

합성영상 처리기술



광주과학기술원 영상통신연구실
호요성, 안정환

기존의 MPEG-1과 MPEG-2 표준은 주어진 데이터 전송량에 따라 임의의 크기로 입력되는 동영상과 음성 정보를 압축하여 생성되는 비트열을 전송한다. 그러나 MPEG-4 표준에서는, 과거의 표준들과 달리 오디오와 비디오 데이터를 화면 내에서 독립된 의미를 가지는 특정한 부분의 객체(Object) 단위로 구성할 수 있다.

그러므로 하나의 한 화면은 사람, 자동차, 나무, 하늘, 새, 구름 등과 같은 다양한 비주얼 오브젝트들(Visual Objects)의 집합으로 이루어진다.

최근 컴퓨터 그래픽에 기반한 삼차원 모델의 사용이 급증하면서 다양한 형태의 멀티미디어 서비스가 창출되고 있다. 이에 따라, 기존의 자연(Natural) 영상이나 음성 부호화에서 탈피한 삼차원 모델을 이용한 좀더 현실감 있는 영화, 게임 등의 미디어 서비스와 네트워크 상에서의 가상 환경 등의 응용이 늘어나고 있다. 따라서 MPEG-4 표준화 작업의 SNHC(Synthetic/Natural Hybrid Coding) 그룹에서는 이러한 멀티미디어 분야의 요구에 따라 가상 공간상에서 인간을 대신할 수 있는 아바타의 구현과 움직임 부호화 방법, 다각형(Polygon) 메시(Mesh) 형태로 표현되는 삼차원 모델을 효율적으로 부호화하기 위한 3DMC(3D Model Coding) 표준화를 수행하고 있다.

삼차원 물체는 일반적으로 다각형 메시 구조를 이용하여 표현하는데, 이것을 크게 기하학(Geometry) 데이터, 연결성(Connectivity) 데이터, 광도 측정(Photometry) 데이터로 분류할 수 있다. 기하학 데이터는 삼차원 물체를 구성하는 각 점들의 위치 정보로 보통 32 Bit의 유동 소수점으로 표현된다. 연결성 데이터

는 주어진 점들이 어떻게 연결되어 다각형을 구성하는지에 대한 정보를 나타낸다. 그리고 광도 측정 정보는 삼차원 모델의 표면에 렌더링(Rendering)이나 셰이딩(Shading)을 하기 위한 정보로서, 각 메시 꼭지점에 대한 색깔, 표면의 법선 벡터(Normal Vector) 성분, 이차원 영상을 매핑(Mapping)하는데 사용하는 Texture 좌표축 등의 정보를 나타낸다. 그러나 주어진 물체를 표현하는데 필요한 모델의 삼각형 수는 보통 10^5 에서 10^7 정도로 그 데이터의 양이 엄청나게 많다. 따라서 이러한 삼차원 데이터를 그대로 전송하거나 저장, 또는 렌더링하는데 많은 어려움이 따른다. 그러므로 이러한 응용에 삼차원 데이터를 효율적으로 압축하여 전송하거나 저장하는 것이 매우 필요하다.

1. 얼굴과 몸체의 움직임

MPEG-4 SNHC에서는 인공적으로 사람을 표현하기 위해 삼차원의 얼굴과 몸체의 모델을 정의하였는데, 이러한 모델은 렌더링을 위해 삼차원 다각형 메시의 형태로 이루어진다. 얼굴과 몸체의 모양, 피부와 표정은 일반적으로 얼굴 정의 변수와 몸체 정의 변수에 의해 정의된다. 그림 1은 얼굴 움직임의 이론에 기반을 두고 근육의 움직임과 밀접한 관계가 있는 가장 기본이 되는 얼굴 특징점을 보여준다. 이 점을 기반으로 각 사람에게 적합하도록 얼굴 모델을 만들고, 압축 과정을 거쳐 수신단에 전송을 한다. 이렇게 사람의 얼굴을 정의한 후, 표정이나 말에 따른 얼굴 특징점에 대한 움직임 변위를 나타내는 얼굴 움직임 변수를 전송하면 수신단은 이 변수

