

기술 동향

3차원 TV와 실감방송^②

— 유럽의 3차원 TV

호요성, 이은경
(광주과학기술원 실감방송연구센터)

1. 서론

디지털 정보화 시대를 맞이하여 인터넷을 이용하는 사용자가 급격히 증가하고 있으며, 대용량의 정보를 초고속 네트워크에 실어 전송하는 다양한 기술이 개발되면서 고품질의 멀티미디어 서비스를 요구하는 사용자들도 늘어나고 있다.

최근 들어, 3차원 게임, 3차원 실감방송, 3차원 교육 프로그램과 같이 3차원 오디오/비디오 정보를 이용한 멀티미디어 응용 서비스들도 출현하고 있는데, 이러한 3차원 정보를 포함한 멀티미디어 응용 서비스는 사용자에게 고품질의 영상과 오디오를 제공할 뿐만 아니라, 2차원 정보만으로 제공할 수 없는 다양한 사용자 친화적 상호작용(user-friendly interaction)도 지원할 수 있다. 하지만, 이러한 멀티미디어 응용 서비스를 실현하기 위해서는 기본적으로 3차원 정보를 효율적으로 표현하고 처리할 수 있는 기술이 개발되어야 한다.

3차원 TV는 3차원 오디오/비디오 정보를 이용하여 사용자가 원하는 임의 시점의 영상을 마음대로 시청할 수 있는 차세대 방송 시스템이다. 현재의 일반적인 3차원 영상은 스테레오 영상을 기반으로 하나의 시점에 대한 3차원 영상만을 제공하지만, 미래의 3차원 TV는 다시점의 3차원 영상 정보를 이용해 자유로운 위치에서의 3차원 영상을 제공할 수 있어야 한다.

3차원 TV는 입력 영상으로 다시점 혹은 단일 시점 비디오를 받아서 3차원 정보를 생성하고, 이를 사용자가 원하는 다양한 형태로 재현한다. 즉, 입력 영상을 부호화기를 이용하여 압축 처리하여 전송하고, 전송 채널을 통해 받은 비트열을 수신단에서 복호하여 원하는 영상을 복원한다.

또한 재구성된 3차원 영상을 재생 장치의 특성에 따라 다양하게 변형하거나 혼합하여 재현한다. 재생 장치의 예로는 2차원 디지털 TV, 스테레오 입체 TV, 다시점 3차원 TV 등이 있다. 자유시점 3차원 TV는 사용자의 요구에 따라 임의 시점의 3차원 장면 영상을 재현해야 하기 때문에 주어진 3차원 정보를 이용하여 중간 시점의 영상(intermediate view reconstruction)을 생성하는 기술도 필요하다.

본 논문에서는 최근 유럽에서 개발하고 있는 3차원 TV의 전반적인 개념 및 기술들을 소개하고, 현재 진행 중인 유럽의 3차원 TV 관련 프로젝트와 앞으로의 연구 방향에 대해 살펴본다. 본 논문의 2장에서는 유럽의 3차원 TV 기술을 소개하고, 3장에서는 유럽에서 3차원 TV를 실현하기 위해 개발하고 있는 기술들을 좀 더 자세히 설명한다.

2. 유럽의 3차원 TV 기술 개발의 역사

유럽의 많은 기업과 대학에서는 2차원 TV와 호환을 유지하면서 3차원 TV를 제공할 수 있는 방송 환경을 구축하기 위해 노력하고 있다. 특히, 1998년부터 PANORAMA 프로젝트를 시작하여 깊이 기반의 스테레오 영상을 이용하여 3차원 TV의 잠재적인 실현 가능성을 보여주었다 [1][2]. PANORAMA 프로젝트는 주로 스테레오 기술을 이용한 영상회의 시스템에 초점을 맞추어 연구했으며, 그 뒤 2002년부터 ATTEST(advanced three-dimensional television system technologies) 프로젝트를 진행하면서 3차원 TV를 실현하기 위해 필요한 기초 기술들을 연구하여 개발했다 [3].

