

자유시점 콘텐츠 구성기술 개발 사례

□ 박구만 / 서울과학기술대학교

1. 서론

자유시점 비디오 기술은 원하는 임의의 각도나 위치, 거리에서 이벤트를 시청할 수 있게 하는 기술이다. 다양한 뷰를 끊김 없이 부드럽게 연결하면서 볼 수 있고 임의 위치에서 공연이나 경기의 관찰이 가능해 실감성을 높여주는 장점이 있다. 이러한 서비스는 많은 수의 카메라들이 배열돼야 하고 영상 구성도 복잡하다. 시점을 바꾸는 과정에서 연속적으로 옆에 있는 카메라를 선택해가거나 다른 위치의 카메라로 자주 스위칭하기 때문에 여러 채널의 비디오가 동시에 전송돼야 한다. 일반적으로 카메라 선택이 자주 바뀌면 비디오 압축 효율도 떨어지게 되고, 초당 데이터 전송량도 더 많아지므로 대용량 콘텐츠를 전송할 수 있는 고속 전송 기술이 필요하다. 대용량, 초고속 5G 무선 전송기술이 서비스되면 스포츠 경기장에서 관객들

이 잘 보이지 않는 각도의 경기 내용을 휴대단말기를 통해서 쉽게 볼 수 있을 것으로 기대되며, 경기 이해와 몰입에 큰 도움이 될 것이다.

자유시점 비디오를 위해서는 카메라별 영상왜곡 보정, 객체 추출, 객체 추적, 객체 분리, 영상 렌더링, 입체음향기술, 대용량 전송기술 등이 필요하다. 또한 콘텐츠 시점을 부드럽게 바꾸고 선택적 객체의 표출, 크기 조절 등이 편리한 사용자 인터페이스가 뒤따라야 한다[1].

기존의 대표적인 자유시점 기술은 1999년에 나온 매트릭스 영화에서 적용했던 “bullet time” 장면을 들 수 있다. 총알이 발사되어 날아가는 장면을 여러 시점에서 보여줌으로써 신선한 충격을 주었다. 총알 실루엣과 여러 각도의 카메라 영상을 합성하여 자유시점 영상을 보여주었다. 2001년에 미식축구 중계에서 카네기멜론대학교와 CBS가 공동으로 적용한 EyeVision기술로 자유시점을 운동경기에 적

용하여 몰입감을 제공한 대표적인 사례이다. 이후 다양한 연구가 진행되어 오고 있다. 2016년 Replay Technologies사는 자유시점 재생기술인 ‘freeD’를 미식축구 중계에 적용하였으며 24개의 4K 카메라를 사용하였다. KDDI[1]는 10개 이내의 카메라를 사용하고, 선수와 배경을 분리하여 자유시점을 제공하는 기술을 개발하였다. 선수를 빌보드(bill-board)개념으로 추출하여 배경과 합성하여 시점이 바뀔 때 해당 각도의 영상을 제공한다. 언급된 것 외에도 2000년 이후 현재까지 다양한 서비스들이 국내외에서 개발되어 스포츠 장면 재생 등에 적용되고 있다.

한편, 개발의 관심이 피사체의 볼륨감을 제공하는 3D 모델링 쪽으로 확장되고 있다. 8i사[2]는 피사체를 모든 방향으로부터의 수십대 카메라들을 사용하여 영상을 획득하였다. 실내에서 촬영하며 피사체와 근접한 거리에서 촬영한다. 국내의 STANS사[3]는 10대 이내의 카메라와 제스처인식 기술을 활용하여 동적 3D모델을 구현하고 있다. 동적 3D 모델이 자유시점 동영상내에 포함되어 관심 인물의 특정 순간을 독립적으로 관찰할 수 있게 했다.

지금까지 자유시점 서비스는 카메라가 많이 필요하고 다채널 전송에 의한 대역폭 요구량이 많아 아직까지 활용이 제한적이었다. 대용량 전송이 가능한 5G기술이 일반화되면 자유시점 서비스도 활발해질 것으로 예상된다.

자유시점 서비스에 필요한 요소기술은 카메라 배열 및 동기화 구동 기술, 카메라 캘리브레이션(영상 보정), 객체 추출, 객체 추적, 객체 분리, 객체와 배경 합성 기술 등이다.