표 1. 16부분으로 이루어진 아바타의 몸체

관절	설명
Pelvis	골반
Waist	허리
Chest	가슴
Neck	목
Rt_up_arm	오른쪽 윗부분 팔
Rt_lower_arm	오른쪽 아랫부분 팔
Rt_hand	오른쪽 손
Lft_up_arm	왼쪽 윗부분 팔
Lft_low_arm	왼쪽 아랫부분 팔
Lft_hand	왼쪽 손
Rt_thigh	오른쪽 넓적다리
Rt_shin	오른쪽 정강이
Rt_foot	오른쪽 발
Lft_thigh	왼쪽 넓적다리
Lft_shin	왼쪽 정강이
Lft_hand	왼쪽 발



그림 2. 걸어가는 아바타

든다.

몸체의 움직임도 얼굴의 움직임과 마찬가지로, 신체 모델을 다시 설정하거나 계산할 필요 없이 관절의 움직임에 맞는 몸체 움직임 변수를 전송하면, 수신단에서는 전송된 변수에 맞게 사람의 실제적인 움직임을 표현한다. 일반적으로 관절들의 움직임 범위는 제한이 없는데, 사람 몸의 실제적인 움직임부터 간단한 사람형태의 모델을 사용하는 네트워크 게임까지 다양한 응용에 적합하도록 디자인되어야 한다. 그림 2는 삼차원 가상 공간상에서 아바타가 걸어가는 모습이다.

2. 삼차원 모델 부호화

삼차원 모델은 일반적으로 삼각형 메시(Mesh) 구조를 이용하여 표현하는데, 이것을 크게 기하학 데이터, 연결성 데이터, 광도 측정 데이터로 분류할 수 있다. 그러나 주어진 물체를 표현하는데 필요한 모델의 삼각형 수는 보통 10^5 에서 10^7 정도로 그 데이터의 양이 엄청나게 많으며, 특히 기하학 데이터는 96비트로 표현되기 때문에 정보량이 아주 크다. 따라서 이러한 삼차원 데이터를 그대로 전송하거나 저장, 또는 렌더링하는데 많은 어려움이 따른다. 그러므로 이러한 응용에 삼차원 데이터를 효율적으로 압축하여 전송하거나 저장하는 것이 필요하다. 이런 문제점을 해결하기 위해 삼차원 모델의 대부분을 차지하고 있는 기하학 정보에 대한 부호화 기

술과 삼차원 모델의 연결성 정보나 기하학 정보를 응용에 따라 다중 해상도 변이(Level-of-Detail, LOD)로 표현하는 기술이 필요하다.

기하학 정보는 삼차원 공간상에서 하나의 점에 대한 좌표값을 표현하기 위해 (X, Y, Z)의 벡터로 표현한다. 일반적으로 하나의 좌표값은 그 주변에 있는 좌표값들과 많은 상관관계를 가지고 있다. 따라서 주위에 있는 꼭지점들을 이용하여 예측 오류를 구하고, 이 예측 오류에 균일 양자화(Uniform Quantizer)를 적용하여 부호화한다. 먼저 전처리 단계에서는 입력되는 삼차원 모델을 분석하여 각각의 구성 성분의 첫번째 꼭지점을 구하고, 구성 성분별로 꼭지점들의 상관관계가 많도록 예측 순서를 정한다. 그리고 이 순서대로 예측 과정을 거쳐 예측 오류를 계산한다.

그림 3은 삼차원 모델을 서로 다른 압축 비율로 부호화한 결과이다. 그림 3(a)는 원래 모델, 그림 3(b)는 8:1로 압축한 결과이며, 그림 3(c)는 12:1로 압축한 결과이다. 결과에서 보듯이 높은 압축율에서도 원래 모델과 유사한 것을 알 수 있다.

LOD기술이란 처음에는 가장 낮은 해상도를 가진 기본 메시로 모델을 표현하고, 여기에 부가적으로 새로운 꼭지점이나 삼각형 면을 더하거나 제거하는 개량 단계를 거쳐 더 좋은 해상도를 가진 물체로 표현한다. 일반적으로 멀리 위치하고 있는 물체에는 낮은 해상도를 가진 메시 구조로 표현하고, 점차 사용자에게 가까이 있는 물체에는 개량 단계를 거쳐 높은 해상도를 가진 물체로 표현하는 기술이다.

