

삼차원 애니메이션 모델의 분할 및 부호화 방법

안정환, 임동근, 호요성
광주과학기술원 정보통신공학과
광주광역시 북구 오룡동 1번지

Segmentation and Compression Techniques for 3D Animation Models

Jeong-Hwan Ahn, Dong-Keun Lim and Yo-Sung Ho
Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-dong, Puk-gu, Kwang-ju, Korea
E-mail: {jhahn,dklim,hoyo}@kjist.ac.kr

요약

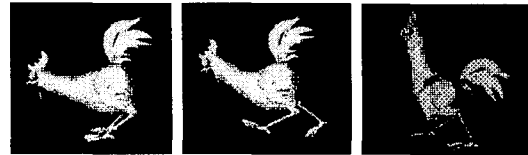
최근 복잡한 실제 사물을 가상 공간상에 표현하기 위해 삼차원 모델을 많이 이용하고 있다. 기존의 삼차원 데이터 처리는 주로 정지 모델에 대해 기하학 정보와 위상학 정보를 표현하거나 다중 해상도(Level of Details, LOD)로 나타내는데 역점을 두었다. 그러나 네트워크를 통한 가상 공간에서 삼차원 애니메이션에 대한 응용이 점차 늘어남에 따라 이러한 데이터를 효율적으로 압축하여 전송하거나 저장할 필요가 생겼다. 본 논문에서는 삼차원 애니메이션 모델의 공간적 또는 시간적 상관 관계를 이용하여 삼차원 모델 정보를 부호화하는 방법을 제안한다. 먼저 주어진 모델의 움직임을 분석하고 이를 (r, θ, ϕ) 의 구 좌표계로 변환한 후 (θ, ϕ) 의 분포에 따라 모델을 분할(Segmentation)한다. 그리고 움직임 벡터는 Affine 변환을 이용하여 삼차원 공간에서의 움직임을 정의한다. Key 프레임에 해당하는 정지 모델의 기하학 정보와 위상학 정보를 압축하고, LOD 기술을 적용하여 손실 혹은 무손실로 부호화하여 전송한다. 또한 Key 프레임 사이의 화면에서는 선형 또는 비선형 보간법으로 각 분할 부분을 복원하고, 이를 조합하여 전체적인 삼차원 모델을 복원한다.

1. 서론

최근 멀티미디어 서비스의 요구가 급격히 증가하면서 가상 공간상에서의 삼차원 애니메이션을 이용한 응용[1]에 많은 관심이 쏠리고 있다. 일반적으로 삼차원 물체는 폴리곤(Polygon) 구조를 이용하여 나타내는데, 이것은 연결성 정보, 기하학 정보, 색상, 법선 벡터, 그리고 텍스처 정보 등으로 이루어진다. 연결성 정보는 폴리곤들이 어떻게 연결되는 지에 대한 정보를 나타내며, 기하학 정보는 폴리곤의 각 꼭지점 좌표값을 가지고 있다. 그리고 색상, 법선 벡터, 텍스처 정보는 삼차원 물체의 표면에 색을 입히거나 렌더링하는데 필요한

정보를 가지고 있다. 동영상 비디오와 같은 스트리밍(Streaming) 미디어 형태의 애니메이션에서는 이러한 각 프레임의 모델 정보를 이용한다.

그림 1은 CHICKEN 모델이며, 3030 개의 꼭지점, 5665 개의 폴리곤, 400 프레임으로 구성된 애니메이션이다. 이러한 삼차원 애니메이션 모델의 정보를 저장하기 위해서는 $14,577,900 \text{ Bytes} (3030 \text{ vertices} \times 400 \text{ frame} \times 3 \text{ coords/vertex} \times 4 \text{ bytes/coord} + 5665 \text{ triangles} \times 3 \text{ indices/triangle} \times 2 \text{ bytes/index})$ 가 필요하다.



