

3차원 비디오 부호화 표준화 동향

광주과학기술원 실감방송연구센터 ■ 호요성 · 이상범

1. 서론

최근 디지털 신호처리 및 전송 기술이 빠르게 발전하면서 다양한 멀티미디어 서비스가 급속하게 늘어나고 있지만, 아직도 3차원 멀티미디어 응용 서비스를 위해서는 관련 기술의 개선 및 신기술의 개발이 절실히 요구되고 있다. 초기 영상 서비스 기술은 방대한 양의 비디오 데이터를 효율적으로 저장하거나 전송하기 위해서 정보를 압축하는 단계부터 발전하기 시작하여 현재는 3차원 입체 TV를 비롯한 실감형 방송 서비스를 제공하는 기술까지 연구되고 있다[1]. 여기서 3차원 비디오란 기존의 단일 시점 카메라를 이용하여 평면적인 장면을 획득하는 방법과는 달리, 여러 대의 카메라나 능동형 센서를 이용하여 얻은 깊이 정보를 포함한 입체적인 동영상을 일컫는다.

3차원 비디오 신호를 획득하기 위해 양안시점(stereoscopic) 카메라, 다시점(multi-view) 카메라, 깊이 카메라, 홀로그래피(holography) 등과 같은 장치들이 개발되었으며, 이를 이용하여 얻은 3차원 비디오 신호를 처리하고 부호화하는 기술들이 연구되고 있다. 특히, 두 대 이상의 카메라를 이용하여 촬영된 다시점 비디오는 최근 3차원 영상처리의 핵심 연구 대상으로 주목을 받고 있다[2]. 3차원 비디오는 사용자들에게 자유로운 시점과 넓은 시야각을 통해 입체감과 임장감을 제공하는 장점을 지니고 있다. 하지만, 다시점 비디오의 경우 카메라 수에 비례하여 데이터양이 증가하고, 깊이 카메라의 경우 장비가 고가라는 문제점이 있기 때문에 다양한 입체 영상 서비스 개발이 상당히 제한적이다.

이에 멀티미디어 콘텐츠의 압축 부호화에 대한 국제표준 규격을 제정하는 국제 조직인 MPEG(moving picture experts group)에서는 2001년 12월부터 3차원 오디오 비주얼(3-dimensional audio-visual, 3DAV)이란 이름으로 3차원 영상처리에 관한 기술을 표준화

하는 작업을 시작했다. MPEG에서는 MPEG-2(ISO/IEC 13818-2)와 MPEG-4(ISO/IEC 14496-2)에서 이미 평행형 다시점 영상 부호화와 3차원 메쉬(mesh) 압축 부호화 기능을 제공했다. 하지만, 좀 더 구체적인 비디오 부호화의 필요성에 따라 3DAV 부호화 작업을 다시점 비디오 부호화(multi-view video coding, MVC)란 이름으로 바꾸고, 2006년 7월에 열린 제77차 MPEG 회의에서부터 MVC에 관한 표준화 작업을 JVT(joint video team)에서 본격적으로 맡아 진행했다. 최근에는 MPEG에서 자유시점 TV(free-viewpoint television)에 대한 표준화 작업을 재개하여 다시점 비디오와 깊이 영상을 기반으로 한 3차원 비디오에 대한 표준화 작업을 활발히 진행하고 있다[1,2].

본 논문에서는 3차원 멀티미디어 서비스를 위한 3차원 비디오 압축 표준화 활동에서 다루는 기술을 전체적으로 살펴보고자 한다. 2장에서는 다시점 비디오 부호화에 대한 일반적인 내용을 소개하고, 최근에 연구되고 있는 주요 기술을 설명한다. 3장에서는 최근 일본을 중심으로 연구되고 있는 자유시점 TV 관련 영상처리 기술을 소개한다. 4장에서는 MPEG에서 진행 중인 3차원 비디오에 관한 표준화 동향과 주요 기술을 살펴본다.

2. 다시점 비디오 부호화

2.1 다시점 비디오 시스템

그림 1은 다시점 비디오 시스템의 전체적인 구성을 나타낸다. 다시점 비디오 시스템은 동일 시간에 여러 대의 카메라를 이용하여 3차원 장면을 획득한 후, 이를 부호화하여 전송하고, 복원한 다시점 영상을 다양한 디스플레이 장치를 통해 재현하는 시스템이다. 우리는 다시점 영상을 이용하여 사용자가 원하는 시점에서의 영상을 선택적으로 재생하거나, 인접한 시점의 두 영상 혹은 다시점 영상 전체를 이용하여 3차원 입체 모니터나 다시점 모니터를 통해 사용자에게 현실감 있는 3차원 영상을 제공할 수 있다.

* 본 연구는 지식경제부 및 정보통신연구진흥원의 대학IT 연구센터(ITRC)의 지원에 의한 것입니다.

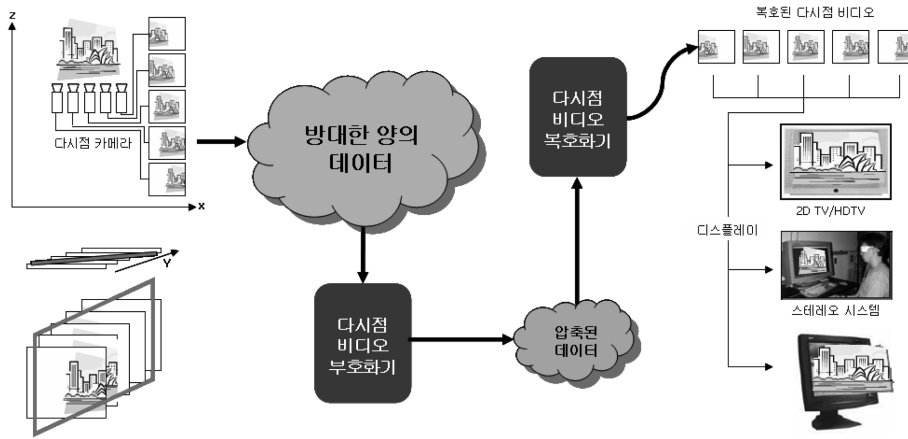


그림 1 다시점 비디오 시스템

한편, 다시점 영상은 기존의 방식과는 달리 여러 대의 카메라를 이용하여 장면을 촬영하기 때문에 데이터의 양이 카메라 수에 비례하여 증가한다. 따라서 다시점 비디오 시스템을 실용화하기 위해서는 다시점 영상을 효율적으로 획득하여 처리하고 부호화할 수 있는 기술을 개발해야 한다.