본 고에 소개할 서비스 예에서는 다양한 시점 선택을 위한 파라미터 정의, 영상과 음향의 연동, 부

드러운 카메라 전환, 스트리밍 기술, 콘텐츠 획득 및 표출 기술이 포함된다.

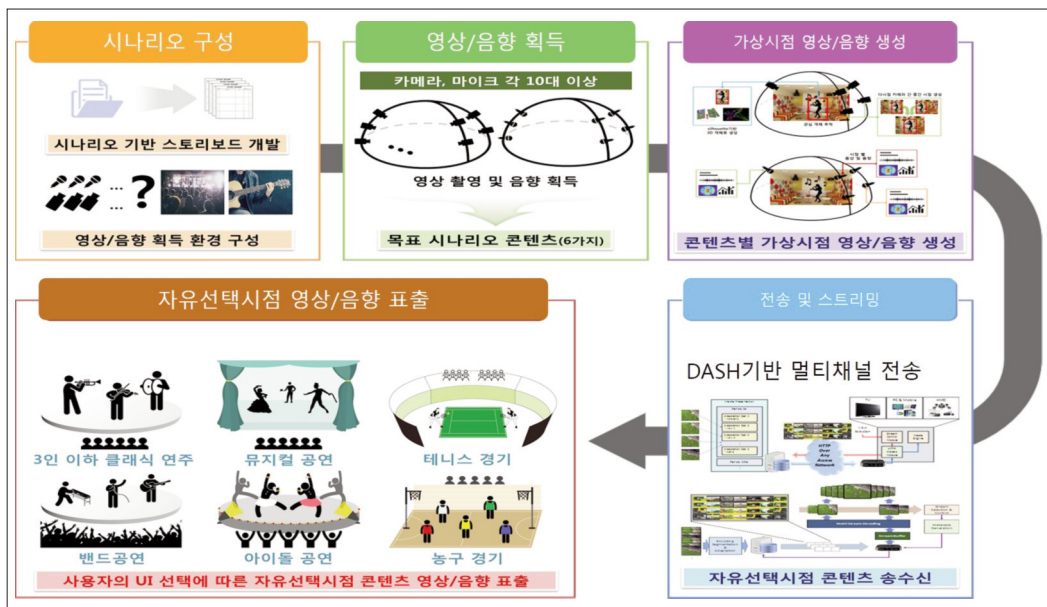
II. 자유시점 서비스 구성

1. 사용자 선택 뷰의 종류와 정의

본 절에서는 사용자와 상호작용이 가능하면서 자유롭게 공간상의 위치, 각도 및 인칭 시점을 선택할 수 있는 뷰의 종류와 파라미터 정의 예를 소개한다.

사용자가 선택할 수 있는 다양한 시점의 종류와 해당 파라미터를 다음과 같이 정의하였다. 여러 파라미터들이 결합되어 다양한 뷰를 표현해주는 것이 가능하다.

- 기존의 한 채널 고정 시점(Conventional single viewpoint)
- 내향시점 시간 고정 자유시점(inward time freezing free viewpoint, bullet time)
- 내향시점 시간 가변 자유시점(inward time varying free viewpoint)
- 전방위 자유위치시점(free position 360 degree spherical outward view)
- 반구형 카메라 배열에 의한 자유시점(hemispheric free viewpoint)
- 줌인/줌아웃 표출(zoom in/zoom out)
- 외향 시점 카메라 배열(outward camera arrangement)
- 동적 3D 모델 기반 객체 관찰시점(inward object view by dynamic 3D modeling)
- 1인칭시점(first person view)
- 기타: 객체 추가, 감소



〈그림 1〉 자유시점 콘텐츠 전송 시스템 개념도

다양한 시점에서 스포츠나 무대공연을 감상하기 위한 전체 시스템 구성을 〈그림 1〉에 나타냈다.

