

3차원 그래픽스 압축을 위한 MPEG 표준화 기술

1. 서론

컴퓨터 그래픽스는 통상적으로 영상의 모델을 생성하고 저장하는 과정을 말한다. 가상의 카메라가 주어졌을 때 3차원 객체, 조명 소스, 조명 모델, 텍스처(texture) 등을 생성하며, 이 과정을 렌더링(rendering)이라고 한다. 3차원 그래픽스는 기하학 구조를 나타낼 때 3차원 데이터를 사용하고, 이는 2차원 그래픽스에 비해 더욱 뛰어난 입체감을 제공한다. 3차원 그래픽스는 도면 설계, 애니메이션, 게임 등의 다양한 분야에서 널리 사용되고 있다.

국제표준화 기구인 MPEG(moving picture experts group)에서는 증강현실(augmented reality, AR)에 관련된 표준화 작업을 진행하고 있으며, 세부 그룹인 3DG(3D graphics)에서 이를 주로 전담한다[1]. 최근에는 신규 기술 분야로 포인트 클라우드(point cloud)와 3차원 프린터를 위한 메타데이터(meta data) 기술에 관한 논의가 시작되었다[2].

본 논문에서는 3차원 그래픽스에 관련된 MPEG 표준화 기술을 살펴본다. 2장에서는 현재 PDAM(proposed draft amendment) 상태인 패턴 기반 3차원 메쉬 부호화(pattern-based 3D mesh coding, PB3DMC) 기술을 설명하고, 3장에서는 MPEG의 3DG(3D graphics) 그룹에서 진행하고 있는 핵심 기술(core experiment, CE)을 살펴본다. 4장에서는 3차원 원격 몰입(tele-immersion)을 소개하고, 5장에서 본 논문을 마무리한다.

2. 3차원 메쉬 부호화 (pattern-based 3D mesh coding, PB3DMC)

실제적인 응용에서 대부분의 3차원 모델은 여러 연결된 요소들로 구성된다. 이렇게 다중 연결된 3차원 모델은 다양한 변환을 통해 여러 반복적인 구조를 가진다. 이를 효율적으로 표현하기 위해서는 반복적인 구조에 존재하는 중복되는 정보를 찾아서 이를 줄여야 한다. 이 장에서는 다중 연결된 3차원 모델을 위한

효율적인 압축 방법을 설명한다[3].

2.1 3차원 모델의 반복적이며 대칭적인 구조의 식별

이 절에서는 PB3DMC에서 사용되는 기술적 용어를 정의하고, 메쉬 부호화에서 반복적으로 나타나는 구조를 식별하는 과정을 설명한다[3]. 3차원 모델에는 두 종류의 반복적인 구조가 있다. 첫 번째 구조는 연결되지 않은 반복적 구조(unconnected repetitive structure)이며, 이는 다양한 위치, 방향, 크기 등에 관계없이 변하지 않는다. 연결되지 않은 반복적 구조는 하나의 연결된 요소(component)에 해당하는 하나의 패턴을 포함한다. 두 번째 구조는 연결된 반복적인 구조(connected repetitive structure)인데, 이는 다양한 위치, 방향, 크기 등에 관계없이 변하지 않는 모든 표면 조각(surface patches)으로 구성된다. 즉, 이는 하나의 연결된 요소(component)에서 발견될 수 있는 반복적인 구조이며, 이를 대칭 구조(symmetric structure)라고도 한다.

PB3DMC에서는 다음과 같이 네 가지 단계로 반복적인 구조를 찾는다. 그림 1은 입력된 3차원 모델을 가정한 것이다.

첫 번째 단계에서는 연결되지 않은 반복적인 구조를 찾는다. 입력된 3차원 모델은 연결되지 않은 반복적인 구조와 이에 해당되지 않은 모든 구성 요소를 포함하는 고유한 부분으로 나뉘어진다. 그림 1에서 구름, 나뭇잎과 오른쪽에 있는 집은 모두 연결되지 않은 반복적인 구조를 나타낸다.