2.2 다시점 비디오 부호화의 응용

2.2.1 자유시점 TV(free-viewpoint television)

자유시점 TV는 조밀한 간격으로 배치된 여러 대의 카메라로 3차원 장면을 촬영하고, 이를 바탕으로 광선 공간(ray space)을 정의한다. 이렇게 정의한 광선 공간을 이용하여 사용자의 요청에 따라 임의의 시점 영상을 자유로이 생성할 수 있다. 최근 일본에서는 자유시점 TV에 대한 집중적인 연구를 진행하고 있는데, 다시점 카메라 시스템을 이용한 자유시점 TV는 스포츠 방송, 게임, 교통 통제 시스템 등 다양한 분야에서 활용될 수 있다.

2.2.2 3차원 TV(three-dimensional television)

다시점 비디오의 또 다른 응용 분야로 3차원 TV를 들 수 있다. 현재의 3차원 영상은 스테레오 영상을 기반으로 하나의 시점에 대한 입체 영상을 제공하지만,

3차원 TV는 다시점 비디오를 이용하여 더 넓은 시야각을 가지는 입체 영상을 제공할 수 있다. 그러므로 시청자는 위치를 바꾸어 보다 다양한 시점의 입체 영상을 자연스럽게 즐길 수 있다. 3차원 TV는 다양한 시점의 입체 영상을 지원하는데 촬영한 시점의 수가 재생되는 시점의 수보다 부족하거나 촬영한 시점 간의 간격이 클 경우, 중간시점의 영상을 생성(intermediate view reconstruction, IVR)하여 3차원 TV의 시점을 자연스럽게 전환할 수 있다.

2.2.3 몰입형 원격회의(immersive teleconferencing)

몰입형 원격회의란 기존의 단일 시점을 기반으로 하는 원격회의 시스템에 다시점 비디오를 적용하여 보다 입장감 있는 영상을 재현하는 원격회의를 일컫는다. 서로 다른 장소에서 촬영한 원격회의 영상을 입체적으로 재생하거나, 원하는 패널의 영상을 자유롭게 선택하여 볼 수 있기 때문에, 보다 현실감 있는 입체 영상 원격회의를 할 수 있다.

2.3 다시점 비디오 부호화 표준화 작업

2.3.1 테스트 영상(test video sequences)

현재 다시점 비디오 부호화 표준화 작업에서 권고하고 있는 테스트 영상은 사용한 카메라의 개수, 카메라

표 1 다시점 비디오 테스트 영상의 특성

테스트 영상	해상도 및 프레임율	교정 여부	카메라 개수	카메라 구성
Ballroom	VGA, 25fps	교정함	8	1차원 평행, 20cm 간격
Exit	VGA, 25fps	교정함	8	1차원 평행, 20cm 간격
Race1	VGA, 30fps	교정하지 않음	8	1차원 평행, 20cm 간격
Flamenco2	VGA, 30fps	교정하지 않음	5	2차원 평행, 직교, 20cm 간격
Uli	XVGA, 25fps	교정하지 않음	8	1차원 평행, 수평, 20cm 간격
Breakdancers	XVGA, 25fps	교정하지 않음	8	1차원 원호, 20cm 간격
Rena	VGA, 30fps	교정함	100	1차원 평행, 5cm 간격
Akko&Kayo	VGA, 30fps	교정하지 않음	100	2차원 배열, 수평 5cm, 수직 20cm 간격

라 배치, 초당 프레임 수, 영상의 해상도, 장면의 복잡도, 카메라의 움직임 등을 고려하여 선정되었다. 또한, 선택된 테스트 영상은 모두 카메라 변수를 포함하고 있으며, 다시점 비디오 표준화 그룹 내에서 이에 대한 검증 작업을 완료했다. 현재 표준화 그룹에서 제공하는 다양한 형태의 다시점 비디오 테스트 영상의 특성을 표 1에 정리했다[3].

이러한 테스트 영상들은 다음 장에서 소개할 부호화 조건 및 참조 소프트웨어를 통해 해당 부호화 알고리즘의 성능을 평가하는데 사용된다. 각 테스트 영상들은 이를 촬영한 기관을 통해 인터넷으로 다운로드할 수 있으며, 특히, ‘Breakdancers’ 영상의 경우에는 깊이 정보도 함께 제공된다[3].

2.3.2 부호화 조건과 참조 소프트웨어(reference software)

2006년 7월에 MVC 표준화 작업이 MPEG에서 JVT로 넘어가면서, 이전에 사용했던 각 영상별 실험 비트율이 바뀌어 통일된 양자화 계수를 이용하도록 권고했다[3]. ‘Rena’와 ‘Akko&Kayo’ 영상은 100대의 카메라를 이용하여 획득한 테스트 영상이므로, 공간적인 밀집도를 이용하면서 성능 실험의 복잡도를 줄이기 위해 한정적인 시점만을 사용한다. ‘Rena’ 영상은 중간의 16시점을 사용하며, ‘Akko&Kayo’ 영상에 대해서는 3x5의 2차원 배열이 선택되어 사용된다. 다시점 비디오 압축 부호화 방식의 성능 실험에는 각 시퀀스의 10초 분량의 영상만 사용한다.