〈그림 2〉는 정의된 자유시점 파라미터를 적용하는 무대 공연의 예를 보여준다. 관객석에서 무대를

바라보는 시점이 가장 기본이고, 다수의 카메라를 배열해서 자유시점 뷰를 생성한다. 무대 위에는 공연자의 입장에서 주변 연주자들을 감상할 수 있는 자유위치시점이 생성되며, 외향 시점의 속성상 다



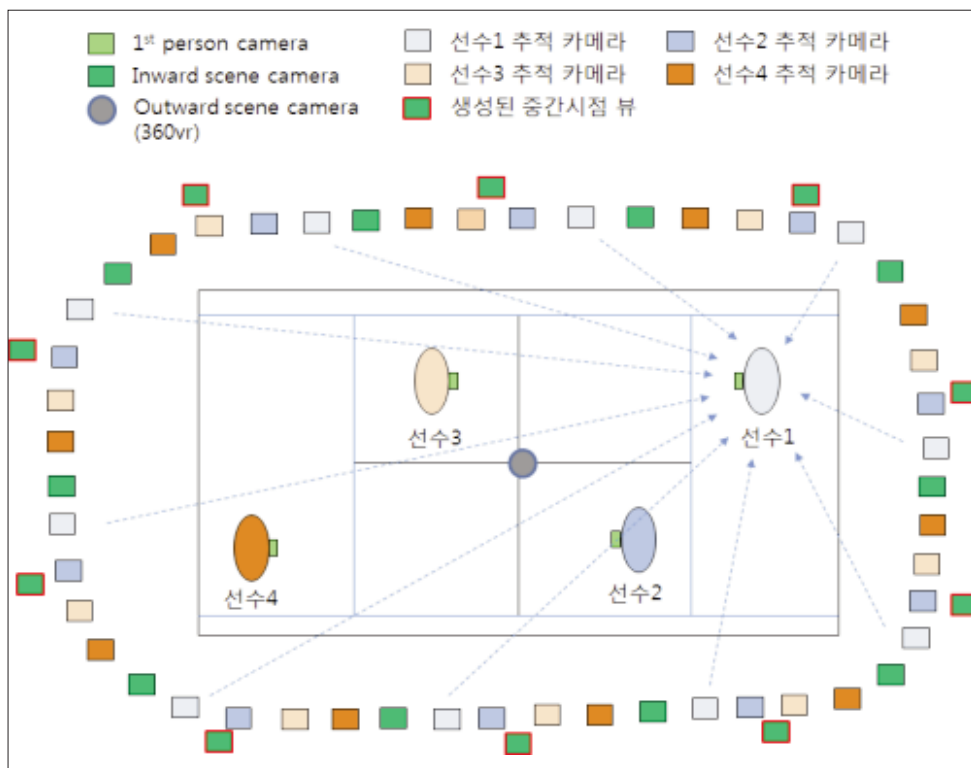
〈그림 2〉 1인칭 자유선택시점/자유위치시점 개념도

수의 카메라를 배열하기 보다 360도 전방위 카메라를 사용할 수 있다. 특히 관객이 무대에 직접 참여하여 주변과 하모니를 맞추는 참여형 1인칭 시점과 체험을 제공해준다. case 2에서는 관심이 있는 연주를 추가하거나 특정 연주를 제거할 수 있다.

〈그림 3〉은 테니스 경기에서 자유시점을 구현하기 위한 카메라 배열을 보여준다. 테니스장을 둘러싼 카메라들은 관객위치에서 경기를 보는 내향시점을 제공한다. 별도로 각 선수들을 전문적으로 추적하는 선수추적용 카메라가 배치된다. 각 선수들을 추적하는 용도의 카메라 영상은 선수들의 동적 3차원 모델을 생성하기 위한 움직임 정보를 추출하는데 사용한다. 선수는 필요시 웨어러블 카메라를 착

용하여 1인칭시점의 영상을 보여줄 수 있다. 자유시점 비디오 내에서 특정 선수의 동적 3D 모델이 생성되면 움직임의 상세한 관찰이 가능해진다. 테니스코트의 중심에 360도 전방위 카메라를 배치하여 자유위치시점의 개념을 표출할 수 있다. 예로 든 카메라 배열의 문제점은 맞은편에 배치한 카메라가 다른 카메라의 영상에 노출되는 점이다. 노출되지 않게 촬영하는 방법이 필요하다. 축구 경기에서는 카메라를 넓은 운동장에 배치하기 때문에 눈에 보이지 않게 할 수 있다.