(a) 프레임 1 (b) 프레임 100 (c) 프레임 200

그림 1. 애니메이션 모델 (CHICKEN Model)

이렇게 애니메이션 모델을 표현하기 위해서는 기존의 삼차원 모델의 다중 해상도 표현과 기하학 또는 위상학 정보의 부호화 방법[2,3] 뿐만 아니라, 애니메이션 모델의 부호화 방법도 필요하다. 본 논문에서는 삼차원 공간상에서의 움직임을 이용하여 애니메이션 모델을 효율적으로 부호화하는 방법을 제안한다.

그림 2는 삼차원 애니메이션 모델을 부호화하는 방법의 개략적인 블록도이다. 먼저 애니메이션 전체 프레임에서 중심이 되는 Key 프레임을 정의하고, 움직임이 비슷한 부분을 찾아 분할(Segmentation)한다. 이때 (x, y, z) 공간상에서의 움직임 벡터를 (r, θ, ϕ) 의 구 좌표계로 변환하고, (θ, ϕ) 의 분포에 따라 모델을 분할한다. 그리고 움직임 벡터는 Affine 변환을 이용하여 삼차원 공간상에서의 움직임을 정의한다. Key 프레임에 해당하는 정지(Static) 모델은 기하학 정보와 위상학 정보를 압축

하고 LOD 기술을 적용하여 손실 혹은 무손실로 부호화하여 전송한다. 그리고 Key 프레임 사이의 화면은 선형 혹은 비선형 보간법을 이용하여 각 분할 부분을 복원하고, 이것을 조합하여 전체적인 삼차원 모델을 복원한다.

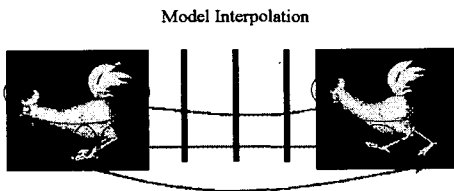
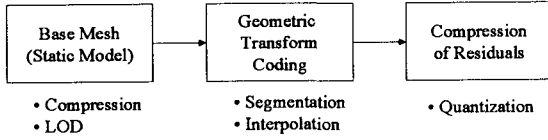


그림 2. 삼차원 애니메이션 부호화의 개략도

2. 삼차원 애니메이션 모델의 분할

본 논문에서는 주어진 삼차원 애니메이션 모델을 분할하기 위해 Spatial Clustering 방법, Topology 에 기반한 방법, 움직임에 기반한 방법을 사용하였다.

2.1 Spatial Clustering 방법

Spatial Clustering 방법에서는 그림 3 과 같이 주어진 C 개의 Polygon 을 N 개의 Cluster 에 적당하게 할당한다. 본 논문에서는 먼저 각 폴리곤의 중심점을 구하고, 각 중심점과 Cluster 의 중심점과의 거리가 최소가 되는 폴리곤을 하나의 Cluster 에 할당한다.

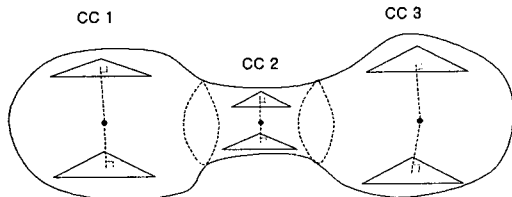


그림 3. Clustering 방법을 이용한 분할 방법

2.2 Topology 에 기반한 분할 방법

Topology 에 기반한 방법에서는, 그림 4 와 같이, 임의의 폴리곤 중에서 하나의 꼭지점을 선택하고, 그 꼭지점을 공유하는 주변의 폴리곤을 시계 방향 혹은 반시계 방향으로 탐색하면서 Decomposition 했을 때 분리되는 것들을 하나의 단위로 간주한다 [3].

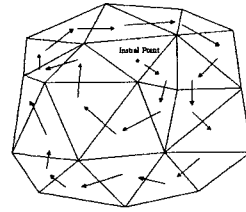


그림 4. Orange-Peeling Decomposition

2.3 움직임에 기반한 분할 방법

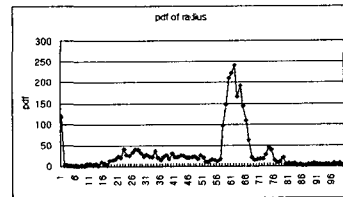
애니메이션 모델의 프레임 사이에 존재하는 시간적인 상관 관계를 이용하여 애니메이션 모델을 분할한다. 먼저 Key 프레임을 정하고, 각 꼭지점에 대한 움직임 벡터 (MV_x, MV_y, MV_z) 를 구하여 식 (1)과 같이 좌표계로 표현한다.