2005년 10월에 다시점 비디오 부호화에 대한 참조 소프트웨어로 Fraunhofer-HHI의 소프트웨어가 선정되었으며, 2006년 2월에 JSVM3.5(joint scalable video

model)를 기반으로 한 참조 소프트웨어가 배포되었다. 이어서 2006년 7월에 다시점 비디오의 병렬처리를 위해 각 시점별로 독립적으로 부호화 하는 JMVM1.0(joint multiview video model)이 배포된 이래, 2008년 1월에는 JMVM6.0이 배포되었다. 기본적으로 시점 방향에 대해서는 IBPBP... 구조를, 시간 방향에 대해서는 계층적(hierarchical) B화면 구조를 사용하고 있다. 그림 2는 다시점 비디오 부호화에서 사용하고 있는 참조 소프트웨어의 기본적인 시공간 예측 구조를 보여주고 있다[4].

그림 2에서 S_n 은 시점축으로 n 번째 시점의 카메라를 의미하고, T_n 은 시간축으로 n 번째 화면을 의미한다. 화살표는 대상 화면을 예측하는데 참조되는 화면들 사이의 관계를 나타내며, 시점 방향으로는 IBPBP 구조를 이용하고, 시간 방향으로는 계층적 B화면 구조를 이용한다. 여기서 계층적 B화면 구조란, 그림 2에 보인 것처럼 하나의 GOP 구조에서 GOP의 양 끝 화면(B_n)을 제외하고 중간 화면들은 모두 B화면으로 이루어진 GOP이다. 즉, 해당 GOP의 중간에 위치한 화면(B_{n+1})이 양 끝 화면을 이용하여 먼저 예측 부호화되고, 다시 해당 중간 화면과 양 끝 화면 사이의 화면들(B_{n+2})이 전후의 부호화된 화면들로부터 계층적으로 부호화되는 GOP 구조를 말한다. 현재 공간적 예측 구조에 대한 부호화 순서는 각 시점을 따로따로 부호화한다. 그림 2의 경우, I 시점인 S_0 를 처음 부호화하고, S_0 의 복원 영상을 이용하여 P 시점인 S_2 를 부호화한다. 마찬가지로 방법으로, S_0 와 S_2 의 복원 영상을 이용하여 B 시점인 S_1 을 부호화한다.

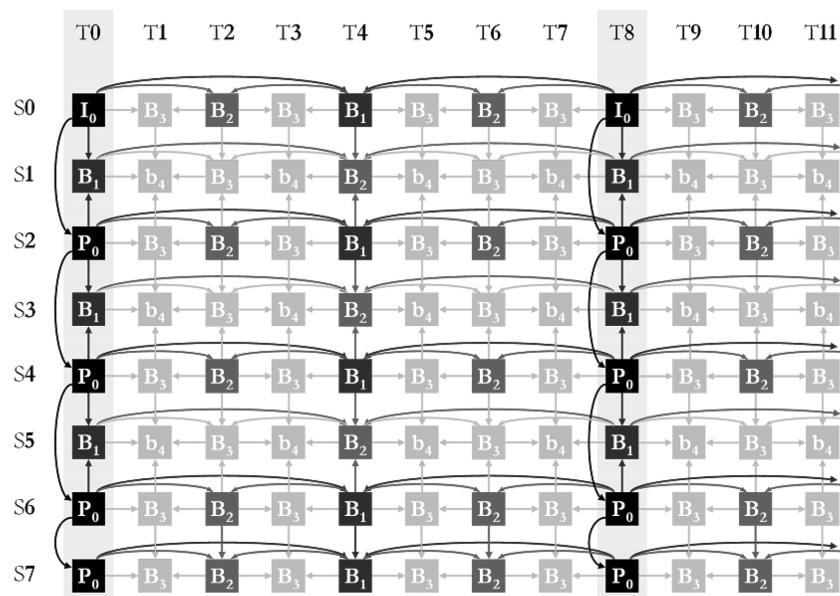


그림 2 다시점 비디오 참조 소프트웨어의 시·공간 예측 구조

2.4 다시점 비디오 부호화 기술

2.4.1 조명보상 기술(illumination compensation)

다시점 비디오는 카메라의 위치, 카메라 모델의 제조상의 차이, 조리개 조절 및 촬영 시 조명 광원의 위치 등의 문제로 각 시점에 대한 조명의 차이가 발생하게 된다. 조명이 일치되지 않은 다시점 영상을 재생하면 시점간 예측 구조에서 이득을 보기 힘들기 때문에, 다시점 비디오 부호화에 큰 장애 요인이 된다. 이러한 문제점을 해결하기 위해 다양한 조명보상 기술이 제안되었는데, 최근 움직임 예측/보상 과정에서 매크로블록 단위로 조명을 보상해 주는 방법(illumination change-adaptive motion compensation)이 JMVM 2.0에 구현되었다[5].

블록 기반의 조명보상 기술은 다시점 비디오 부호화의 테스트 조건 하에서 부호화 효율이 평균 0.2dB 정도 개선되는 효과를 보였다. 하지만 이 방법은 블록 단위로 조명을 보상하기 때문에 블로킹 현상(blocking artifact)이 발생하게 된다. 이를 해결하기 위해 2007년 5월에 배포된 JMVM 4.0에는 B_Skip과 Direct_16x16 모드에 대해 조명보상 관련 정보를 유추하여 사용하는 방법과 블록 단위의 조명 보상으로 인해 발생하는 블로킹 현상을 줄이기 위해 복원 영상을 만드는 과정에서 경계 부분에 대한 필터의 강도(boundary filtering strength, bS)를 조정하는 방법이 추가로 구현되었다[6]. 이 방법으로 조명 보상으로 인한 블로킹 현상을 줄이면 부호화 효율이 좋아지는 않더라도 주관적인 화질의 개선에는 도움이 된다.

2.4.2 움직임 정보 생략 모드(motion skip mode)

다시점 비디오는 하나의 장면을 여러 대의 카메라로 찍은 것이므로 인접하는 시점 사이의 상관도가 매우 높다. 움직임 생략 모드는 시점 사이의 상관도를 고려하여 이전에 부호화된 인접 시점의 움직임 정보를 공유함으로써 부호화 효율을 높이는 방법이다.