각 선수별로 여러 방향에서 추적하는 카메라 영상을 활용하여 선수의 운동궤적을 관찰할 수 있을 뿐만 아니라 3D 모델링도 가능하며 객체 관찰 시점을 제공



〈그림 3〉 스포츠 중계영상에서의 자유선택시점 개념

한다. 동적 3D 객체 관찰시점은 기본적으로 정지 상태에서 인물의 3D 모델을 획득한다. 여러 방향에서 촬영한 영상을 이용해 동작 정보를 얻어낼 수 있다. 영상에서 관심객체의 골격 정보를 추출하여 3D 정지 모델에 적용하면 동적 3D 모델을 생성할 수 있다.

2. 카메라 배열과 보정

내향 시점 영상을 획득하는데 있어서 수평라인에 함께 배열된 카메라들의 높이를 동일하게 하고 가상 시점도 생성하기 위해 기하학적 영상보정을 해야 한다. <그림 4>에 카메라 수평 보정 전과 후의 개념을 나타내었으며 다음과 같이 수평 보정을 실행하였다.

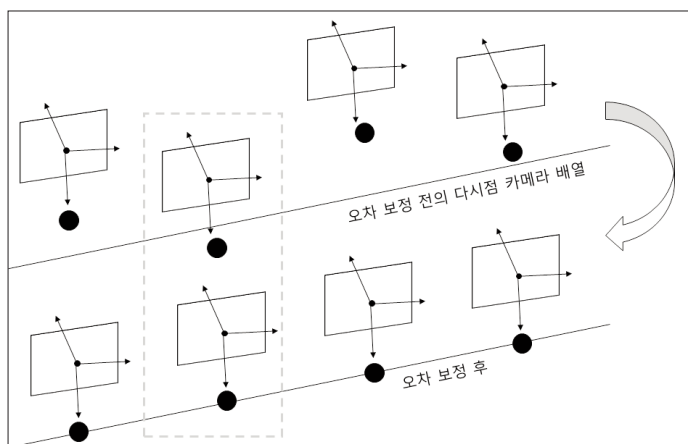
- ① 다시점 영상 중 기준 시점을 선택하여 이상적인 카메라 배열에서의 매개변수를 예측
- ② 예측된 카메라 매개변수와 원본 카메라 매개변수를 이용하여 보정 변환을 계산하고 영상에 적용
- ③ 기하 오차 보정 결과를 확인하기 위해 4번째 카메라에 물리적 오차를 두어 이미지를 획득

- ④ 세 번째 카메라를 기준 시점으로 선택하고, 이상적인 카메라 매개변수를 계산하여 원본 영상에 적용
- ⑤ 기하 오차 보정을 통해 4번째 시점 이미지가 다른 시점과 동일하게 이미지가 보정되었음을 확인

3. 중간시점 영상 생성

카메라 배열로 다양한 방향의 멀티 뷰를 생성할 수 있지만 카메라를 많이 사용해야하는 문제가 있다. 중간시점을 생성하면 카메라 개수를 충분히 줄일 수 있다. 카메라마다 깊이 영상을 구하는 과정은 다음과 같다.

- 1) 3차원 공간상의 원점 좌표나 카메라로부터 객체까지의 거리 축을 Z축이라 했을 때 깊이가 영상 및 가상 시점 생성을 위해 시차(disparity) 값과 z 값을 조정한다.
- 2) 시차 값과 z값의 조정은 다음과 같다.
 - ① 가상 시점이 성공적으로 생성되는 최적의