그림 1 입력 3차원 모델

두 번째 단계에서는 첫 번째 단계에서 찾은 각 연결되지 않은 반복적인 구조 내에서 하나의 기본 패턴을 선택한다. 다른 반복은 그와 관련된 기본 패턴으로 만들어진다. 따라서 입력된 3차원 모델은 다음과 같은 세 부분으로 나누어진다.

- (A) 연결되지 않은 반복적인 구조의 기본 패턴(그림 2)
- (B) 기본 패턴의 반복으로 구현된 경우(그림 3)
- (C) 연결되지 않은 반복 구조에 포함되지 않는 유일한 부분(그림 4)

그림 5에 나타낸 것처럼, (A)의 기본 패턴은 대칭성에 따라 두 종류로 더 세밀하게 나눌 수 있다. 즉, 대칭 구조를 포함하지 않는 기본 패턴 (D)와 대칭 구조를 포함하는 기본 패턴 (E)로 구분할 수 있다.

그림 6에 보인 것처럼, 연결되지 않은 반복 구조에 포함되지 않는 유일한 부분 (C)도 대칭 구조를 포함하지 않는 (F)와 대칭 구조를 포함하는 (G)로 세분할 수 있다.



그림 2 연결되지 않은 반복적인 구조의 기본 패턴 (A)



그림 3 기본 패턴의 반복으로 만들어진 경우 (B)



그림 4 연결되지 않은 반복적인 구조에 포함되지 않는 유일한 부분 (C)

세 번째 단계에서는 대칭적인 구조를 식별한다. 객체는 관련된 기본 패턴을 참조하여 부호화되기 때문에, 3차원 모델의 기본 패턴이나 고유한 부분에서 대칭적인 구조를 확인한다. 결과적으로, 그림 7에 보인 것처럼, 세 번째 단계의 입력 모델은 두 번째 단계에서 찾은 모든 반복적인 구조가 실현된 경우를 제외한다.

반복적인 구조를 찾는 두 번째 단계와 비슷하게, 네 번째 단계에서는 대칭적인 구조를 가지는 기본 패턴과 구현된 객체를 찾는다. 그림 8에 보인 것처럼, 네 종류의 데이터를 정의할 수 있다.

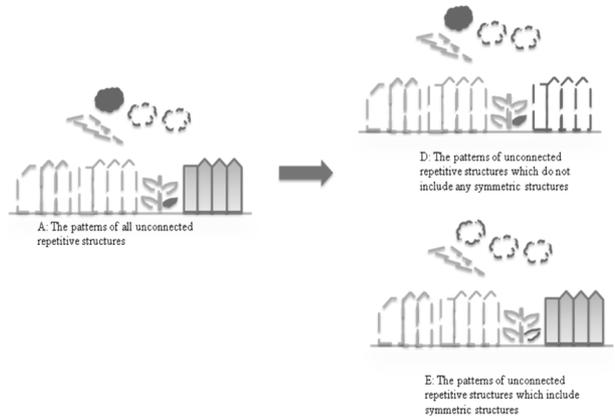


그림 5 대칭성에 따른 두 종류의 기본 패턴 (D)와 (E)

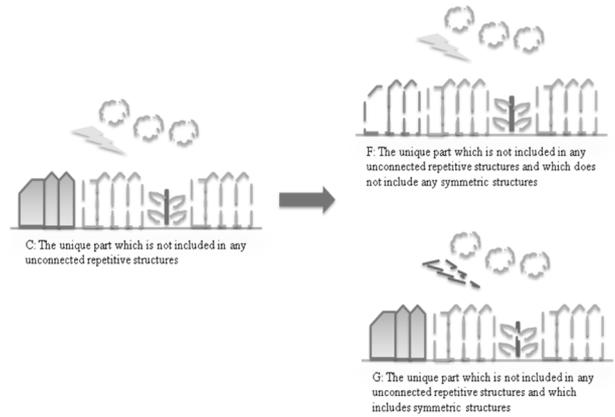


그림 6 대칭성에 따른 비반복 부분의 세분화 (F)와 (G)



