 1  | Nagoya University  | Champagne Tower   | 80    | 1280×960  |
| 2  |                    | Pantomime         |       |           |
| 3  |                    | Dog               |       |           |
| 4  | Holografika        | Rabbit            | 91    | 1280×768  |
| 5  | Blender,           | Killing           |       |           |
| 6  | Tampere University | Flowers           |       |           |
| 7  | NICT               | Kimono            | 64    | 1920×1080 |
| 8  | Hasselt University | Soccer-linear 1   | 8     | 1600×1200 |
| 9  |                    | Soccer-linear 2   |       |           |
| 10 |                    | Soccer-corner     | 7     | 1920×1080 |
| 11 | Poznan University  | Poznan Blocks     | 10    | 1920×1080 |
| 12 |                    | Poznan Team       |       |           |
| 13 | Nagoya University  | Soccer-Game Goal  | 10    | 1920×1080 |
| 14 |                    | Soccer Game Side  |       |           |
| 15 | Hasselt University | San Migue         | 200   | 1920×1080 |
| 16 | NICT               | Bee               | 185   | 1920×1080 |
| 17 |                    | Shark             |       |           |
| 18 | Poznan University  | Poznan Laboratory | 45×45 | 1920×1080 |
| 19 | Nagoya University  | Akko & Kayo       | 15    | 640×480   |

표 1에서 8번부터 14번까지는 자유시점 비디오 시퀀스이고 나머지는 초다시점 비디오 시퀀스이다. 해상도는 대부분이 Full HD급(1920×1080)이다. 자유시점 비디오 시퀀스들은 전부 ground truth 깊이 맵이 제공되지 않고 초다시점 비디오는 제공되는 것도 있고 제공되지 않는 시퀀스도 있다. 그림 5는 일본 Nagoya 대학의 카메라 시스템 구조이다. 카메라의 수는 80개이며 5cm 간격으로 평행하게 배열되어 있다 [8, 9].

그림 6은 이 시스템을 이용하여 촬영한 “Champagne Tower” 색상 영상과 깊이 추정을 통하여 생성한 깊이영상을 나타낸다. 깊이영상은 FTV 그룹에서 기존에 개발한 깊이 추정 소프트웨어(depth estimation reference software, DERS)를 이용하여 생성된 데이터이다. 깊이영상은 3개의 시점에 한해서만 제공되었다. 그림 7은 벨기에 Hasselt 대학이 제작한 “San Miguel” 시퀀스이다 [10-12]. 이 시퀀스는 200개 시점 영상이 모두 광선 추적(ray tracing) 방식으로 생성되었다. 이를 이용한 깊이 영상은 ground truth 데이터가 된다. “San Miguel”과 “Champagne Tower”의 깊이영상을 비교해보면 깊이 추정은 개선 여지가 많은 것을 확인할 수 있다.

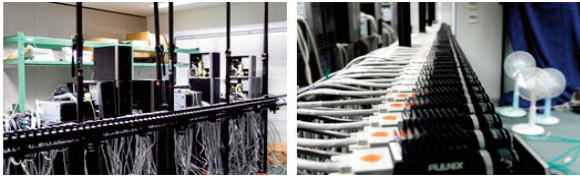


그림 5. Nagoya 대학 80-시점 카메라 시스템



그림 6. “Champagne Tower” 시퀀스 색상영상(좌), 깊이영상(우)

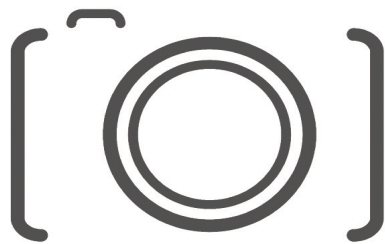






그림 7. “San Miguel” 시퀀스 색상영상(좌), 깊이영상(우)

#### 4. 기술 실험 및 평가 일정

##### 4.1 초다시점 비디오 기술 실험

초다시점 비디오 기술은 대체적으로 모든 시점을 전부 부호화하는 대신에 일부 시점만 부호화하여 전송하고 복호기에서 원본 시점 데이터들을 복호화한다. 시점의 수를 적게 전송할수록 데이터적으로는 이득이지만 결과로 복원되는 영상의 화질은 저하된다. 따라서 최적의 시점 수 전송은 많은 실험과 결과로 논의가 필요하다.그림 8은 초다시점 비디오의 기본적인 프레임워크를 나타낸다 [13].

1D 선형 또는 수렴형 형태로 넓은 범위를 촬영할 수 있는 카메라 시스템을 가정한다.