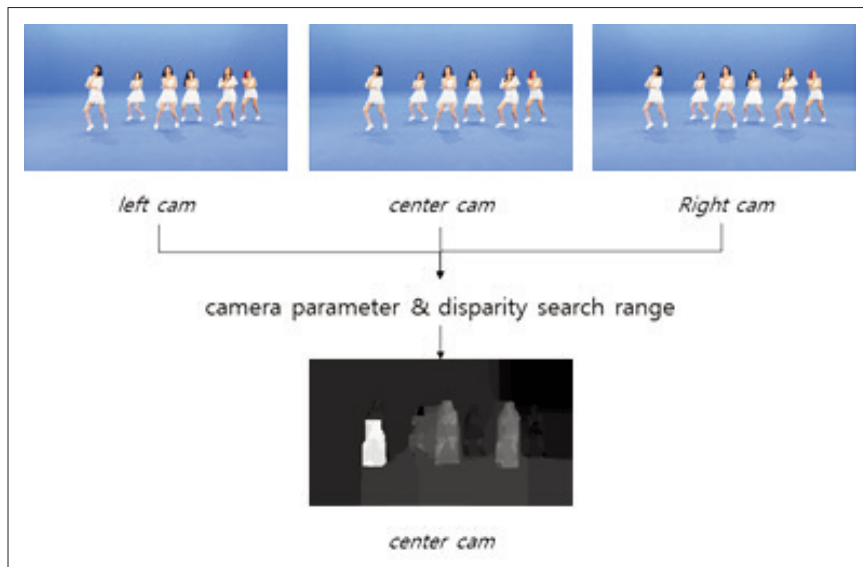


<그림 4> 기준 시점으로부터의 오차 보정

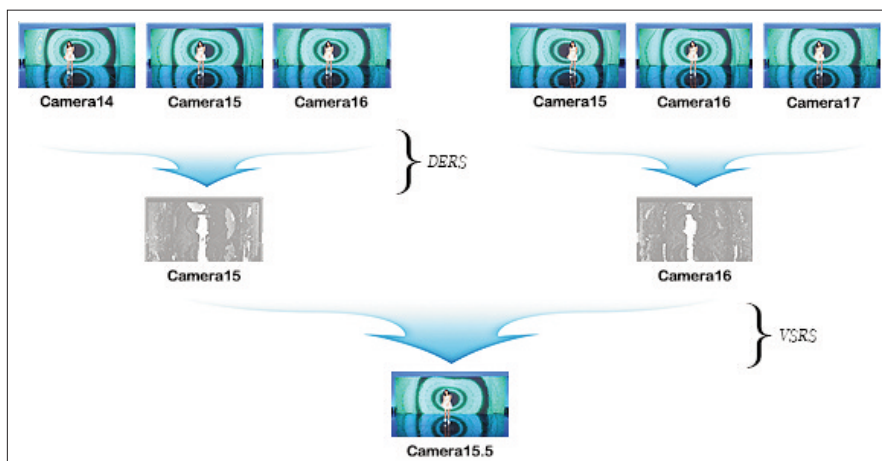
- 가장 최소 z 값을 찾음
- ② 깊이 맵이 성공적으로 생성되는 최적의 최대 z 값을 찾음
 - ③ z 값을 모든 시점에 대해 적용하기 위해 각 시점마다 해당 z 값을 도출하는 시차 값을 유도

〈그림 5〉는 인접하는 3개의 카메라로부터 깊이 맵을 생성한 예를 보여준다. 카메라 파라미터와 깊이 맵 및 색상 영상을 이용하여 각 카메라 사이의 중간 영상을 생성하였다.

〈그림 6〉은 DERS(Depth Estimation Reference Software)와 VSRS(View Synthesis Reference



〈그림 5〉 깊이 영상 생성 방법



〈그림 6〉 DERS와 VSRS를 이용한 깊이 영상 및 가상 시점 생성

〈표 1〉 가상 시점 영상 생성 결과

frame	cam15	cam15.5(virtual view)	cam16
696			

Software)를 이용하여 깊이 영상 및 가상시점을 생성한 예를 보여준다. 15번째 카메라와 16번째 카메라 사이의 영상이 생성된 예이다.

〈표 1〉은 가상시점 영상 생성 결과를 보여준다.

4. 동적 3D 모델 구성

특정 연주자나 선수를 자세히 관찰할 수 있도록 동적 3차원 객체 모델을 선택 시점의 하나로 정의하였다. 〈그림 7〉은 동적 3차원 객체 모델을 구성하기 위한 기본 정지 3D 모델이 생성된 예이다.

