



FTV 3차원 부호화를 위한 표준화 기술 동향

호요성 광주과학기술원 교수

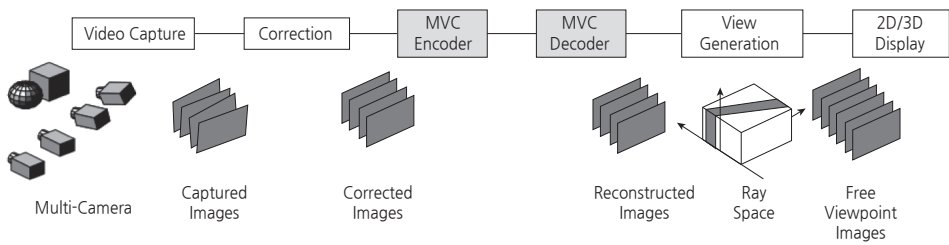


1. 머리말

최근 3차원 영화나 광고가 제작되면서 3차원 기술에 대한 관심이 증가하고 있다. 3차원 기술은 기존의 2차원 비디오에 입체감을 더한 정보를 의미한다. 사람은 좌우 눈을 통해 정보를 받아들일 때 존재하는 약간의 시차를 통해 물체의 입체감을 느낄 수 있다. 이러한 원리를 기반으로 3차원으로 제작된 영상들을 감상할 때 입체안경을 통해 좌우 눈에 각각 다른 영상이 보이게 하는 것이다. ‘Avatar’, ‘Gravity’와 같은 3차원 영화들의 흥행을 기반으로 소비자 시장에서 3차원 기술의 인지도는 높아졌다. 이에 따라 삼성과 LG는 스마트 TV에 3차원 기술을 추가하기 위해 개발을 하고 있으며, 머리에 착용하고 3차원 영상을 즐길 수 있는 Oculus Rift, Microsoft HoloLens, Gear VR 등과 같은 장치도 개발되고 있다. 3차원 비디오 기술은 앞으로 더욱 다양한 분야에서 적용 가능하고 동시에 3차원 기술의 수요 또한

커질 것으로 전망하고 있다.

3차원 영상은 2차원 영상에 깊이 정보를 추가하기 때문에 깊이 영상을 필요로 한다. 또한, 3차원 비디오는 2차원과 같이 단일 시점에서 촬영한 정보뿐만 아니라, 다시점에서 촬영한 정보를 사용해야 한다. 하지만 촬영할 수 있는 시점의 수는 제한되어 있기 때문에 촬영한 시점의 영상 데이터를 가지고 중간시점 영상 데이터를 생성하는 기술이 필수적이다[1]. 이를 가상시점 영상 합성이라고 하며, 이 과정에서 생성하고자 하는 가상시점과 인접한 시점의 색상영상과 깊이영상을 기반으로 3차원 워핑과 보간 방법이 사용된다. 일반적으로 3차원 영상은 단일 시점 카메라를 통해 촬영한 영상을 기반으로 생성되기 때문에 시청자에게 제한적인 시점 영상만을 제공해야 하는 단점이 존재한다. 제한된 시점 문제는 아직까지 우리가 실생활에서 볼 수 있는 것처럼



[그림 1] 자유시점 영상 시스템의 기본 구조

사물을 완벽하게 볼 수 없다는 한계를 가지고 있다.