그림 3은 인접 시점과 현재 시점간의 매크로블록간의 상관관계를 보여준다. 이미 부호화된 인접 시점에서 T0 시간대의 매크로블록 P0'이 T1 시간대의 P1'을 참조하여 부호화된 경우에 움직임의 정도가 MV_{nbr} 이고, 현재 시점의 T0 시간대의 매크로블록 P0가 T1 시간대의 매크로블록 P1을 참조하여 부호화하는 경우의 MV_{cur} 는 MV_{nbr} 와 유사하고, 두 시점간의 상대적인 변위 dv 와 dv' 도 유사하다고 가정한다. 이 경우 P0를 부호화할 때에 필요한 움직임 정보를 P0로부터 공유하여 사용할 수 있다는 것이 바로 움직임 생략 모드의 중심 아이디어이다. 움직임 정보 생략 모드 기술

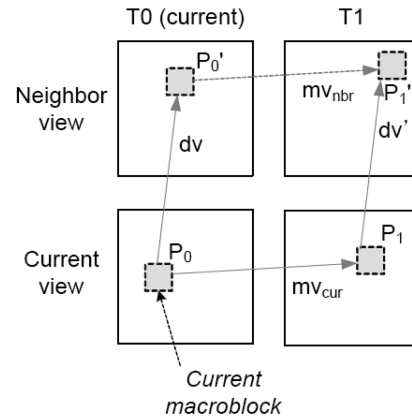


그림 3 인접한 두 시점에서의 매크로블록간의 상관관계

은 크게 인접 시점에서 상용 매크로블록을 찾는 단계와 상용 매크로블록으로부터 움직임 정보를 유도하는 단계로 나뉘어 수행된다. 이 기술은 테스트 조건을 준수하여 실험한 결과, 움직임 정보 공유 모드 기술은 평균적으로 0.18dB 정도의 부호화 효율을 향상시켰다.

2.4.3 적응적인 참조 화면 필터링(adaptive reference filtering)

다시점 카메라 시스템은 여러 대의 카메라를 독립적으로 조절해야 하기 때문에 시점 간 화면의 초점이 맞지 않는 문제가 발생할 수 있다. 한 시점의 영상은 전경에 초점이 맞춰 촬영되고, 다른 한 시점의 영상은 배경에 초점이 맞춰져 촬영되면, 두 영상 사이의 상관도는 떨어지게 되며, 인접 시점을 참조하여 부호화하는 시점 방향의 예측 효율이 떨어지게 된다. 이를 보정하기 위해 최근 참조화면의 전경과 배경에 대한 필터링을 통해 마치 참조화면이 부호화할 화면과 동일한 초점거리에서 촬영된 것처럼 보정하는 방법이 제안되었다[7]. 이를 통해 초점의 불일치로 인해 발생하는 참조화면과 부호화할 화면 사이의 상관도 저하를 보정할 수 있어 부호화 효율의 향상을 기대할 수 있다.

3. 자유시점 비디오 압축 부호화

3.1 광선 공간(ray space)의 정의

3차원 공간을 가로지르는 광선(ray)은 세 개의 공간 좌표(x, y, z)와 두 개의 방향 좌표(θ, ϕ)를 사용하여 5차원 변수 공간 내에 한 점으로 사상(mapping)될 수 있으며, 이와 같은 변수 공간을 “광선 공간”이라고 한다. 광선 공간 내에서 영상을 얻는 과정은 다음과 같다. 영상을 획득하기 위해 핀홀(pin hole) 카메라를 임의의 위치에 놓고, 그 위치에서 광선의 세기(intensity)

정보를 얻는다. 이 과정에서 5차원 변수 공간 정보는 중복도가 높으므로 실제 응용에서는 잘 사용되지 않고, 대부분의 3차원 영상 응용에서는 투사(projection) 방법을 사용하여 만든 4차원 변수 공간 정보를 사용한다.

앞서 언급했듯이, 3차원 공간 내의 하나의 광선(ray)은 광선 공간에서 하나의 점으로 표현될 수 있다. 광선 공간은 많은 시점의 영상들로 구성되고 광선 공간의 한 단면은 대응하는 시점의 실제 영상을 나타낸다. 3차원 공간 내의 시각 정보는 광선들로 표현되고, 광선 데이터는 3차원 공간 내의 임의의 시점에서 보이는 장면의 모든 정보를 포함한다. 그러므로 광선 공간을 샘플링하면 임의의 시점의 영상을 현실감 있게 재생할 수 있다.

광선 공간은 가상의 공간이지만 실제 공간(real space)과 직접적으로 연관되어 있다. 광선 공간은 카메라 변수들을 가진 다시점 영상들을 모아서 만들 수 있다. 자유시점 TV 시스템에서는 두 종류의 광선 공간을 사용하는데, 하나는 직교(orthogonal) 광선 공간이고, 다른 하나는 구형(spherical) 광선 공간이다. 직교 광선 공간은 평행 카메라 배치를 사용한 자유시점 TV 시스템에서 사용되고, 구형 광선 공간은 수렴 카메라 배치를 사용한 자유시점 TV 시스템에서 사용된다.

3.2 광선 공간 기반의 자유시점 TV 시스템

광선 공간은 자유시점 TV를 구현하는데 적합한 3차원 입체 영상 표현 방법이다. 광선 공간을 이용하면 임의 위치의 기하학 정보를 사용하지 않고도 해당 위치에서의 가상 영상(virtual view)을 쉽게 만들 수 있다. 광선 기반 자유시점 TV 시스템은 광선 획득(ray acquisition), 광선 처리(ray processing)와 광선 디스플레이(ray display)로 구성된다.