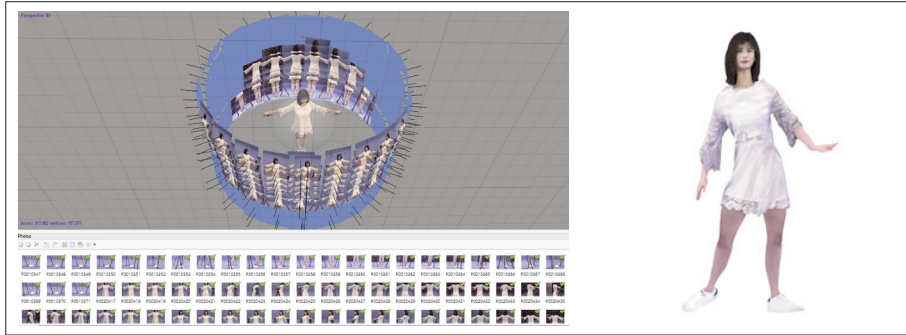
다시점 영상을 이용한 3차원 객체 구성을 위해 다음과 같은 과정을 거친다.

- ① 관심 객체에 대해 다시점으로 이미지 획득
- ② 입력된 다시점 영상을 정렬
- ③ cloud point들을 조밀하게 밀집시킴으로써 매쉬(mesh)를 생성
- ④ 매쉬에 원본 영상에 대한 텍스처를 입힘으로써 3차원 객체 구성

3차원 객체 복원 실험을 위해 총 240장의 다시점 영상을 적용하였다. 〈그림 8〉은 영상 촬영과 객체



〈그림 7〉 3차원으로 복원된 객체에 대한 실루엣

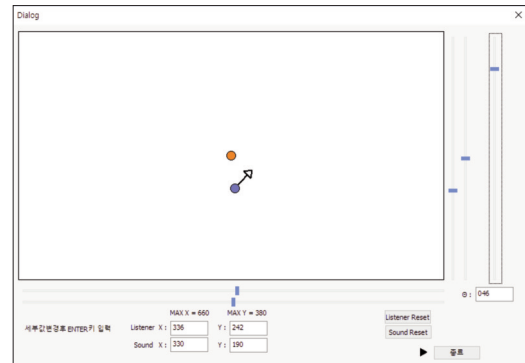


〈그림 8〉 동적 객체에 대한 3차원 복원

모델을 복원한 예를 보여준다. 3D 객체가 완성된 후 각 방향에서 촬영한 피사체의 움직임 인식 결과를 적용하여 동적 3D 모델을 완성하였다.

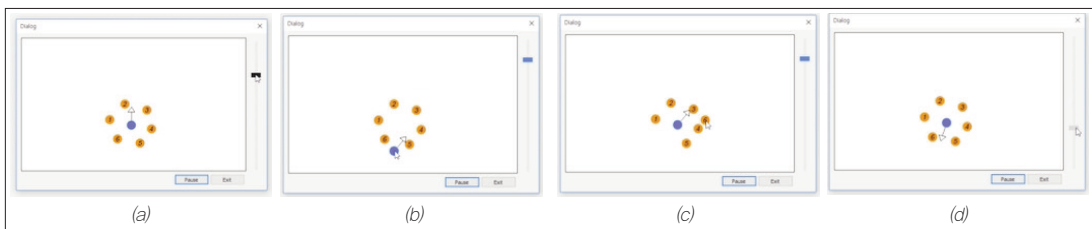
5. 자유선택시점과 연동되는 음향 표현

무대에서 공연자들마다 개별적으로 음원을 취득한다. 취득된 음원들을 감상자의 위치 및 시점 방향에 따라 믹싱하여 표출하는 것을 개념적으로 표시하면 〈그림 9〉와 같다. 아래 파란색 점은 감상자이고 화살표는 바라보는 방향이다. 위에 주황색 점은 음원이다. 음원과 감상자의 위치를 자유롭게 재배치하면 이 조건에 맞는 입체 공간 음향이 표출되도록 믹싱과 인터페이스가 구현된다.