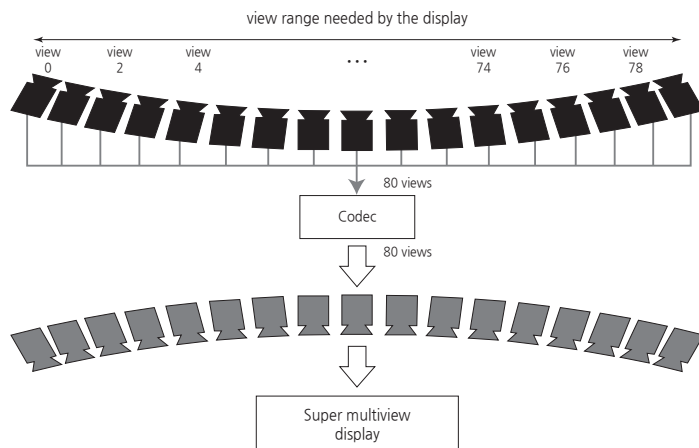
시청자들에게 자유로운 시점 영상을 제공하기 위해 MPEG에서는 3차원 비디오 부호화 (3D video coding, 3DVC)와 자유시점 영상(Free-viewpoint Television, FTV) 기술에 관한 연구를 진행하고 있다 [2][3]. 자유시점 영상에 관한 표준화 작업은 2001년 MPEG 미팅 이후로 지속적으로 진행되고 있다. 자유시점 영상의 첫 번째 형태로는 다시점 비디오 부호화(MVC, Multi-view video coding)로 진행되었으며, 2009년에는 3차원 다시점 비디오 부호화(3DMVC, 3D multi-view video coding)로 발전되었다[4][5]. [그림 1]은 기본적인 자유시점 비디오 시스템 구조를 나타내고 있다. 송신부에서는 다시점 카메라를 이용하여 3차원 장면을 촬영하며, 다시점 카메라로부터 생성된 많은 데이터는 전송과 저장을 위해 압축된다. 수신부에서는 부호화된 영상들을 복호기를 통해 보간 방법을 사용하여 복원하게 되는데, 이렇게 생성된 다양한 시점 영상들을 자유시점 영상이라 한다. 자유시점 영상은 PC를 이용한 모니터나 태블릿과 같은 장치들을 통해 시청할 수 있다. 다시점 카메라를 통해 촬영한 영상 데이터를 전송할 때 대역폭 초과 문제와 전송시간 지연 문제를 일으키므로 효율적인 부호화 방법이 필요하다. 이에 따라 MPEG(Moving Picture Experts Group) 국제 표준화 기구에서는 자유시점 영상을 위한 다

양한 부호화 방법에 관한 내용도 논의하고 있다. 자유시점 영상을 위한 기술로는 크게 초다시점 영상(SMV, Super multi-view)과 자유탐색 영상(FN, Free-navigation) 등이 있다.

본고에서는 FTV 그룹에서 진행한 자유시점 영상 제공을 위한 3차원 비디오 부호화 기술 동향을 소개한다. 세부적으로 초다시점 영상을 위한 부호화 기술과 자유탐색 영상을 위한 합성을 설명한다. 또한, 표준화 작업에서 사용되는 실험 영상들과 테스트를 위한 환경 설정에 대해 설명한다.

2. 자유시점 영상을 위한 응용 프로그램

초다시점 영상과 자유탐색 영상을 기술은 시청자에게 자유로운 시점의 영상을 제공하는 것을 목적으로 한다. 초다시점 영상과 자유탐색 영상은 이와 같은 공통적인 부분이 있는 반면, 각 시스템이 추구하는 구체적인 목표는 다르다. 초다시점 영상은 다시점 카메라로 촬영한 영상의 정보를 효율적으로 부호화하는 것을 목표로 하고 있으며, 자유탐색 영상은 넓은 시점의 영상을 이용하여 곡선 영역이나 카메라의 기하학적 구조가 변형되는 구간에서 효율적으로 가상시점 영상을 생성하는 것을 목표로 하고 있다. 두 가지 방법 모두 자유로운 시점 영상을 제공하는 최종적인 목적은 같지만, 서로 다른 장점을



[그림 2] 초다시점 영상 시스템의 구조

가지고 있다. 본 장에서 각 방법에 대해 구체적으로 설명하도록 한다.

2.1 초다시점 영상(SMV, Super multi-view)