ATTEST 프로젝트에서는 우선 3차원 TV의 전체

적인 구조를 정의하고, 3차원 콘텐츠 생성에서부터 압축 부호화, 고속 전송, 다양한 재현 방법에 이르는 3차원 TV의 시스템 개념을 정립했다. 뿐만 아니라, 3차원 TV의 전체적인 구조를 최적화하기 위해서 인간의 3차원 인지 능력에 대한 연구도 함께 진행했다. 사용자 중심의 시스템 구성은 사용자 하여금 좀 더 사실적인 시각 효과를 제공할 수 있다. <그림 1>은 ATTEST 프로젝트에서 제안한 3차원 TV의 전체적인 시스템 개념도이다.

이어 2004년 9월부터 유럽의 대학과 연구소가 컨소시엄을 구성하여 3DTV 프로젝트를 수행하고 있다. 3DTV 프로젝트는 완전한 3차원 장면 모델을 만들고, 이를 재현하기 위한 무안경식 입체 디스플레이와 홀로그래피 등 3DTV에 관련된 포괄적 연구를 수행하고 있다.

3DTV 프로젝트는 이전의 두 프로젝트와는 다르게, 스위스 ETH Zurich 대학은 단독으로 Brick 카메라 시스템을 이용해서 3차원 장면을 복원하는 기술을 개발하고 있다. 이는 여러 위치에 Brick 카메라 시스템을 집중형(convergence)으로 배치해서 3차원 장면 영상을 획득하고, 이를 하나의 3차원 모

델로 융합하는 기술로서, 원하는 시점의 영상을 자유롭게 생성하는데 중요한 기술이다.

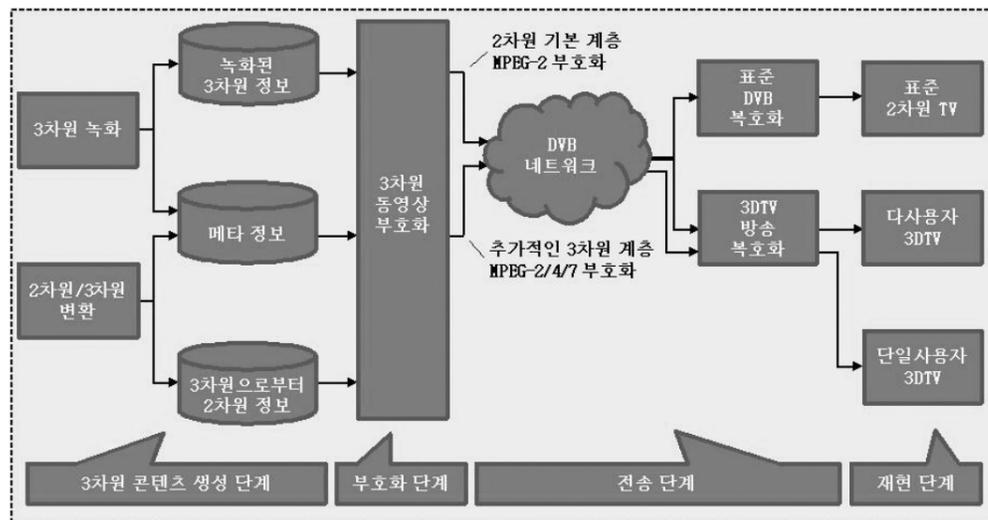
3. 유럽의 3차원 TV 기술 개발

앞 장에서 설명한 것처럼, 유럽의 3차원 TV 기술은 크게 ATTEST 프로젝트와 3DTV 프로젝트, 그리고 ETH Zurich 대학의 3차원 장면복원 프로젝트로 나누어진다. 본 장에서는 각 프로젝트에 대한 기술을 좀 더 자세히 살펴보고자 한다.