복잡한 삼차원 모델을 단순화된 기본 메시로 만들기



(a) 원래 모델



(b) 8 : 1 압축



(c) 12:1 압축

그림 3. 기하학 압축 부호화

위해 전역 오류의 크기에 기반을 둔 메시 병합과 방법을 이용하였다. 이 방법은 삼차원 물체의 기하학적인 특성을 이용하여 제거될 수 있는 메시의 꼭지점을 찾아 제거하고 삼각형 메시를 단순화시킨다. 그리고 메시 분할과 병합 변환을 통하여 사용자나 응용에 적합하게 다양한 해상도를 가진 모델로 표현한다. 그림 4와 같은 메시 구조에서 꼭지점 v_1 이 제거될 수 있는 꼭지점이라면 이것을 제거하여 단순화된 메시 구조를 만든다. 역으로 단순화된 메시에 v_1 의 꼭지점을 첨가하여 원래의 모델을 만들며, 이와 같이 단순화된 삼차원 모델에 점진적으로 데이터를 전송하여 삼차원 모델을 점진적으로 전송할 수 있다.

그림 5는 "Bunny" 모델에 대한 다중 해상도 변이를 보여준다. 2531개의 꼭지점을 가진 원래의 모델에 메시 병합 방법을 적용하여 1156이나 748개의 임의의 꼭

지점을 가진 단순화된 모델을 만들었다. 또한 단순화된 모델에 메시 분할 방법을 적용하여 점진적으로 고 해상도를 가진 모델을 복원하였다.

3. 결론

사용자들은 여러 사람들이 함께 참가하여 상호 작용을 할 수 있는 보다 실제적인 가상 공간을 기대한다. 즉 기존 방식의 수동성에서 서로 대화하는 멀티미디어로의 발전을 요구하는 것이다. 따라서 가상 환경에 있는 멀티미디어 정보를 가지고 서로의 상호 작용을 실현하기 위해선 무엇보다도 사용자를 표현할 수 있는 아바타가 가장 중요하며, 삼차원 모델을 효율적으로 전송, 저장하는 것이 필요하다. 이같이 개발된 기술의 주된 응용