〈그림 9〉 현장내향시점 음향 및 시각정보 표출

〈그림 10〉은 음원과 감상자의 위치를 자유롭게 바꾸어 입체음향을 구성하는 예를 보이고 있다. (a)는 6개의 연주자 음원이 있고 가운데 감상자를 보여주고 있고 화살표 방향을 보고 있다. 이 경우 위



〈그림 10〉 음원과 감상자의 위치를 자유롭게 바꾸어 입체음향을 구성하는 예

쪽 두 명의 음향이 주도적으로 표출된다. (b)는 감상자가 아래로 이동한 경우이고, (c)는 연주자와 음원이 모두 움직이는 예이다. (d)에서는 아랫방향의 연주자 음향이 강조되는 경우이다.

6. 서버기반 인터넷 스트리밍 서비스 기술

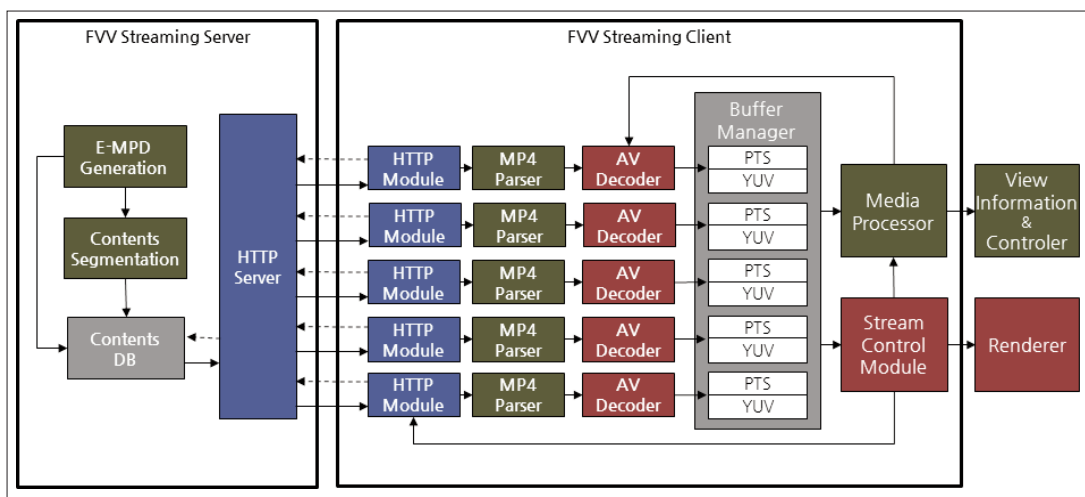
사용자의 시점 이동을 위한 다중 영상 송수신 시스템은 <그림 11>과 같다[4]. FVV 스트리밍 서버는 자유선택시점 콘텐츠를 다중 영상 전송이 가능하도록 부호화, 패킷구성(encapsulation), 세그멘테이션 과정을 거쳐 스트리밍 포맷으로 저장한다. 업로드된 각 비디오 스트림을 클라이언트가 수신할 수 있도록 E-MPD(Expended-Media Presentation Description)를 생성해서 저장한다. 저장된 E-MPD와 세그먼트된 영상들은 HTTP 서버 모듈을 통해 클라이언트에 전송된다.

FVV 스트리밍 클라이언트는 다중 영상 스트리밍 및 시점 이동을 위해 사용자가 선택한 시점과

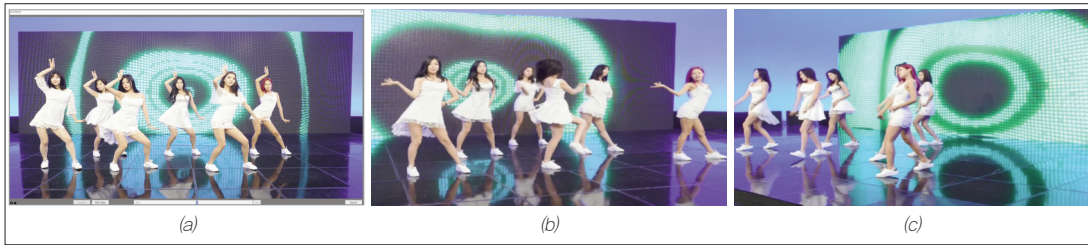
인접한 시점의 영상을 스트리밍할 수 있도록 각 시점별 HTTP 모듈, MP4 파서, AV 디코더가 개별적으로 동작된다. 버퍼 매니저를 통해 각 시점별 영상의 동기화를 맞추며, 미디어 프로세서, 스트림 컨트롤 모듈을 통해 사용자가 선택한 시점과 인접한 시점 영상의 재생 및 시점 전환이 가능하도록 했다.