1) ATTEST 프로젝트

① 콘텐츠 생성 단계

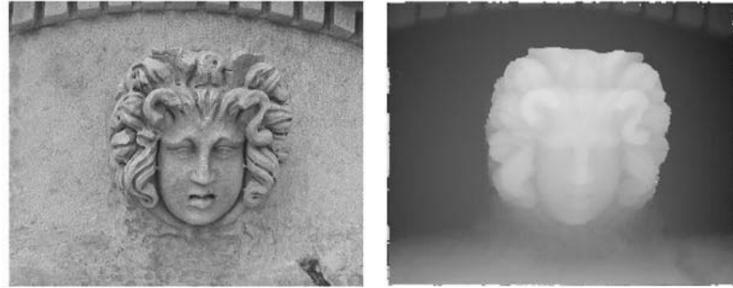
3차원 TV의 콘텐츠를 생성하는 방법은 크게 깊이 카메라를 이용한 방법과 2차원 영상을 이용한 방법으로 나뉜다. 우선 깊이카메라를 이용한 방법은 카메라에 부착된 깊이 센서에서 빛을 보내고, 이 빛이 객체에 부딪혀서 반사되어 돌아오는 빛의 시간을 측정하여 카메라와 객체 간의 거리를 화소 단위로 획득하는 방법이다. <그림 2>는 깊이카메라를 통해 획득한 색상영상과 깊이 맵(depth map)이다.



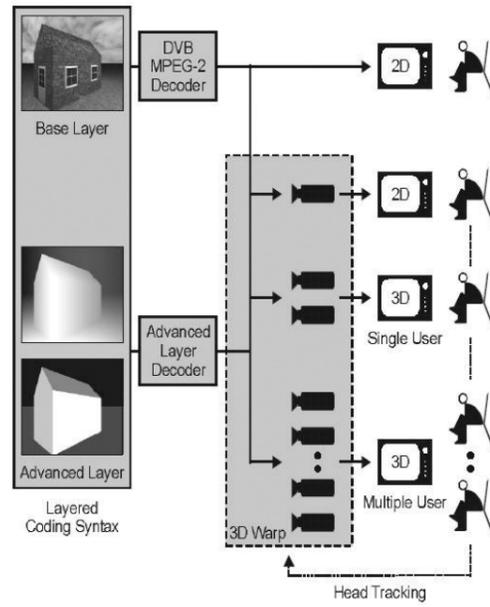
<그림 1> 3차원 TV 시스템 개념도



<그림 2> 깊이 카메라를 통해 획득한 색상영상과 깊이 맵



〈그림 3〉 2차원 영상을 이용해 생성한 색상영상과 깊이 맵



〈그림 4〉 단계적 형태의 부호화 기술

한편, 2차원 영상을 이용한 방법은 2차원 비디오를 이용하여 깊이 맵을 생성하는 방법이다. 3차원 컴퓨터 비전 기술을 이용해 2차원 비디오 안에서 객체의 움직임을 계속적으로 추적하면서, 3차원 공간상의 기하학적 특성을 분석하여 3차원 깊이정보를 생성하는 방법이다. 〈그림 3〉은 이러한 2차원 비디

오를 이용해 깊이 맵을 생성한 결과이다.

② 부호화 및 전송 단계

ATTEST 프로젝트의 가장 중요한 기술 중의 하나는 3차원 TV를 위한 정보의 표현 방법 및 부호화 기술이다. 3차원 영상 정보를 부호화하기 위해서는

많은 사항들이 모두 고려되어야 하는데, 우선 기존의 2차원 TV와의 호환성을 유지해야 하고, 각 사용자의 환경에 따른 계위성(Scalability)도 고려되어야 했다.