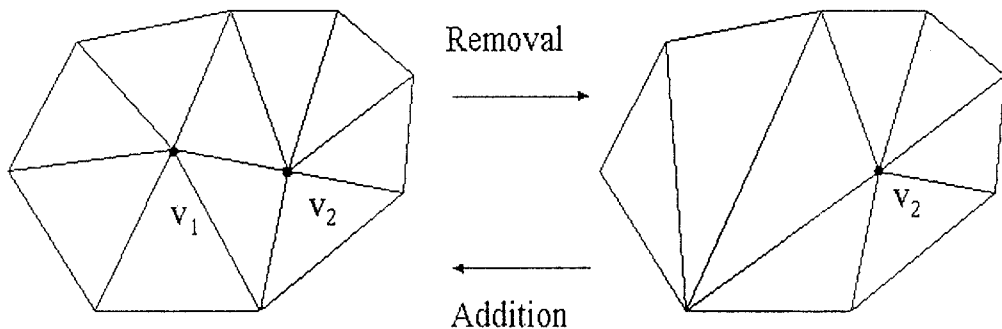


그림 4. 메시 병합과 분할 방법

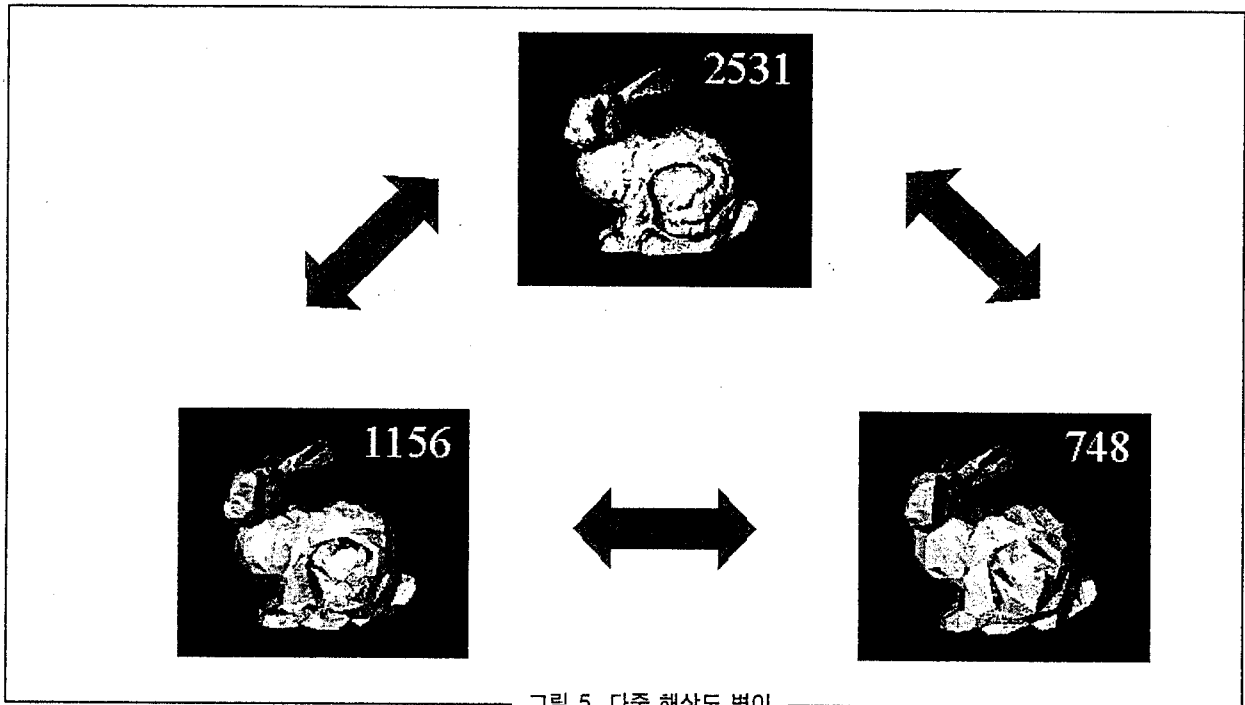


그림 5. 다중 해상도 변이

분야로는 이동체용 멀티미디어 데이터 서비스, 인터넷 서비스, 원격지 의료 및 교육 시스템, 대화형 TV, 대화

형 게임기, TTS(Text-to-Speech), 디지털 비디오 디스크(DVD) 등을 예상해 볼 수 있다.

저 | 자 | 약 | 력



호요성 (扈堯盛, Yo-Sung Ho)
E-mail : hoyo@kjst.ac.kr
1981년 02월
서울대학교 공과대학 전자공학과 (학사)
1983년 02월
서울대학교 대학원 전자공학과 (석사)

1983.03.~1995.09.
한국전자통신연구소 선임연구원

1989년 12월 University of California, Santa Barbara
Dept. of Electrical and Computer Engr. (박사)

1990.01.~1993.05. 미국 Philips 연구소 Senior Research
Member

1995.09.~현재 광주과학기술원 정보통신공학과 부교수

◆ 연구분야: 디지털 신호처리, 영상신호 처리 및 압축, 초저속 영상
통신, 디지털 TV와 고선명TV 방식, 멀티미디어 통신



안정환 (安廷桓, Jeong-Hwan Ahn)
E-mail : jhahn@kjst.ac.kr
1996년 02월
한국항공대학교 항공통신정보공학과 (학사)
1998년 02월
광주과학기술원 정보통신공학과 (석사)

1998.03.~현재
광주과학기술원 정보통신공학과 박사과정

◆ 연구분야: 3차원 영상 부호화, 디지털 신호처리, 영상신호 처
리 및 압축, 초저속 영상통신

우편번호: 500-712
광주광역시 북구 오룡동 1번지,
광주과학기술원 정보통신공학과
전화번호: 062-970-2263
FAX번호: 062-970-2204