그림 7 연결되지 않은 반복적인 구조의 객체를 제외한 세 번째 단계의 입력

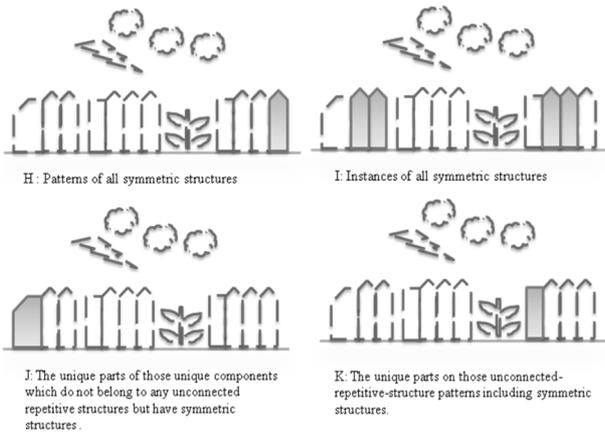


그림 8 네 번째 단계의 결과

- (H) 대칭적인 구조를 가지는 기본 패턴
- (I) 기본 패턴으로 만든 대칭적인 구조를 가지는 구현 사례
- (J) 연결되지 않은 반복적인 구조에 속하진 않지만, 대칭적인 구조를 가지는 고유한 부분
- (K) 대칭적인 구조를 포함하는 반복적인 구조의 기본 패턴에서 고유한 부분

2.2 객체 재구성 모드

전송되는 비트스트림은 구현된 모든 객체의 데이터를 포함해야 응용분야에 따라서 비트스트림의 크기, 복호화 효율, 또는 에러 내성 등을 모두 고려하였을 때의 효율성 또한 중요하다. 따라서 비트스트림에서 구현된 객체의 데이터, 즉, 기본 패턴 ID, 반사 변환(F), 전이 변환(T), 회전 변환(R), 스케일링 변환(S) 등을 복원하는 데에 두 가지 옵션이 제안되어 있는데, 이들은 모두 장단점이 존재한다[3].

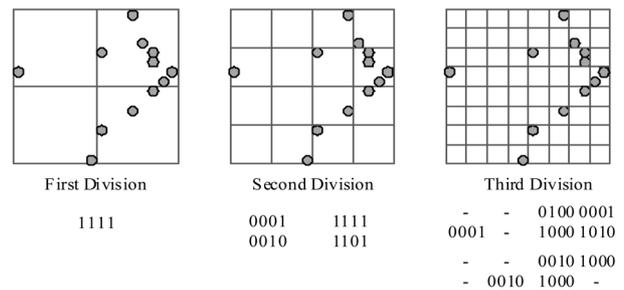
첫 번째 옵션은 기초 객체 데이터 모드(elementary instance data mode)인데, 이 모드를 사용하여 한 객체의 패턴 ID, 반사 변환, 전이 변환, 회전 변환, 스케일링 변환 정보를 ID, F, T, R, S, ID, F, T, R, S 순서로 비트스트림에 함께 묶는다. 이 모드는 복호기에서 일부 변환 손실을 복구할 수 있고, 온라인 복호화가 가능한 장점이 있다. 즉, 객체가 압축된 비트스트림을 실제 읽는 동안에, 압축된 모든 비트스트림의 판독이 완료될 때까지 기다릴 필요가 없이, 각각의 객체를 복호화할 수 있다. 또한, 빠른 처리 속도를 가지며 버퍼가 필요없는 장점이 있다. 하지만, 이 모드는 3차원 모델의 압축된 비트 크기가 큰 단점이 있다.

두 번째 옵션은 그룹 객체 데이터 모드(grouped instance data mode)로, 한 객체의 기본 패턴 ID, 반사 변환, 전이 변환, 회전 변환, 스케일링 변환 부분이 ID, ID, F,