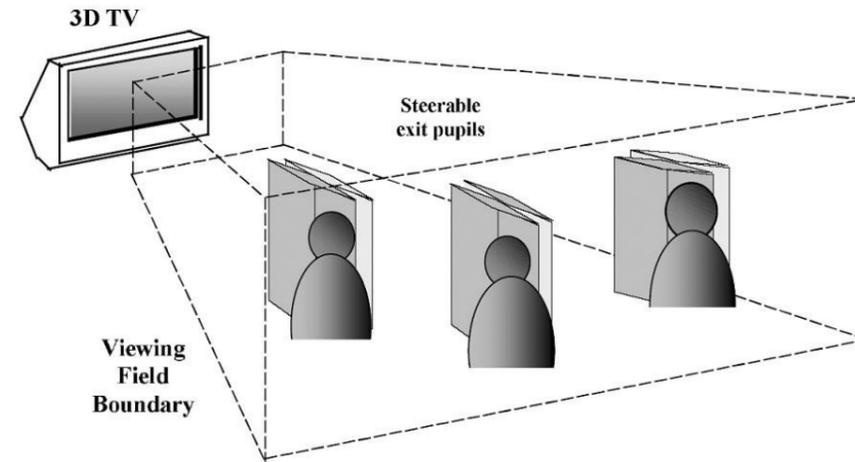
뿐만 아니라, 각 사용자의 재현 장치가 2차원 장치인지 3차원 장치인지 고려해야 했다. 이러한 다양한 조건을 충족시키기 위해서 ATTEST 프로젝트는 〈그림 4〉와 같이 단계적 계층(layer) 형태의 부호화 기술을 정의했다.

3차원 영상 부호화 기술은 기본적으로 기본 계층(base layer)과 최소한 하나 이상의 확장 계층(advanced layer)으로 구성되어 있다. 기존의 2차원 TV를 사용하는 사용자와의 호환성을 유지하기 위해서 기본 계층은 MPEG-2 기반 DVB(digital video broadcasting) 표준을 이용해 부호화하고 복호화했다.

그리고 확장 계층은 3차원 TV를 가진 사용자를 위해 추가적인 정보를 전송해야 하는데, 이 확장 계층에서는 최소한 기본 계층에서 전송한 색상 영상에 상응하는 깊이 맵도 전송해야 한다. 영상의 크기가 큰 콘텐츠이거나 폐색영역(occlusion)이 많은 영상과 같이 콘텐츠의 특성에 따라 분할 마스크나 폐색영역의 영상 정보와 같은 추가적인 정보를 더 전송해야 한다.

③ 재현 단계

3차원 영상을 재현하기 위해서는 단일 사용자 재현 방법과 여러 사용자 재현 방법을 별도로 고려해야 한다. 고품질의 3차원 영상을 효과적으로 재현하기 위해서는 사용자의 위치에 대한 제한이 필요하다. 왜냐하면 3차원 영상을 특정한 위치에서 시청해야만 3차원 입체 영상의 효과를 좀 더 느낄 수 있기 때문이다. 사용자가 여러 명인 경우에는, 사용자의



〈그림 5〉 다사용자 3차원 TV 재현

움직임 추적을 통해 사용자의 시점에 맞는 영상을 동시에 재현하는 기술을 고려해야 한다 [4].

<그림 5>는 사용자가 여러 명인 경우, 재현 방법을 나타내는 그림이다. 재현을 위해서는 최대 4명의 사용자까지 재현이 가능하며, 사용자는 1m에서 3m 사이에 위치해야 하고, 사용자의 위치에 따라서 좌우 30° 정도의 시야각을 고려하여 재현한다. 사용자의 위치를 알기 위해서 HMD(Head Mounted Displays)를 이용했다.