$$MV_r = \sqrt{MV_x^2 + MV_y^2 + MV_z^2}$$

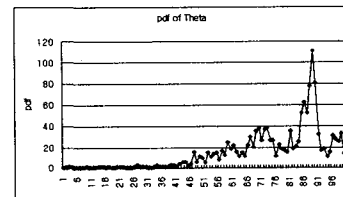
$$MV_\theta = \tan^{-1}(MV_y / MV_x) \quad (1)$$

$$MV_\phi = \cos^{-1}(MV_z / r)$$

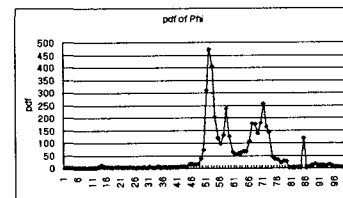
그림 5 는 첫번째 프레임과 10번째 프레임 사이의 (r, θ, ϕ) 움직임 벡터 분포를 나타낸 것이다. 이때 MV_θ 와 MV_ϕ 는 $0 \leq MV_\theta \leq \pi/2$, $-\pi \leq MV_\phi \leq \pi$ 의 범위이다.



(a) Radius MV_r 의 분포



(b) MV_θ 의 분포



(c) MV_ϕ 의 분포

그림 5. 움직임 벡터 $(MV_r, MV_\theta, MV_\phi)$ 의 분포

그림 6은 삼차원 공간상에서 회전 운동을 하는 모델을 보여준다.

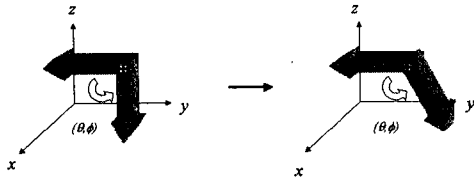


그림 6. 삼차원 좌표계에서의 움직임

이때 움직임 벡터 (MV_x, MV_y, MV_z) 를 구 좌표계로 변환하여 살펴보면, 그림 7에 보인 것처럼, 비슷한 운동을 하는 부분은 어느 일정한 부분에 몰려 있다. 따라서 같은 움직임을 갖고 있는 부분을 찾아서 의미있는 움직임 단위로 Grouping 할 수 있다.

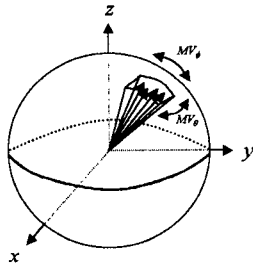
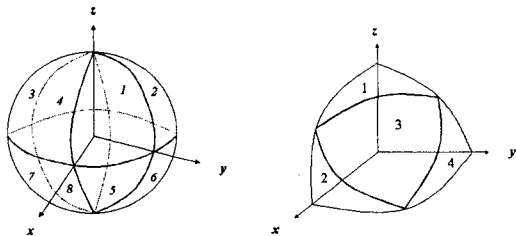


그림 7. 비슷한 움직임을 갖는 (MV_x, MV_y, MV_z) 분포

움직임을 좀더 세분화하여 분석하기 위해 그림 8이나 그림 9와 같이 일정한 면적을 갖도록 구 표면을 분할하여, 구 표면의 영역을 나눈다.

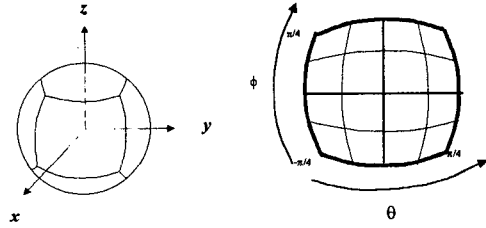
그림 8에 보인 8-4 분할[2]은 처음에 단위 원을 8등분으로 나누고, 각 면이 같은 면적을 갖도록 4개의 영역으로 계속 나누는 방법이다. 반면, 그림 9에 보인 6-4 분할은 처음에 구 표면을 정육면체로 생각하여 6등분으로 나누고, 각 면이 같은 면적을 갖도록 계속해서 일정하게 4개의 영역으로 나누는 방법이다. 각 영역의 실제 중심점을 전송하는 대신에 인덱스(Index)만을 부호화하여 전송해도 수신단에서는 부호화기와 같은 방법을 사용하여 중심점을 복원할 수 있다.