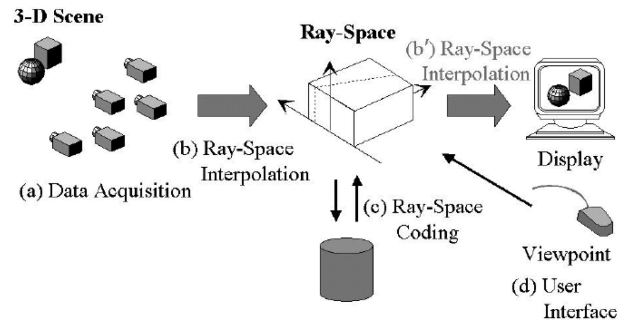


그림 4 자유시점 TV 시스템 구조도

그림 4는 광선 공간을 이용한 자유시점 TV의 시스템 구조도를 나타낸다. 광선 공간을 이용한 자유시점 TV의 주요 처리 과정은 다시점 영상 획득, 영상 보간(view interpolation), 영상 부호화, 자유시점 영상의 재현 과정으로 구성된다. 그림 4에 보인 것처럼 자유시점 TV는 여러 대의 카메라를 사용하여 3차원 장면의 영상 정보를 획득한다. 다수의 카메라로 획득한 정보는 3차원 장면을 구성하고 복원하기 위해 광선 공간으로 사상되어 처리된다.

3.3 다시점 영상 획득과 광선 공간상의 표현

자유시점 TV 시스템에서는 다수의 카메라를 여러 방향으로 배치하여 다시점 영상을 획득한다. 평행 카메라 배치의 경우, 각 카메라로부터 획득한 영상들은 광선 공간 내에서 수평으로 평행하게 위치한다. 그림 5는 여러 대의 카메라로부터 획득한 영상을 광선 공간 내에서 재구성한 것을 나타낸다.

3.4 중간 시점 영상의 보간(intermediate view interpolation)

자유시점 TV 시스템을 실현하기 위해서는 많은 카메라로부터 획득된 많은 시점의 영상이 필요하다. 하

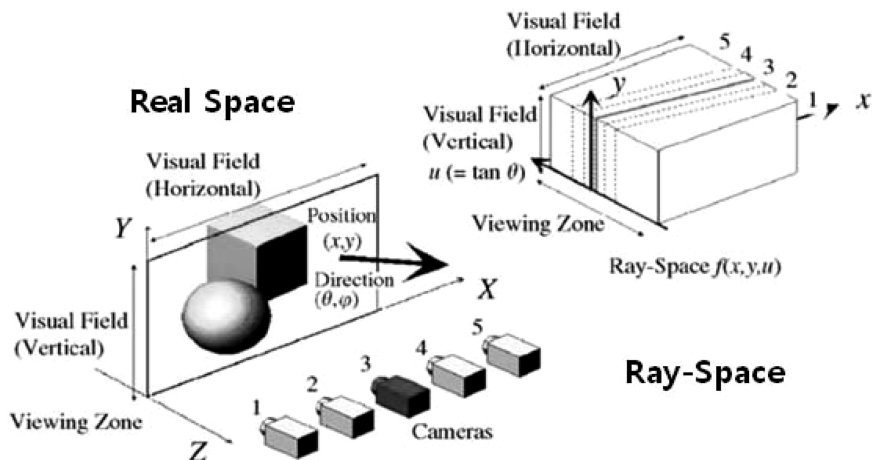


그림 5 광선 공간을 이용한 자유시점 TV 시스템의 영상 획득 과정

지만 실제로 사용할 수 있는 카메라의 수는 제한적이기 때문에, 제한된 개수의 카메라로부터 획득된 영상을 이용하여 실제 카메라로부터 획득되지 않은 임의의 가상 시점의 영상을 복원해야 하는데, 이러한 과정을 시점 보간(view interpolation)이라고 한다.

시점 보간 방법은 크게 PMI(pixel-based matching interpolation) 방법과 BMI(block-based matching interpolation) 방법으로 구분된다[8,9]. PMI 방법은 가장 가까이 존재하는 EPI(epipolar plane image) 선(line)의 쌍으로부터 선형 방정식(linear equation)을 사용하여 가장 최적의 화소(pixel)쌍을 찾은 후, 이를 이용하여 보간하는 방법이다. BMI 방법은 가까이 존재하는 EPI 라인 쌍에서 한 화소씩 이동하면서 최적의 화소 블록 쌍을 찾은 후, 블록 쌍의 가운데 화소를 이용하여 보간하는 방법이다.

3.5 영상 부호화

자유시점 TV 시스템은 다시점의 영상과 카메라 변수 등 상당히 많은 데이터를 사용한다. 이러한 데이터는 카메라 수에 비례하여 증가하기 때문에 효율적인 데이터 압축은 자유시점 TV 시스템 구현에 있어 매우 중요하다. 가장 간단한 자유시점 TV 데이터의 부호화 방법은 각 시점의 영상을 기존의 단일 시점에 대한 부호화 방법을 사용하여 독립적으로 부호화하는 것이다. 그러나 이 방법은 여러 시점의 영상들이 가지는 상관도를 무시하고 독립적으로 부호화하기 때문에 비효율적이다. 그러므로 자유시점 TV 시스템에서는 모든 데이터를 광선 공간의 구조를 고려하여 효율적으로 부호화해야 한다.

3.6 자유시점 영상의 재현

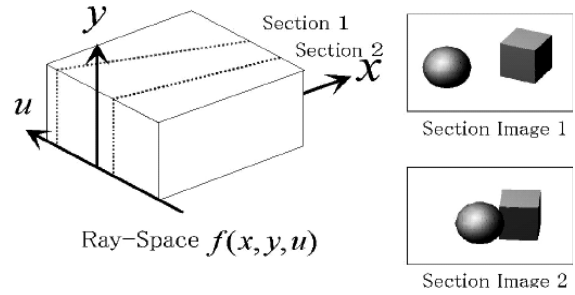


그림 6 임의의 시점 영상의 재현

자유시점 TV의 영상 재현 과정은 매우 간단하다. 그림 6에서 볼 수 있듯이, 사용자가 재현하고자 하는 시점의 영상은 광선 공간의 한 면을 절단하여 획득할 수 있다. 획득된 영상은 2차원 또는 3차원 디스플레이 장치를 통해 재현된다.