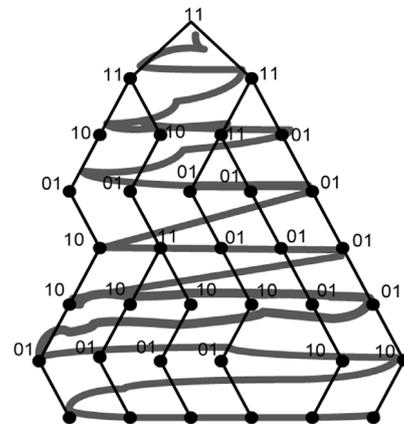
F, T, T, R, R, S, S 형태로 비트스트림에 함께 묶인다. 이 모드는 압축된 3차원 모델의 비트스트림 크기가 상대적으로 작지만, 복호기가 전송 오류에 약하고, 압축된 비트스트림을 모두 읽고 난 후에야 복호화를 시작할 수 있다. 또한, 처리 속도가 느리고, 버퍼가 필요하다는 단점이 있다

2.3 객체 변환의 복원

객체 변환은 반사, 회전, 변환, 스케일링 부분으로 이루어진다. 반사 부분을 1 비트로 표시하고, 회전 부분은 세 개의 Euler 각도로 표시한다. 변환 부분은 3축의 이동 벡터로 표현되며, 마지막으로 스케일링 부분은 객체의 균일 스케일링 인자 S로 표시한다[3].

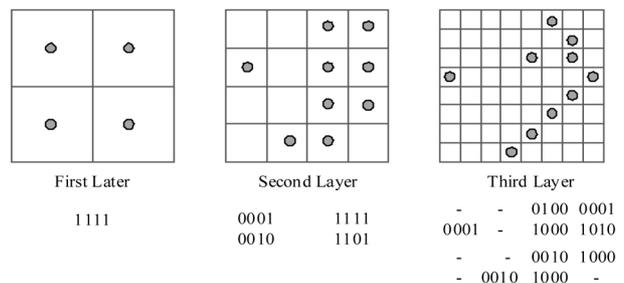


(a) 쿼드 트리 구조를 위한 공간 세분



(b) 이진 트리의 폭 우선 탐색

11110001111111010010000101000001101010000010100010000010



(c) 쿼드 트리의 재구성

그림 9 쿼드 트리 작업

그룹 객체 변환 모드를 사용하는 경우, 객체 변환 벡터는 그림 9(a)에 보인 옥트리(octree, OT) 분할을 이용하여 압축된다. 각 팔진 노드 분할은 자식 노드가 비어 있는 지 여부를 나타내기 위한 1 비트를 포함하여 총 8 비트 길이의 점유 코드로 표현된다. 옥트리 점유 코드는, 그림 9(b)에 보인 것처럼, 폭을 우선적으로 조사하여 만든다[3].

이러한 객체 전이 벡터를 복호화하려면, 그림 9(c)에 나타난 것처럼, 위에서 아래로 폭을 우선적으로 조사하여 옥트리를 복원한다. 먼저, 첫 번째 긴 점유 코드를 복호하여 맨 상단 계층을 구한다. 만약 점유 코드에 하나 이상의 "1"을 가지는 노드가 있으면, 복호기는 그 노드의 자식으로 추가하기 위해 필요한 수의 기호를 복호한다. 모든 잎(leaf) 노드가 하나의 "1"을 가지거나 단말 코드가 복호될 때까지 이 과정을 계속 수행한다. 점유 코드 시퀀스는 여러 구간으로 나뉘어서 서로 다른 확률 모델을 이용하여 복호한다. 기타 자세한 기술적인 내용은 참고문헌 [3]을 참조하기 바란다.

3. MPEG 3DG 그룹에서 진행하고 있는 핵심 기술 (core experiment, CE)

3.1 CE1: 그래픽 툴 라이브러리(graphics tool library, GTL)

CE1의 목적은 MPEG-B와 MPEG-C란 이름으로 이미 존재하는 표준인 RMC(reconfigurable media coding)의 구조에 기반을 둔 그래픽 툴 라이브러리(GTL)를 제공하는 것이다[4]. 또한 GTL 프로젝트 참가의 진입 장벽을 낮추기 위해 GTL 사용지침서를 제공하며, 세 부적으로 RMC 복호기의 설계와 구현을 돕는다. ETRI, 한양대학교, Institute Telecom 등이 CE1에 참여하고 있다.