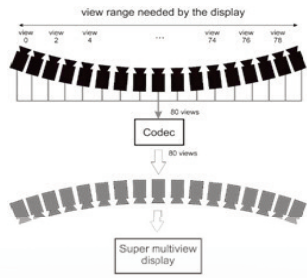


그림 8. 초다시점 비디오 프레임워크

현재 최고의 압축 성능을 내는 3차원 영상 부호기는 HEVC(High Efficiency Video Coding)을 기반으로 확장한 3D-HEVC이다. 3D-HEVC는 같은 수의 색상 비디오와 깊이 비디오를 부호화하여 데이터를 전송한다. 이 부호기에서는 시점 간 연관성,색상영상과 깊이영상의 유사성 등이 참조 정보로 이용된다.제안될 초다시점 비디오 기술은 3D-HEVC를 이용하여 부호화했을 때의 결과와 비교된다. 3D-HEVC를 이용할 때의 결과는 미리 약속된 환경 설정을 바탕으로 해야 한다. 대표적으로 P-I-P 시점 간 예측 구조, 8개의 GOP(Group of Pictures), 24 프레임마다 인트라 지정 등의 설정이다. 기술을 제안하는 기관은 rate-point 비트스트림,복호기 실행 파일,결과 비디오 클립,복호화 결과 영상의 PSNR(Peak Signal-to-Noise Ratio) 등을 제출해야 한다.

##### 4.2 자유시점 비디오 기술 실험

자유시점 비디오 기술은 넓은 간격으로 비선형적으로 배치된 10개 정도의 카메라 시스템에서 임의시점 영상 합성을 하기 위해 필요하다. 영상 합성에 깊이 영상과 같은 보조 데이터도 함께 전송된다. 복호화가 마무리되면 렌더러로 연결되어 임의시점 영상이 합성되고 디스플레이에 입력되게 된다.

다. 그림 9는 이러한 자유시점 비디오 프레임워크를 표현한다. 카메라는 간격이 넓게 떨어지고 시점 간 연관성이 낮은 상태를 가정한다. 예를 들면 축구장 같은 스포츠 경기장에 카메라를 둘레로 설치하면 TV 시청자는 원하는 시점에 카메라가 존재하지 않더라도 그 시점의 합성영상으로 스포츠를 즐길 수 있게 된다.

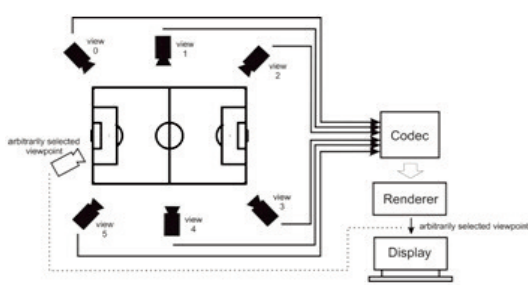


그림 9. 자유시점 비디오 프레임워크

현재 사용되는 대부분의 영상 합성 기술들은 깊이영상을 이용하여 3차원 정보를 생성한다. 자유시점 비디오에서는 시점 간 연관성이 매우 낮고 높은 수준의 합성 화질이 필요하기 때문에 깊이영상 기반의 합성 방법은 비효율적일 수 있다. 자유시점 비디오를 위하여 영상 포맷을 변경하게 된다면 부호화가 HEVC 기반일 때를 기준으로 하여 비디오 압축 성능 결과 또한 검토되어야 한다.

실험 결과 비교 대상은 3D-HEVC를 부호기로 사용하며, 영상 합성 참조 소프트웨어(view synthesis reference software, VSRS)를 이용하여 합성된 데이터이다. 자유시점 비디오 기술을 실험할 때 이용되는 3D-HEVC의 환경 설정은 초다시점 비디오 기술 실험에 활용되는 3D-HEVC 환경 설정과 동일하다. 제안하는 기술은 3차원 공간에 있는 어느 시점이든 합성 영상을 생성할 수 있어야 한다. 기술을 제안하는 기관은 rate-point 비트스트림,복호기 실행 파일,렌더러 실행 파일, 결과 비디오 클립, 복호화 결과 영상의 PSNR 등을 제출해야 한다.

##### 4.3 표준화 일정 및 기술 평가 방법

2015년 6월 미팅에서 최종 실험 시퀀스와 3D-HEVC 실험 평가 기준 데이터가 결정될 예정이며 2015년 10월 미팅에는 FTV 참여 기관들이 기술들을 제안하고 이를 평가하게 된다. 객관적 평가로는 rate-point, PSNR 등을 비교하게 된다. 초다시점 비디오 기술과 자유시점 비디오 기술은 목적이 서로 다르기 때문에 평가 또한 구분되어 이루어진다.

초다시점 비디오 기술의 평가 경우 양안식 모니터가 준비되며 복호화된 영상들의 비디오 클립들이 합쳐져 전송된 모든 시점 영상들에 대하여 sweep을 생성하게 된다. Sweep은 같은 시간 축을 바탕으로 시점들의 영상을 연결한 시퀀스이다. 스테레오 영상 간 간격은 각 시퀀스마다 따로 미리 결정된다. Sweep의 연결 속도는 1 시점 당 2 프레임이다.그림 10은 초다시점 비디오 기술을 평가하기 위한 sweep 형태를 나타낸다.

자유시점 비디오 기술 평가는 2D 모니터를 이용할 예정이다.전송된 데이터를 바탕으로 생성된 임의시점 영상들은 합쳐져 sweep을 생성하게 한다. Sweep의 시작점은 평가 좌상이 임의로 정한다.Sweep의 연결 속도는 1시점 당 1 프레임이다.그림 11은 자유시점 비디오 기술 평가 sweep 형태를 표현한다.