3.7 사용자 상호작용

일반 TV 시스템과 달리 자유시점 TV 시스템의 중요한 특징 중의 하나는 사용자 상호작용이 가능하다는 점이다. 이는 임의의 시점에 대한 영상이 재현 가능한 자유시점 TV의 특징 때문에 가능하다. 자유시점 TV 시스템은 사용자의 위치에 대응하는 장면을 디스플레이 장치에 업데이트하여 화면에 재현한다. 그림 7은 사용자의 위치에 대응하는 장면을 실시간으로 업데이트하는 시스템을 나타내고 있다. 그림 7에서 사용자가 오른쪽에 있을 때 디스플레이 장치는 오른쪽 시점에 대한 영상을 재현하고, 사용자가 왼쪽으로 이동함에 따라 재현되는 영상 역시 시점이 왼쪽으로 이동하는 것을 알 수 있다. 또한 시점이 수평 방향뿐만 아니라 수직 방향으로도 이동할 수 있음을 나타내고 있다.

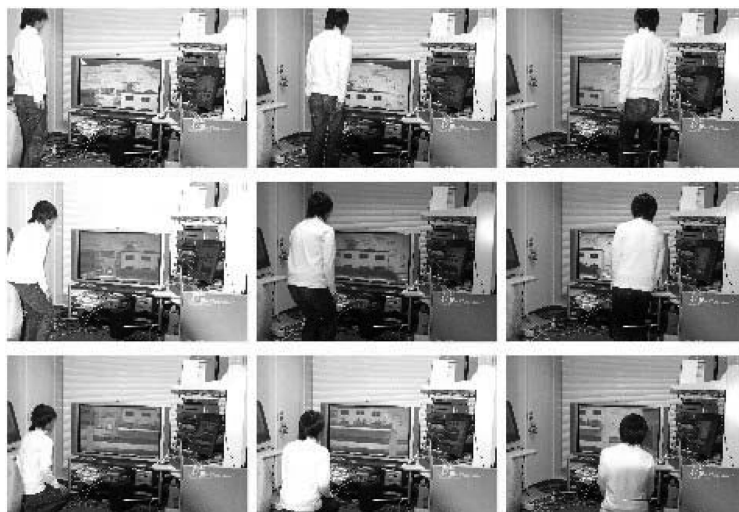


그림 7 사용자의 위치에 대응하는 장면을 업데이트하는 시스템

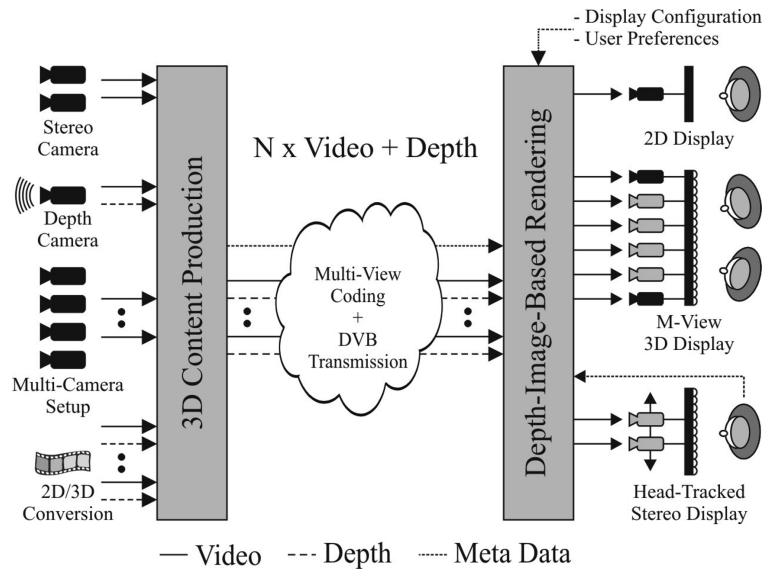


그림 8 일반적인 자유시점 TV 시스템의 구조도와 데이터 형식의 예

4. 3차원 비디오 부호화

4.1 3차원 비디오의 개념

자유시점 TV를 위한 표준화의 첫 번째 초점은 3차원 비디오이다. 3차원 비디오란 스테레오 카메라, 다시점 카메라, 깊이 카메라를 이용하거나 2D/3D 변환 과정을 통해 획득한 모든 영상을 말하며, 단일 시점의 일반 카메라로부터 획득한 기존의 2차원 영상과는 달리 다양한 시야각을 가지는 3차원 장면 재현이 가능하다.

그림 8은 3차원 비디오를 전송하는 일반적인 자유시점 TV 시스템의 예를 보여준다[10]. 다양한 장치를 통해 획득한 3차원 비디오를 부호화하여 전송하고, 수신단에서는 깊이영상 기반 렌더링 방법을 이용해 다양한 형태의 디스플레이에 재현할 수 있다. 여기서 다시점 입체 디스플레이는 N개의 시점을 동시에 사용자에게 제공한다. 그러나 시스템의 효율성을 위해 실제 시점 영상보다 적은 수 $K(K \leq N)$ 개의 시점 영상만을 전송하고, K개의 시점에 대한 추가적인 깊이 데이터를 제공한다. 수신 측에서는 N개의 시점을 디스플레이 하기 위해, 전송된 K개의 시점과 깊이 정보, 깊이 영상 기반 렌더링 방법을 이용하여 N개의 시점을 생성한다. 즉, 수신단에서 고화질의 중간영상을 생성하여 사용자로 하여금 실감나는 3차원 장면을 재현하도록 한다. 또한, 이러한 구조는 자유시점 기능과 다시점 디스플레이를 위한 시점 생성을 가능하게 한다.

그림 8과 같은 시스템은 시야각이 작은 3차원 비디오(20° 이하)를 획득해야 한다는 제약이 있다. 또한 이러한 시스템은 수신단에 카메라 교정(rectification) 작

업이 포함되어선 안된다. 이는 만약 카메라 교정 작업이 필요하다면, 부호화 단의 입력 시점에서 이미 수행되어야 함을 의미한다.