④ 사용자 인지 평가

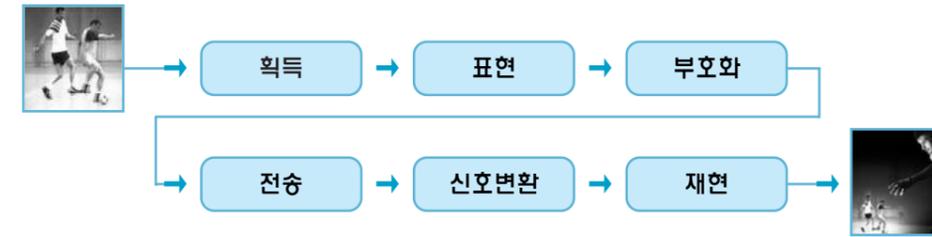
효과적인 3차원 입체 영상 체험을 위해서는 사용자의 인지 능력에 따른 깊이감 제공이 무엇보다 중요하다. 시청자에게 다양한 깊이감을 제공할 경우, 그 시각적인 효과 또한 향상된다. 그러나 지나친 깊이감 제공으로 인해 사용자가 두통이나 눈의 피로함뿐만 아니라 다른 신체적 불편함을 겪을 수 있

다. ATTEST 프로젝트에서는 이러한 부작용을 해결함과 동시에 시각적 효과를 극대화할 수 있는 시스템을 개발하기 위해서 인간의 시각 및 인지 시스템에 대한 연구를 함께 진행했다 [5].

2) 3DTV 프로젝트

3DTV 프로젝트는 3차원 TV의 실현을 위해 교육, 시뮬레이션, 의학, 게임, 멀티미디어, 문화 등 다양한 분야로 그 응용 범위를 확대해 나가고 있다. <그림 6>은 그러한 3DTV에서 제안하고 있는 3DTV를 이미지화한 것이다 [6].

3DTV는 시청자가 원하는 시점의 사실적인 장면 생성을 통해 사용자에게 몰입감을 줄 수 있는 방송 시스템을 의미한다. 3DTV 시스템은 <그림 7>에서 보는 것처럼, 획득, 표현, 부호화, 전송, 신호 변환, 마지막으로 재현 과정으로 나누어진다.



<그림 7> 3차원 TV 시스템 구성도

① 획득 단계

장면 획득 기술은 하나의 카메라 혹은 여러 대의 카메라를 이용하여 장면 영상을 획득하고 모델링하는 기술이다. 뿐만 아니라, 이미 알고 있는 객체를 장면 안에서 추적하거나 경로를 추정하는 기술도 장면 획득 기술에 포함된다. 다시점 카메라를 이용해 장면을 획득할 경우, 먼저 다시점 카메라 간의 동기를 맞추어 하나의 장면을 동시에 획득한 후, 3차원 장면 모델을 생성한다 [7].

여러 대의 카메라로 영상을 획득할 때, 조명의 차이와 같은 문제점 때문에 다시점 영상이 각각 다르게 획득될 수 있다. 이러한 문제는 다시점 영상을 이용해 하나의 3차원 모델을 생성할 때 큰 문제가 된다. 따라서 이러한 다시점 영상 사이의 차이를 해결하기 위해, 모든 다시점 영상에 조명 효과를 일정하게 적용할 수 있는 재조명 기술도 개발되고 있다. 사용자에게 맞는 정보를 제공하기 위해서는 사용자의 얼굴을 인식하는 기술뿐만 아니라, 얼굴의 특징 인식 및 움직임 분석 기술도 연구해야 한다.

<그림 8>은 터키 Bilkent 대학에서 개발한 다시점 영상 획득 장치이다 [6]. 총 8대의 동기화된 카메라

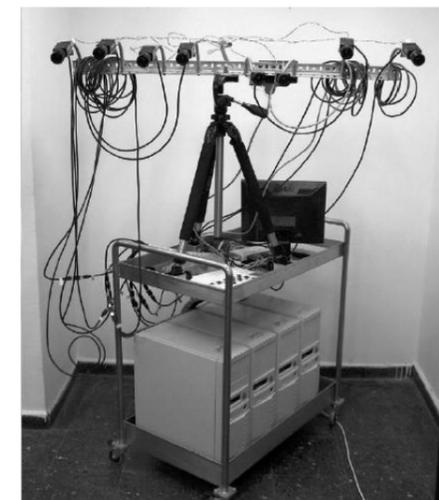
로 구성되어 있고, 640×480의 영상을 초당 30장씩 획득할 수 있다. 카메라 확장 또한 용이한 형태이다.