다시점 카메라를 이용해 촬영한 영상을 기반으로 시청자에게 다양한 시점의 영상을 제공할 수 있지만, 카메라 설치를 위한 물리적 제약과 영상 정보 전송을 위해 필요한 대역폭, 전송 시간 등의 기술적 한계가 존재한다. 초다시점 영상 시스템은 이러한 문제를 해결하기 위해 영상 정보를 효율적으로 부호화하여 전송하는 것을 목적으로 한다. 효율적으로 영상 정보를 부호화하기 위해 영상 정보를 전송할 때 일부 영상만을 압축해서 전송하며, 수신단은 받아들인 정보를 기반으로 다른 시점의 영상을 복원하여 새로운 시점의 영상을 생성한다. 부호화된 정보를 복호화 과정을 통해 다시점 카메라로 촬영한 영상과 유사한 수의 가상 시점 영상을 생성하기 위해서는 송신단에서 가능한 많은 양의 초기 영상 정보를 보내야 한다. 이러한 이유 때문에, 현재 초다시점 영상 생성 방법에 관한 주된 연구 내용은 높은 부호화 효율을 갖는 부호화기를 개발하고 이를 통해

촬영한 영상 정보의 손실을 최소화하여 전송하는 것을 목표로 하고 있다.

초다시점 영상을 위한 다시점 카메라는 각 카메라 가격이 정렬되어 있으며 수평 시차가 일치하는 밀집된 구조를 가지고 있다. 이러한 구조적 특성 덕분에 시청자는 3차원 안경의 도움 없이 자유롭게 시점을 선택할 수 있으며, 실제 사물의 앞을 걸어 다니면서 감상하는 것 같은 ‘walk-around feeling’을 체험할 수 있다.

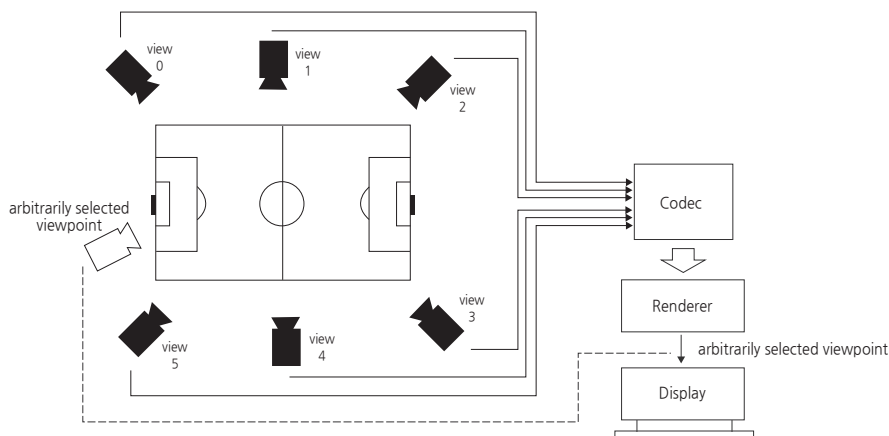
초다시점 영상은 일반적으로 80시점 이상의 다시점 카메라를 이용하여 촬영한 영상 정보를 의미한다. FTV는 80시점의 카메라를 이용하여 촬영한 영상을 통해 자유로운 시점의 영상을 제공하는 것을 초다시점 영상으로 정의하고 있다. [그림 2]는 초다시점 영상에 사용되는 카메라 시스템을 나타내며, 카메라들은 선형적 또는 곡선형의 1차원 배열 구조를 사용한다.

2.2 자유탐색 영상(FN, Free-navigation)

가상시점 영상을 만들기 위해서는 이웃한 위치에서 촬영한 색상 영상과 이에 부합하는 깊이 영상



[그림 3] 자유탐색 영상 시스템의 생성 과정



[그림 4] 자유탐색 영상 시스템 구조

이 필요하다. 하지만 스포츠 중계나 광범위한 영역에서 촬영이 필요한 경우에는 초다시점 카메라 시스템을 구축하는데 많은 제약이 존재한다. 자유탐색 영상 시스템은 이러한 문제를 극복하고자 고안된 시스템으로 [그림 3]에 나타나 있다. 넓은 영역에서 촬영한 영상을 부호화 과정을 거쳐 전송하게 되면, 수신단의 복호기를 통해 생성된 가상시점 영상을 사용하여 시청자에게 자유로운 시점의 영상을 제공할 수 있게 된다. 자유탐색 영상은 서로 다른 카메라 장비를 이용하거나 다른 경로에 설치된 카메라를 통해 촬영된 영상을 통해 생성된 가상 시점의 영상을 시청자가 아무런 제약 없이 선택하여 시청하는 것을 의미한다. 자유탐색 영상을 시청하기 위해 기존 2차원 영상, 입체영상, 자동 입체영상 그리고 통합 촬영 시스템에 적용했던 디스플레이 장치

를 사용한다.