4.2 3차원 비디오의 응용

앞서 언급했듯이, 3차원 비디오는 기존의 2차원 비디오와는 달리 한 장면을 여러 시점에서 시청 가능하도록 하는 자유시점 변환 기능을 제공한다. 이러한 기능을 바탕으로 3차원 비디오를 다양한 분야에 효율적으로 이용할 수 있다. 예를 들어, 보다 실감나는 입체 영상을 기반으로 하는 게임이나, 전시 또는 이벤트에 효율적으로 이용할 수 있다. 한 사물을 다양한 각도에서 실시간으로 볼 수 있다는 장점을 이용하면, 보다 현장감 있는 학습용 소프트웨어 또한 이용이 가능하다. 의학적으로는 실감형 가상수술 교육 프로그램을 제작하여 보다 효과적인 교육을 수행할 수 있다. 뿐만 아니라, 입체영화나 입체형 TV 방송용으로 다양하게 이용할 수 있다.

4.3 3차원 비디오 처리 기술

3차원 비디오는 다양한 시점의 입체 영상을 제공한다는 점에서 기존의 2차원 영상보다 발전된 시스템이라 할 수 있으며, 다시 말해 이는 영상의 획득/처리/전송/재생 등 모든 과정에서의 발전을 의미한다. 그림 9는 3차원 비디오의 주요 기술을 기능별로 도식화한 것이다. 여기서는 각 기능에 필요한 기술을 획득하는 과정부터 재생하는 마지막 단계까지 간략하게 설명한다.

3차원 비디오는 기존의 2차원 영상과 달리 사용자에게 입체감을 제공해야 하기 때문에 두 대 이상의 카

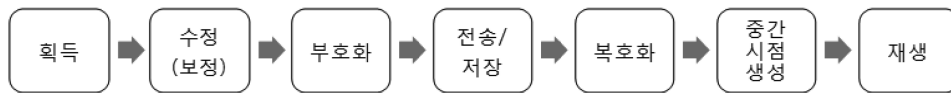


그림 9 3차원 비디오의 주요 기술

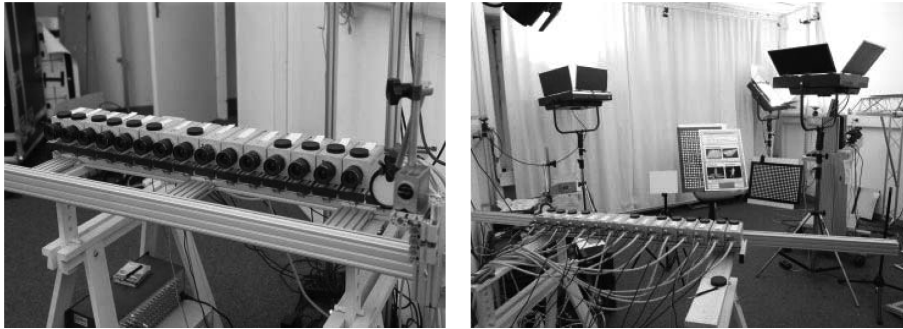


그림 10 3차원 비디오 테스트 영상의 획득을 위한 카메라 구조

메라로 영상을 획득한다. 이때, 3차원 비디오는 입체감을 느끼기 위한 최소한의 시점인 양안식 입체 화면을 기본 조건으로 한다. 그림 10은 3차원 비디오 테스트 영상의 획득을 위한 카메라 구조의 한 예인 다시점 카메라 시스템을 나타내고 있다. 그림 10과 같이 카메라의 간격은 인간의 두 눈 간격인 6.5cm와 비슷해야 한다. 또한, 보다 입체적인 화면을 제공하기 위해서는 피사체를 카메라에서 3m 이내에 두는 것이 좋고, 양 끝 카메라와 피사체가 20° 정도의 각도를 이루도록 한다.

그림 10과 같은 카메라 시스템을 이용하여 3차원 비디오를 획득한 후, 카메라 배치의 미세한 오차로 인해 발생하는 화면 간 기하학적 오차를 보상하는 카메라 정렬(rectification)을 수행하고, 시점 간 색상차를 줄이는 색상 보정(color correction)을 수행한다. 이렇게 여러 오차를 보정한 영상은 깊이 영상과 함께 다시점 비디오 부호화기에 의해 부호화된다. 앞서 언급했듯이, 다시점 비디오 부호화는 다시점 카메라로 획득한 영상을 시점 사이의 상관도를 이용하여 효율적으로 부호화한다.

3차원 비디오에서 새롭게 정의된 형식에 의해 전송된 비트스트림은 수신단에서 영상으로 복원된다. 이후 수행되는 작업은 임의의 시점의 영상을 복원하기 위한 중간영상 생성 방법이다. 다시점 카메라로 여러 시점에서 영상을 획득할 수는 있지만, 부드러운 시점 변환을 보장하기 위해서는 다시점 카메라의 시점보다 많은 시점의 영상을 생성해야 한다. 중간 시점의 영상은 영상 기반 렌더링 방법[11]을 이용하여 여러 가지 방법으로 생성할 수 있지만, 3차원 비디오의 환경에 적합한 방법으로 3차원 워핑 방법[12]이 주목을

받고 있다. 3차원 워핑은 깊이 정보를 기반으로 3차원 공간에 참조영상을 복원하고 원하는 시점으로 재투영하는 방법을 이용한다. 이 과정에서 발생하는 빈 공간(hole)이나 폐색 영역(occlusion)은 중간영상 생성 방법의 가장 큰 문제점이며, 빈 공간과 폐색 영역을 얼마나 효율적으로 처리하는지에 대해 많은 연구가 진행되고 있다.

4.4 3차원 비디오의 국제 표준화 동향

최근 3차원 비디오의 국제 표준화 그룹은 해당 기술의 탐색 실험을 위한 테스트 데이터를 요청(call for test material, CfT)하였다[13]. 또한, 다시점 카메라로 획득한 영상과 그 영상을 이용하여 중간 영상을 생성하는 영상 합성(view synthesis) 프로그램과 각 영상에 해당하는 깊이 영상(depth video)을 생성하는 프로그램을 요청하였다. 현재 3차원 비디오의 테스트 영상의 조건에 부합하는 영상에는 Microsoft 연구센터에서 제공한 'Breakdancers'와 'Ballet' 영상이 있다. 이 두 영상은 8개 시점의 색상 or 컬러 영상과 함께 각 시점에 해당하는 깊이 영상이 있기 때문에 깊이 영상의 생성 프로그램의 비교 대상으로 선택되었다.