② 표현 단계

3차원 장면 표현 기술은 획득한 영상 정보를 점 표현 방법, 메쉬 표현 방법, 그리고 부피 표현 방법과 같은 다양한 표현 방법을 이용해 3차원 모델을 생성하는 것을 의미한다 [8].



<그림 6> 3DTV의 3차원 체험



<그림 8> 다시점 영상 획득 장치

3차원 장면을 표현하기 위해서는 실제 장면처럼 보이게 하는 렌더링 방법과 전송할 때의 전송 가능한 형태로 표현이 가능한지를 고려해야 한다. 3DTV 프로젝트에서는 사면체 부피(tetrahedral volume)를 이용해 3차원 객체를 표현하는 기술, 영상 기반 렌더링 방법을 이용해서 3차원 장면을 표현하는 기술, 3차원 영상 기반 비디오 객체를 계위화하여 표현하는 기술도 개발했다.

영상 신호뿐만 아니라, 음성 신호와의 상호작용을 고려하기 위해 영상에서 머리와 손의 움직임 및 행동을 인식하여 이를 동기화된 음성 신호와 합성하는 기술도 개발했다. <그림 9>는 사면체 부피를 이용해 F-16 비행기 모델을 계위화하여 표현한 결과 영상이다.

③ 부호화 단계

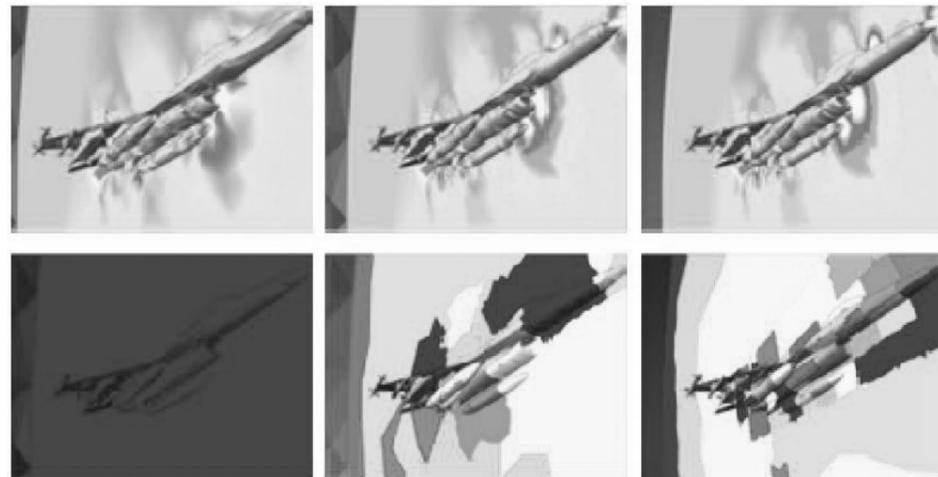
부호화 기술은 3차원 TV를 서비스하기 위해 3차원 장면을 효율적으로 압축하고 전송하는 기술을 의미

한다. 3차원 정보를 이용해 생성한 3차원 영상을 전송하기 위해서는 시스템의 다양성과 전송환경의 성능 차이를 극복할 수 있는 데이터의 압축, 변환 및 전송 분야에 걸친 여러 가지 요소기술이 필요하다.

3DTV 프로젝트에서는 스테레오 3차원 비디오 영상을 자동으로 분할하는 기술과 다시점 영상 전체를 end-to-end 방식으로 전송할 수 있는 비디오 부호화기를 개발했다. 뿐만 아니라 자유시점 비디오를 부호화하고 렌더링하기 위한 다양한 3차원 메쉬 부호화, 워터마킹, 홀로그램 부호화 기술도 개발하고 있다.