7. 자유시점 구현 결과

<그림 12>는 좌우로 부드럽게 시점을 바꿀 수 있는 무대공연의 내향시점을 구현한 구동 소프트웨어를 보여준다. 구동 소프트웨어의 하단에 있는 슬라이드 바를 좌우로 움직임에 따라 좌측부터 우측까지 원하는 방향에서 무대 공연을 감상하도록 인터페이스가 구성되었다. (a)는 정면에서 바라본 무대영상이고, (b)와 (c)는 각각 왼쪽과 오른쪽의 방향으로 회전하면서 바라본 무대영상이다.



<그림 11> 자유시점 영상 전송 시스템



〈그림 12〉 내향 시점 표출 결과

Ⅲ. 결 론

본 고에서는 기존의 자유시점 연구에 대한 동향을 간단히 분석하였고, 자유로운 시청자의 뷰 선택 개념을 확장하는 서비스 개발 사례를 소개하였다. 제시한 시스템은 한가지 콘텐츠에서 모두 적용할 필요는 없고, 톨박스 개념처럼 콘텐츠마다 제각기 필요한 기술들을 취사선택할 것이다. 자유선택시점 기술이 지향하는 바는 다음과 같이 정리할 수 있다.

- 자유선택성: 무대 공연, 스포츠, 체험형 게임 등에서 시청자가 원하는 시점과 위치에서 몰입감이 높은 콘텐츠를 제공함으로써 현장감과 체험성을 높임
- 영상과 음향의 연동: 공간상의 위치 및 시점을 자유롭게 선택하고 영상과 음향을 연동시켜 몰입감이 높은 실감 콘텐츠 개발

- 능동적 참여형: 콘텐츠 내에서 직접 참여하는 자유선택시점을 제공하여 새로운 체험을 가능하게 함
- 지능형 영상인식 기술의 활용: 관심 객체에 대한 동작 인식 기술을 동적 3D 모델에 적용
- 상위 개념의 멀티미디어기술: 기존의 UHD, 3D, Free viewpoint 기술을 뛰어넘는 새로운 실감미디어

자유시점 관련 서비스는 기본적으로 많은 카메라를 필요로 하는데 디지털 카메라의 시장점유율이 높은 일본에서 꾸준히 개발되고 있다. 카메라 수를 가급적 최소화하면서도 다양한 뷰를 제공할 수 있는 기술 개발이 필요하다. 아직까지도 본 기술은 적용 분야에 따라 실시간, 주문형 또는 상황 재생 위주로 응용하였다. 아직까지 수작업과 반자동 작업이 포함될 수밖에 없지만 궁극적으로 자동화와 실시간 처리가 가능해질 것으로 기대한다.

참 고 문 헌

- [1] Houari Sabirin, et. al., "Toward Real-Time Delivery of Immersive Sports Content," IEEE Multimedia, vol.25, No.2, April-June, pp.61-70, 2018.
- [2] www.8i.com
- [3] www.stans.co.kr
- [4] S. Song, Y. Park and J. Wee, "DASH-based Streaming Client for View Switching in Free-viewpoint Video Systems," in Proc. ICUFN2019, July, 2017, pp.312-314.

필자 소개



박구만

- 1984년 : 한국항공대학교 전자공학과 공학사
- 1986년 : 연세대학교 전자공학과 석사
- 1991년 : 연세대학교 전자공학과 박사
- 1991년 ~ 1996년 : 삼성전자 신호처리연구소 선임연구원
- 1999년 ~ 현재 : 서울과학기술대학교 전자IT미디어공학과 교수
- 2006년 ~ 2007년 : Georgia Institute of Technology Dept., of Electrical and Computer Engineering, Visiting Scholar
- 2016년 ~ 2017년 : 서울과학기술대학교 나노IT디자인융합대학원 원장
- ORCID : <https://orcid.org/0000-0002-7055-5568>
- 주관심분야 : 컴퓨터비전, 실감미디어