자유탐색 영상은 넓은 촬영범위를 기반으로 곡선 영역이나 가려진 영역에서 촬영된 것 같이 다양한 가상시점 영상을 생성하는 것이 최종적인 목표이다. 이를 위해 깊이 영상을 이용하여 실제 촬영한 영상과 인접한 영역에서의 가상시점 영상을 생성하는 방법에 대한 소프트웨어 개발도 진행되고 있다[6]. 자유탐색 영상을 위한 카메라 시스템 구조는 [그림 4]에 나타나 있다. 자유탐색 영상 시스템은 카메라 사이의 간격을 넓게 설치할 수 있도록 6~10개 정도로 적은 수의 카메라를 사용한다. 카메라 사이의 간격은 baseline이라고 하는데 자유탐색 영상 카메라 시스템은 넓은 baseline을 이용해 촬영을 수행한다. 적은 수의 영상을 기반으로 가상시점 영상을 생성하여 시점의 자유도를 높여 시청자에게 제공하는

<표 1> 초다시점 영상 시스템을 위한 실험 영상

순서	제작기관	이름	시점	해상도	fps	재생시간	카메라 구조
1	Nagoya 대학	Champagne Tower	80	1280x960	29.4114	10초 300 프레임	linear 1D parallel
2	Nagoya 대학	Pantomime	80	1280x960	29.4114	10초 300 프레임	linear 1D parallel
3	Holografika	Big Buck Bunny flowers	91	1280x768	24	8초 192 프레임	45 degree arc
4	Holografika	Big Buck Bunny butterfly	91	1280x768	24	8초 123 프레임	45 degree arc

<표 2> 자유탐색 영상 시스템을 위한 실험 영상

순서	제작기관	이름	시점	해상도	fps	재생시간	카메라 구조
1	Hasselt 대학	Soccer-linear2	8	1600x1200	25	22초 550 프레임	1D parallel (no aliened)
2	Hasselt 대학	Soccer-corner	7	1280x1080	25	22초 550 프레임	120 deg Corner, arc
3	Poznan 대학	Poznan Blocks	10	1280x1080	25	40초 1000 프레임	100 deg, arc around the scene
4	Holografika	Big Buck Bunny flowers no_blur	91	1280x1080	24	8초 192 프레임	45 degree arc

것이 자유 탐색 영상 시스템의 목적이다.

3. 성능 검증 위한 실험 영상

초다시점 영상 시스템과 자유탐색 영상 시스템을 이용하여 생성한 가상시점 영상의 화질 및 시스템 성능을 검증하기 위해 FTV 그룹에서는 지정된 실험 영상을 사용한다.

〈표 1〉은 초다시점 영상 시스템 성능 검증에 사용되는 실험 영상들의 자세한 특성을 나타내고 있다. 실험 영상은 카메라로 직접 촬영하거나 컴퓨터 그래픽을 통해 만들어지며, 각 실험 영상마다 서로 다른 카메라 구조를 사용한다. 초다시점 카메라 시스템을 위한 실험 영상과 자유탐색 카메라 시스템의 실험 영상은 두 시스템의 목적이 다르기 때문에 서로 다르게 지정되어 있다. 〈표 2〉는 자유탐색

영상 시스템을 위한 실험 영상들을 나타낸다.

초다시점 영상 시스템 검증을 위해 ‘Champagne Tower’는 0~79, ‘Pantomime’은 0~79, ‘Big Buck Bunny flowers’는 5~84 그리고 ‘Big Buck Bunny butterfly’는 5~84 시점의 영상을 사용한다. 또한, 자유탐색 영상 시스템에는 ‘Soccer-linear 2’는 1~7, ‘Soccer-corner’는 1~7, ‘Poznan Blocks’는 2~8 그리고 ‘Big Buck Bunny Flowers’는 6, 19, 32, 45, 58, 71, 84번 시점 영상을 사용한다. 자유탐색 영상은 지정한 시점의 영상을 사용하여 생성하는 이상적인 가상시점의 수를 제안한다. ‘Soccer-linear 2’는 14, ‘Soccer-corner’는 22, ‘Poznan Blocks’는 12 그리고 ‘Big Buck Bunny Flowers’는 12개의 가상시점 영상을 생성하여 시스템의 성능을 검증하도록 되어 있다[7].