④ 재현 단계

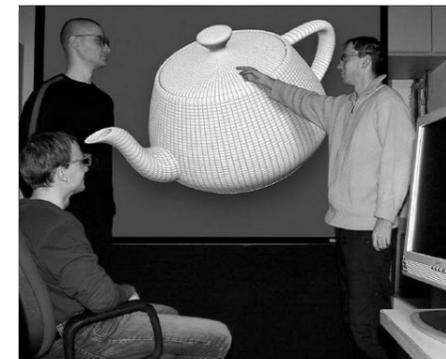
3차원 영상 재현은 인간이 눈으로 보는 세계를 보다 충실히 표현하고, 마치 현장에 있는 것처럼 느끼며 감상할 수 있도록 3차원 장면을 실감나게 재현하는 기술이다. 또한 3차원 영상 재현은 사용자가 취하는 행동에 맞추어 다양한 시나리오와 상호



<그림 9> 사면체 부피를 이용한 3차원 영상 표현

작용을 제공하여 현실감을 제공한다. 3DTV 프로젝트에서는 사용자의 위치를 추적하여 사용자의 시점에 따른 3차원 영상을 자동으로 생성할 수 있는 재현 장치를 개발하고 있다 [9].

<그림 10>은 체코 Plzen 대학에서 개발한 stereowall 이다. 2.5×1.8m 크기의 스크린에 3차원 객체를 보여줄 수 있는 장치이며, 안경방식으로 OpenGL과 DirectX를 이용해서 3차원 장면을 표현한다.



<그림 10> Stereowall

<그림 11(a)>는 독일 Tuebingen 대학에서 개발한 상호작용이 가능한 Fogscreen을 이용한 파노라믹 3차원 모델 획득 및 모델링 시스템이다. 3차원 영상을 획득하기 위해 모바일 로봇 카메라를 설치하거나 사람이 획득 장치를 들고 직접 이동하면서 3차원 정보를 획득한다.

획득 장치는 3대의 레이저 스캐너와 전방향(Ominidirectional) 카메라로 구성되어 있다. 획득한 3차원 영상은 <그림 11(b)>의 오른쪽 그림에 있는 ForScreen™을 이용해 장면을 재현한다.

3) 3차원 장면 복원

스위스 ETH Zurich 대학의 3차원 장면 복원 기술은 획득하고자 하는 객체나 장면을 중심으로 다시점 카메라를 집중형(convergence)으로 배치한다. 획득 장치는 한 대의 고화질 카메라와 영사기, 그리고 두 대의 흑백의 스테레오 카메라가 하나의 시스템으로 구성되어 여러 위치에 배치된다.



(a) 3차원 영상 획득



(b) FogScreen을 이용한 3차원 영상 재현

<그림 11> 상호작용이 가능한 파노라믹 3차원 모델 획득 및 모델링 시스템

영상은 일반적인 비디오 영상과 미리 정의된 패턴을 투사한 장면을 획득한 비디오 영상을 동시에 획득한다. 스테레오 영상을 이용해 3차원 깊이정보를 획득하고, 이렇게 획득한 영상과 깊이정보를 이용해 점 기반의 3차원 모델로 융합하여 장면을 복원하는 시스템이다 [10].

〈그림 12〉는 3차원 장면 획득을 위한 카메라 시스템, 고화질 카메라를 통해 획득한 색상정보, 그리고 흑백의 스테레오 카메라에 일정한 패턴을 투사하여 획득한 영상을 보이고 있다. 여기서 일정하게

투사된 패턴의 크기와 모양을 이용해 3차원 깊이정보를 획득할 수 있다.