3.2 CE3: 패턴 기반 메쉬 부호화(pattern-based 3D mesh compression, PB3DMC)

CE2는 2장에서 설명한 PB3DMC 기술에 관련하여 반사, 전이, 회전, 스케일링 등의 반복되는 패턴의 객체 변수들을 압축 부호화하는 효율적인 방법을 정하는 것이 주된 목적이다. 또한 3차원 모델의 꼭지점에 연계된 법선 벡터, 색상, 텍스처 좌표 등의 추가적인 속성들을 효율적으로 압축하는 방법을 찾는다. 세부적으로 참여기관들이 PB3DMC의 비트스트림과 부호기/복호기의 설계 및 구현을 제공한다. Technicolor, 한양대학교, Institute Telecom, UPM 등이 CE3에 참여하고 있다.

3.3 CE4: 실시간 복원을 위한 3차원 메쉬 압축

3차원 메쉬는 다수의 깊이 영상으로부터 실시간적으로 복원될 수 있다. 3차원 원격 몰입(tele-immersion)은 실제 사람과 다른 객체가 가상공간으로 몰입되게 메쉬 데이터를 실시간으로 전송할 수 있다. CE4에서는 부호화 시간, 복호화 시간, 유휴-왜곡 성능, 파일 크기 등을 평가한다. 깊이 카메라로부터 실시간으로 복원된 메쉬 데이터를 이용해 실험이 진행된다. CWI, ETRI, 한양대학교 등이 CE4에 참여하고 있다.

4. 3차원 원격 몰입(tele-immersion)

4.1 3차원 원격 몰입에 대한 요구사항(requirements)

3차원 원격 몰입은 분산된 원격지에 존재하는 사용자들이 높은 수준으로 표현된 자신의 3차원 표현을 통해 서로 상호작용을 가능하게 하는 기술이다. 이러한 시스템은 산업과 학계간의 공동연구를 통해 꾸준히 개발되었다[5, 6].

3차원 원격 몰입은 고수준의 다시점 영상회의 시스템과는 차이가 있다. 3차원 원격 몰입은 2차원 또는 3차원 메쉬 또는 점군으로 사람 객체에 대한 표현을 수집한다. 그러한 3차원 표현 방법의 효과적인 압축과 전송이 3차원 상에서 인공지능에 의해 생성된 합성 콘텐츠와의 표현과 구성을 가능하게 한다. 그림 10은 다시점의 깊이 카메라에 의해서 획득된 사용자 시나리오를

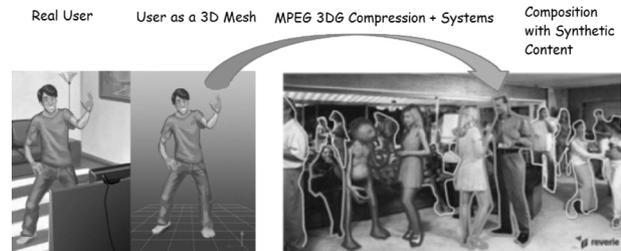


그림 10 원격으로 복원된 사용자



그림 11 실제 복원된 콘텐츠와 실제 세계에 표현된 합성 콘텐츠

나타내며, 그림 11은 실제 사람과 합성 콘텐츠를 혼합하여 실제 시스템을 구현한 것이다[7]. 이러한 시스템은 실제적인 3차원 회의와 가상세계의 통합에 대한 잠재력을 가진다. 이것은 3차원의 기하학 메쉬 표현이 사용된다는 사실 때문에 다시점, 충돌 감지 및 효과적인 렌더링(rendering)을 가능하게 한다.

4.2 네트워크 구조

이 절에서는 실제적인 응용의 가능성을 예증하기 위해서 기본적인 전송 압축 파이프라인을 설명한다. 그림 12는 전송부와 수신부의 파이프라인을 보여준다. 초기에 사람에 대한 3차원 정보가 실시간으로 획득되면, 소리, 움직임, 외형 등의 다른 정보도 획득된다. 이 시스템은 복원된 3차원 정보를 300 ms 이하의 지연시간으로 8 화면/초(frames per second, fps)까지 전송과 수신이 가능하다.