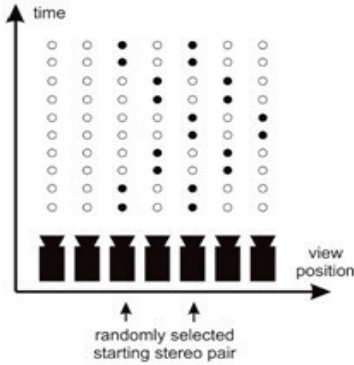


그림 10. 초다시점 비디오 기술 평가

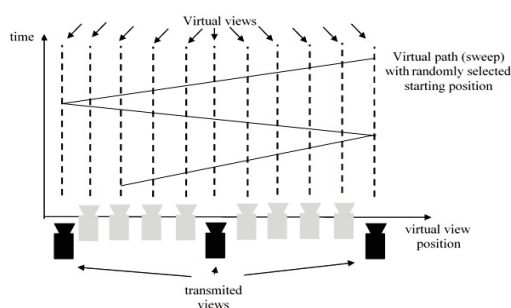


그림 11. 자유시점 비디오 기술 평가

#### 5. 결론

본고에서는 MPEG FTV 그룹의 초다시점 및 자유시점 비디오 표준화 동향에 대해살펴보았다. 현재까지 개발된 다시점 및 3차원 비디오 기술들은 미래의 다양한 응용 제품에 적용하기에 한계가 많기 때문에 새로운 기술들을 개발하고 이를 표준화하는 작업이 매우 중요하다. 초다시점 비디오는 100개 가까이 되는 무수히 많은 시점의 데이터를 효과적으로 처리해야 하며 자유시점 비디오는 시점 간 이동이 많은 영상에서 임의시점 영상의 합성 화질 증대가 중요한 문제이다. FTV 그룹에서는 부호화 및 임의시점 영상 합성이 상대적으로 용이한 좁은 카메라 간격 및 평행 배치 구조 틀에서 벗어나 다양한 카메라 구조에서 촬영된 영상에 적용될 수 있는 기술들을 개발할 예정이다. 2015년 6월과 10월에 열릴MPEG 회의에서는 기술 평가에 사용될 실험 영상,평가 방법등을 최종 결정할 예정이다.



호요성 교수  
소속 : 광주과학기술원 실감방송연구센터  
주 연구분야 : 영상신호처리, 영상압축, 3DTV, MPEG/HEVC  
E-mail : hoyo@gist.ac.kr  
http://vclab.kjist.ac.kr/

#### 참고문헌

[1] M. Tanimoto, M.P. Tehrani, T. Fujii, and T. Yendo, “Free-Viewpoint TV,” IEEE Signal Processing Magazine, vol. 28, no. 1, pp. 67–76, Jan. 2011.

[2] M. Tanimoto, T. Senoh, S. Naito, S. Shimizu, H. Horimai, M. Domanski, A. Vetro, M. Preda, and K. Muller, “Proposal on a New Activity for the Third Phase of FTV,” ITU-T/ISO/IEC JCT3V-E0258, July 2013.

[3] Y.S. Ho and Y. Song, “MPEG Activities for 3D Video Coding,” Proc. of Asia-Pacific Signal and Information Processing Association (APSIPA), Plenary 2.3(1–4), Dec. 2014.

[4] Y.S. Ho and Y. Song, “Overview of the State-of-the-art 3D Video Coding Technologies,” Global 3D Tech Forum, vol. 3, pp. 28–29, Oct. 2014.

[5]Requirements Group, “Experimental Framework for FTV,” ISO/IEC MPEG N15048, Oct. 2014.

[6] K. Wegner, T. Senoh, and G. Lafruit, “Description of Exploration Experiments on Free-viewpoint Television (FTV),” ISO/IEC MPEG N14551, July 2014.

[7] M. Tanimoto, K. Wegner, G. Lafruit, “AHG on FTV (Free-viewpoint Television),” ISO/IEC MPEG M35442, Feb. 2015.

[8] M. Tanimoto, T. Fujii, and N. Fukushima, “1D Parallel Test Sequences for MPEG-FTV,” ISO/IEC MPEG M15378, April 2008.

[9] M. Tanimoto and T. Fujii, “FTV-Free Viewpoint Television,” ISO/IEC MPEG M8595, July 2002.

[10] L. Jorissen, P. Goorts, N. Michiels, M. Dumont, B. Bex, S. Rogmans, and G. Lafruit, “EE1 and EE2 on San Miguel,” ISO/IEC MPEG M33164, April 2014.

[11] P. Goorts, M. Javadi, S. Rogmans, and G. Lafruit, “San Miguel Test Images with Depth Ground Truth,” ISO/IEC MPEG M33163, April 2014.

[12] P. Goorts, P. Carballeira, S. Garcia, K. Wegner, F. Moran, and G. Lafruit, “[FTV-AHG] EE2: San Miguel Results,” ISO/IEC MPEG M35111, Oct. 2014.

[13] Requirements Group, “Draft Call for Evidence on FTV,” ISO/IEC MPEG N15095, Feb. 2015.