여러 위치에서 획득한 색상영상과 깊이 맵을 모두 3차원 공간에 투영시키면 3차원 점으로 구성된 3차원 장면 영상을 획득할 수 있고, 이 정보를 이용해 원하는 시점에서의 3차원 장면 영상을 생성할 수 있다. 〈그림 13〉은 점 정보를 이용해 임의 시점의 영상을 생성한 결과 영상이다.



〈그림 12〉 3차원 장면 획득 장치 및 획득 영상



〈그림 13〉 3차원 복원 정보를 이용해 생성한 임의 시점의 3차원 장면 영상

4. 결론

이 논문에서는 유럽에서 3차원 TV를 구현하기 위해 수행한 여러 프로젝트에서 연구개발한 3차원 영상처리 기술을 대략적으로 살펴보았다. 단일 시점 카메라나 다시점 카메라로부터 3차원 장면 모델을 획득하고, 획득된 3차원 장면 영상을 편집하고 융합하여 3차원 장면 콘텐츠를 생성했다.

이렇게 생성된 3차원 장면 콘텐츠를 고용량의 고속 전송망을 통해 보낸 뒤에, 2차원 TV와의 호환성과 사용자의 단말 환경에 따른 재현 방법을 고려하여 다양한 사용자의 요구에 맞는 적응형 재현 서비스를 가능케 했다.

이와 같이 사용자 상호작용을 지원하는 3차원 TV는 방송과 통신 기술의 융합체로서 점점 늘어가는 시청자의 다양한 요구에 부응할 수 있는 미래형 방송 서비스의 방향을 제시하고 있으며, 앞으로 경제적, 산업적, 사회적으로 많은 기여를 할 것이다. 특히, 고품질의 방송 서비스를 제공하여 인간의 삶의 질을 한층 더 향상시키고, 차세대 성장동력인 IT 산업의 저변을 확대하고 활성화하는데 크게 기여할 것이다.

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 지식경제부 및 정보통신연구진흥원의 대학IT연구센터(ITRC)의 지원에 의한 것이다.

참고 문헌

- [1] DISTIMA, European RACE 2045 Project, <http://www.tnt-uni-hannover.de/project/eu/distima>, 1992-1995.
- [2] PANORAMA, European ACTS AC092 Project, <http://www.tnt-uni-hannover.de/project/eu/panorama>, 1995-1998.
- [3] A. Redert, M. O. Beeck, C. Fehn, W. Ijsselstein, "ATTEST-Advanced Three-dimensional Television System Technologies," Proceedings of the First International Symposium on 3D Data Processing, Visualization and Transmission, pp. 313-319, 2002.
- [4] E. Kurutepe, M. Civanlar, and A. Tekalp, "Interactive Multi-view Video Delivery with View-point Tracking And Fast Stream Switching," Lecture Notes in Computer Science Volume, vol. 4105, pp. 586-593, 2006.
- [5] W. Usselstein, H. Ridder and J. Vliegen, "Subjective Evaluation of Stereoscopic Images: Effects of Camera Parameters And Display Duration," IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, pp. 225-233, 2000.
- [6] 3DTV EU NoE Project, <http://www.3dtv-research.org>
- [7] E. Aguiar, C. Theobalt, M. Magnor, H. Seidel, "Reconstructing Human Shape And Motion from Multi-view Video," European Conference on Visual Media Production, p. 44-51, 2005.
- [8] A. Alatan, Y. Yemez, U. Gudukbay, X. Zabulis, K. Muller, C. Erdem, C. Weigel, A. Smolic, "Scene Representation Technologies for 3DTV-A Survey," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, pp. 1587-1605, 2007.
- [9] L. Onural, T. Sikora, and A. Smolic, "An Overview of A New European Consortium: Integrated Three-Dimensional Television - Capture, Transmission and Display (3DTV)," Proceedings of European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, 2004.
- [10] M. Waschbüsch, S. Wurmlin, D. Cotting, M. Gross, "Point-sampled 3D Video of Real-world Scenes," Proceedings of Graphite, p.469-477, 2006.