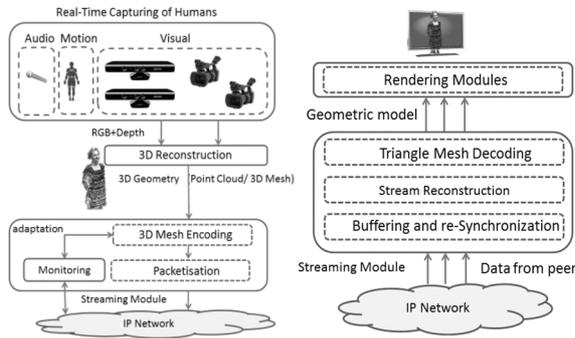


그림 12 3차원 정보를 이용한 전송 파이프라인

4.3. 파이프라인에 대한 요구사항

3차원 원격 몰입을 구현하기 위해 인터넷 상에 생성한 메쉬 구조의 실시간 스트리밍을 다루는 구체적인 해결법이 필요하다. 전체가 3차원의 메쉬 표현이라고 가정하면 그것은 이러한 점들을 가리키는 점과 면의 나열이다. 다시점과 깊이 비디오 대신에 자연스러운 가상 세계의 통합을 가능하게 해 준다.

일반적으로 낮은 응답시간은 3차원 원격 몰입에서 대역폭과 실제적인 연산의 계산으로 구성된 파이프라인의 대부분을 차지하는 중요한 요소이다. 일반적으로 300ms 이하의 응답시간이 영상회의 시스템에서 기본적인 요구사항이다. 이것은 일반적으로 그래픽스 부호화 시스템이 실시간으로 동작하지 않기 때문에 어려운 부분이다.

유연한 입출력의 표현도 중요하다. 데이터는 효과적으로 획득부에서 복원 블록으로 스트리밍 엔진을 통해 렌더러(renderer)로 지연없이 흘러야 하고, 최소

한의 파이프라인 지연 시간과 유연한 입출력을 위해 다른 스트림과 가능한 동기화를 맞춰야 한다. 이 유연성은 스트림 사이의 미디어 동기화를 위해 유용하다. 전통적으로 입출력 유연성은 그래픽스 코덱에서 비효율적이다.

시스템이 대역폭, 획득된 스트림의 수준이 갑자기 줄어드는 경우와 같은 네트워크 상황에 따라 적응할 수 있어야 한다. 또한 혼잡한 네트워크에 대해서 패킷 손실에 대한 강인함이 요구된다. 어떠한 수준의 저하가 발생할 수도 있다. 하지만 수신자 측에서 패킷 손실이 발생하는 순간에도 메쉬 구조가 복원 가능해야 한다. 또한, 메쉬 스트림은 너무 많은 대역폭을 사용해서는 안 된다. 압축의 일부 형태가 이러한 사항을 요구하는데 원본의 압축되지 않은 메쉬 구조는 한 프레임에 15 MB까지 사용될 수 있기 때문이다. 물체의 기하학적인 특성의 사전 정보는 추정 가능하다. 프로그램은 영상회의와 유사하게 계속적으로 현재 획득되고 있는 어떠한 메쉬라도 전송해야 한다. 이것은 사전에 적재된 아바타가 없다는 것을 의미한다. 그리고 일반적으로 일관성이 있는 역동적인 메쉬 정보가 다루어질 필요가 있다.

실시간으로 획득된 메쉬 구조가 지원되어야 한다. 바로 획득된 비디오와 대조적으로 장면들은 일반적으로 고정된 숫자의 점으로 구성되어 있다 (320×240, 640×480). 획득된 메쉬 구조들은 다른 개수의 점을 가질 수 있다.

4.4. 복잡한 3차원 원격 몰입 시스템의 구조

거시적인 관점에서 이러한 기술의 가능성을 보여주기 위해서 3차원 원격 몰입 시스템의 구조를 살펴본다. 3차원 원격 몰입 시스템의 구조는 현재 유럽 산업과 연구재단의 컨소시엄에 의해서 개발되고 있다. 그림 13은 전역적인 3차원 원격 몰입 시스템의 구조를 나타낸다[7].