실험 영상을 이용하여 가상시점 영상을 생성할

경우, 3D-HEVC 부호기를 사용하여 부호화된 정보를 전송하여 복호화된 정보를 통해 가상시점 영상을 생성한다. 가상 시점 영상 생성은 깊이 영상을 기반으로 하는 VSRS 소프트웨어를 이용하며 4.0 버전을 적용한다. 3D-HEVC를 이용하여 영상을 부호화할 경우 시점 간 부호화 모드를 사용하며 시간 축 예측 때 GOP(Group of Pictures) 크기는 8로 지정하며 intra period는 24프레임으로 설정한다. 부호화 입력 영상으로 8비트로 이루어진 데이터를 사용하며 합성 영상 최적화(VSO, View synthesis optimization)는 사용하지 않은 상태에서 작업을 수행한다.

4. 맺음말

본고에서는 FTV 3차원 비디오 부호 기술 동향에 대해 살펴보았다. 현재 FTV를 활성화하기 위해 중점적으로 연구되는 초다시점 영상과 자유탐색 영상에 대해 설명하였다. 각각의 시스템의 최종 목표는 자유로운 시점의 영상을 생성하는 것이지만, 각 시스템 서로 다른 장점을 이용하여 가상시점 영상을 생성한다. 시스템의 성능을 검사하기 위해 사용되는 다양한 실험 영상들의 종류와 특징들을 표를 통해 확인했다. 마지막으로 실험을 위해 사용되는 소프트웨어를 설명하고 부호화 과정에서 필요한 변수 값 설정 방법을 알아보았다. 초다시점 영상과 자유탐색 영상 시스템은 2018년 대한민국 평창에서 개최되는 동계올림픽과 2020년 일본 도쿄에서 열리는 하계올림픽에 적용하기 위해 많은 연구가 진행되고 있다. 자유시점 영상은 일반적인 스포츠 중계나 영화, 게임 등 여러 산업 분야에 적용하여 다양한 콘텐츠를 제작할 수 있는 발전 가능성을 가지고 있다. 이에 따라 국내외 학계 및 기업에서 지속적으로 자유

시점 영상 시스템에 대해 많은 관심을 가지고 관련된 산업 기반을 마련함으로써 앞으로 많은 인력이 이와 관련된 연구를 진행할 수 있는 환경을 제공해야 한다. 

[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, no. 6, Nov. 2007.
- [2] M. Tanimoto, 'Overview of Free Viewpoint Television,' Signal Processing: Image Communication, vol. 21, no. 6, pp. 454-461, July 2006.
- [3] M. Tanimoto, 'FTV (Free viewpoint TV) and Creation of Ray-Based Image Engineering,' ECTI Transaction on Electrical Engineering, Electronics and Communications, vol. 6, no. 1, pp. 3-14, Feb. 2008.
- [4] Video Group, 'Introduction to Multi-view Video Coding,' ISO/IEC JTC1/SC29/WG11, N7328, July 2005. <http://www.chiariglione.org/mpeg/technologies/mp-mv/index.htm>
- [5] Video Group, 'Introduction to 3D Video,' ISO/IEC JTC1/SC29/WG11, N9784, May 2008. <http://www.chiariglione.org/mpeg/technologies/mp3d/index.htm>
- [6] K. Wegner, O. Stankiewicz, M. Tanimoto, M. Domanski, 'Enhanced View Synthesis Reference Software (VSRS) for Free-viewpoint Television,' ISO/IEC JTC1/SC29/WG11 MPEG2013/M31520 Oct. 2013, Geneva, Switzerland.
- [7] G. Lafruit, K. Wegner, M. Tanimoto, 'Final Draft Call for Evidence on FTV,' ISO/IEC JTC1/SC29/WG11 MPEG2014/M36580 June 2015, Poland, Warsaw.