최근 열렸던 제84차 MPEG 회의에서는 3차원 비디오 표준화 그룹의 요청에 대한 응답으로, 제시된 조건에 부합하는 3차원 비디오 테스트 영상과 다시점 카메라 교정 방법에 대한 소프트웨어, 다시점 영상을 바탕으로 각 시점의 깊이 영상을 생성할 수 있는 소프트웨어, 그리고 중간 시점의 영상을 합성할 수 있는 소프트웨어 등 다양한 기술들이 기고되었다[14]. 특히, 테스트 영상들은 테스트 데이터의 조건에 부합하도록 촬영되고 보정되었으며, 이에 따라 회의에서 기고된 테스트 자료들을 이용한 다양한 기술 검증이 가능해졌다.

5. 맺음말

본 논문에서는 3차원 비디오 부호화에 대한 일반적인 내용을 소개하고, 관련되는 기술의 표준화 동향에 대해 살펴보았다. 3차원 비디오는 사용자에게 자유시점 TV나 3차원 TV와 같은 보다 입체감 있는 영상을 사용자에게 제공할 수 있지만, 처리해야 할 데이터량이 방대하여 효율적인 부호화 및 처리 기술이 요구되는 분야이다. 현재 진행 중인 다시점 비디오 부호화에 대한 표준화 작업은 2008년까지 최종 국제 규격안(Final Draft International Standard, FDIS)이 마련될 예정이다. 이와 더불어 최근에는 다시점 비디오와 깊이 영상을 이용한 자유시점 TV에 관한 연구가 진행되고 있고, 관련 기술을 테스트하기 위한 데이터를 수집하는 단계에 이르렀기 때문에 관련 기술 개발이 더욱 가속화될 것으로 예측된다. 이 기술은 기존의 부호화 방법을 기반으로 하여 깊이 영상까지 포함하는 보다 진보된 형태의 부호화 방법을 포함한다. 또한, 수신단에서는 보다 효율적으로 입체 영상을 재생하기 위한 중간영상 생성 방법이 중점적으로 논의될 것이다. 3차원 비디오의 실용화 기술에 대한 연구는 더욱 활발히 진행될 예정이며, 자유시점 TV, 3차원 TV나 휴대 단말기 분야와 DMB 등의 사업 분야에서도 그 활용이 기대된다.

참고문헌

- [1] 호요성, 이천, "자유시점 TV와 3차원 비디오 국제 표준화 동향," TTA Journal, vol. 116, pp. 86-94, 2008.
- [2] 호요성, 오관정, "다시점 비디오 부호화," TTA Journal, vol. 115, pp. 93-100, 2008.
- [3] ISO/IEC MPEG & ITU-T VCEG, "Common Test Condition for Multiview Video Coding," JVT-U211, Oct. 2006.
- [4] ISO/IEC MPEG & ITU-T VCEG, "Joint Multiview Video Model (JMVM) 1.0," JVT-T208, July 2006.
- [5] ISO/IEC MPEG & ITU-T VCEG, "CE11: Illumination Compensation," JVT-U052, Oct. 2006.
- [6] ISO/IEC MPEG & ITU-T VCEG, "Joint Multiview Video Model (JMVM) 4.0," JVT-W207, April 2007.
- [7] ISO/IEC MPEG & ITU-T VCEG, "Adaptive Reference Filtering for MVC with decoder complexity analysis," JVT-X060, July 2007.
- [8] Droese, M., Fujii, T., and Tanimoto, M., "Ray Space Interpolation based on Filtering in Disparity Domain,"

Proc. of 3D Conference, pp. 213-216, June 2004.

- [9] Tehrani, M. P., Fujii, T., and Tanimoto, M., "Off-set Block Matching of Multi-view Images for Ray Space Interpolation," The Journal of the Institute of Image Information and Television Engineers, vol. 58, no. 4, pp. 540-548, April 2004.
- [10] ISO/IEC MPEG & ITU-T VCEG, "Multi-view video plus depth(MVD) format for advanced 3D video systems," JVT-W100, April 2007.
- [11] Shum, H.Y., Kang, S.B., "A Review of Image-based Rendering Techniques," IEEE/SPIE Visual Communications and Image Processing (VCIP), pp. 2-13, 2000.
- [12] McMillan, L., An Image-based Approach to Three-dimensional Computer Graphics, Technical Report, Ph.D. Dissertation, UNC Computer Science TR97-013, 1999.
- [13] Pereira, F. and Ebrahimi, T., "The MPEG-4 Book," Prentice Hall PTR, pp. 13-15, 2002.
- [14] ISO/IEC JTC1/SC29/WG11, "Description of Exploration Experiments in 3D Video Coding," N9783, Ar-champs, France, May 2008.



호 요 성

1981 서울대학교 전자공학과 학사
1983 서울대학교 전자공학과 석사
1989 Univ. of California, Santa Barbara, Department of Electrical and Computer Engineering, 박사
1983~1995 한국전자통신연구소 선임연구원
1990~1993 미국 Philips 연구소, Senior Research

Member

1995~현재 광주과학기술원 정보통신공학과 교수
관심분야: 디지털 신호처리, 영상신호 처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV, MPEG 표준, 3차원 TV, 실감방송
E-mail : hoyo@gist.ac.kr



이 상 범

2004 경북대학교 전자전기공학부 학사
2006 광주과학기술원 정보통신공학과 석사
2006~현재 광주과학기술원 정보통신공학과 박사과정

관심분야: 다시점 깊이맵 예측, 깊이영상 기반 렌더링, 3차원 TV, 디지털 신호처리, 영상신호처리

E-mail : sblee@gist.ac.kr