그림 13에서 황색 모듈은 사용자의 표현과 손짓, 정적인 물체를 획득하고 복원하기 위해 사용된다. 그들은 이러한 디지털화된 표현을 얻기 위한 센서와 상호작용하는 소프트웨어 모듈이다. 디지털화된 표현은 동작 결과, 시간 종속적인 3차원 기하, 손짓 데이터 등이다. 반면에, 적색 모듈은 정적이고 동적인 물체와 함께 장면을 구성하고 표현하기 위해서 사용된다. 또한 청색 모듈은 컴퓨터로 조종하는 캐릭터, 실제와 합성 콘텐츠 사이의 상호작용의 새로운 형태를 가능하게 해주는 부가적인 확장이다. 이러한 콘텐츠의 데이터 흐름의 효과적인 의사소통이 MPEG에서 제공되는 공동으로 이용할 수 있는 방법으로 제공된다[7].

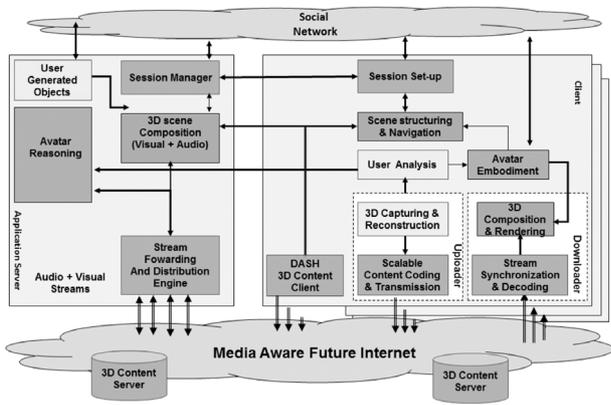


그림 13 전역적인 3차원 tele-immersion 시스템 구조

5. 맺음말

본 논문에서는 3차원 그래픽스 압축을 위한 MPEG 표준화 기술을 살펴보았다. 표준 제정 상태인 패턴 기반 3차원 메쉬 부호화 기술을 먼저 소개했고, MPEG 3DG 그룹의 핵심 기술을 설명하였다. 또한 3차원 원격 몰입에 대해서도 살펴보았다. 3차원 그래픽스 압축은 정확성있는 도면 설계나 미디어에서 수준 높은 특수 효과 등을 위해 필요한 기술이기 때문에 앞으로 더욱 발전할 것이다.

참고문헌

- [1] J. Han and S. Kim, "Text of white paper on MPEG-V," ISO/IEC JTC1/SC29/WG11, N14187, Jan. 2014.
- [2] S. Lee, J. Choi, and B. Koo, "Draft requirement and use

cases for the 3D printing metadata," ISO/IEC JTC1/SC29/WG11, M31505, Oct. 2013.

- [3] 3DG subgroup, "Study text on ISO/IEC 14496-16/PDAM4 pattern based 3D mesh compression," ISO/IEC JTC1/SC29, N14196, Jan. 2014.
- [4] 3DG subgroup, "Description of 3DG core experiments," ISO/IEC JTC1/SC29/WG11, N14195, Jan. 2014.
- [5] Requirements group, "MPEG-4 requirements, version 19," ISO/IEC JTC1/SC29/WG11, N14194, Jan. 2014.
- [6] R. Mekuria, P. Cesar, L. Bivolarski, and M. Preda, "3D tele-immersion use case and associated draft requirements," ISO/IEC JTC1/SC29/WG11, N14197, Jan. 2014.
- [7] G. Kurillo and R. Bajcsy, "3D teleimmersion for collaboration and interaction of geographically distributed users", Elsevier Virtual Reality, vol. 17, no. 1, pp 29-43, March 2013.

약 력



호요성

1981년 서울대학교 공과대학 전자공학과 졸업(학사)
 1983년 서울대학교 대학원 전자공학과 졸업(석사)
 1990년 Univ. of California, Santa Barbara, Dept. of Electrical and Computer Engineering (박사)
 1983년~1995년 한국전자통신연구소 선임연구원
 1990년~1993년 미국 Philips 연구소, Senior Research

Member

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