(a) 8개 분할

(b) 4개 분할

그림 8. 구 표면의 8-4 분할 방법



(a) 6개 분할

(b) 4개 분할

그림 9. 구 표면의 6-4 분할 방법

이렇게 전송된 움직임 벡터는 각 분할 단위를 대표하는 값이므로, 수신단에서는 같은 분할 단위에 속해 있는 폴리곤에 대해서는 전송된 움직임 벡터를 적용하여 삼차원 모델을 애니메이션할 수 있다.

3. 삼차원 모델의 보간 방법

주어진 삼차원 모델을 움직임에 따라 분할하고, 분할된 부분에 해당하는 대표 움직임 벡터와 Key 프레임 모델을 전송하면, 수신단에서는 전송된 정보를 이용하여 Key 프레임 사이에 해당하는 모델들을 복원할 수 있다. 본 논문에서는 식(2)와 같이 12개의 변수를 가지고 있는 Affine 변환식으로 움직임 벡터를 모델링하였다.

$$\begin{pmatrix} \hat{v}_x(x, y, z) \\ \hat{v}_y(x, y, z) \\ \hat{v}_z(x, y, z) \end{pmatrix} = \begin{pmatrix} x'-x \\ y'-y \\ z'-z \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_5 & a_6 & a_7 \\ a_9 & a_{10} & a_{11} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} a_4 \\ a_8 \\ a_{12} \end{pmatrix} \quad (2)$$

여기서 (x, y, z) 는 이전 key 프레임에서 각 삼각형 메시에 해당하는 기하학 정보값이고, (x', y', z') 현재 프레임에서의 값이다. 각 메시의 움직임 벡터 (v_x, v_y, v_z) 는 두 Key 프레임을 이용하여 구할 수 있기 때문에, 식 (3)과 같이 오류 식을 정의할 수 있다. 여기서 N 은 분할 개수이다.

$$E(a) = \sum_{i=1}^N \{ [v_x(x_i, y_i, z_i) - \hat{v}_x(x_i, y_i, z_i)]^2 + [v_y(x_i, y_i, z_i) - \hat{v}_y(x_i, y_i, z_i)]^2 + [v_z(x_i, y_i, z_i) - \hat{v}_z(x_i, y_i, z_i)]^2 \} \quad (3)$$

따라서 식(3)에 식(2)을 대입하면 식(4)와 같이 되고 여기에 Least Square 방법을 적용하면 12개의 Affine 변수를 구할 수 있다.

$$E(a) = \sum_{i=1}^N \{ [v_x(x_i, y_i, z_i) - (a_1x + a_2y + a_3z + a_4)]^2 + [v_y(x_i, y_i, z_i) - (a_5x + a_6y + a_7z + a_8)]^2 + [v_z(x_i, y_i, z_i) - (a_9x + a_{10}y + a_{11}z + a_{12})]^2 \} \quad (4)$$

4. 실험 결과

본 논문에서 제안한 삼차원 애니메이션 모델의 분할과 부호화 방법을 실험하기 위해 앞에서 언급한 CHICKEN 모델을 사용하였으며, 전체 400 프레임 중에서 첫번째 프레임과 10 번째 프레임에 대해서 살펴보았다.

4.1 분할 결과

우선 첫번째 프레임에 Spatial Clustering 방법을 적용하여 2개와 4개로 분할한 결과를 그림 10에 제시하였다. 그림 10에서 보듯이, 공간적으로 나눈 각 단위는 어떤 의미있는 부분이라기보다는, 근처에 있는 폴리곤을 하나의 집합으로 단순하게 묶었음을 알 수 있다. 따라서 하나의 분할 단위내에 서로 다른 움직임을 갖는 부분이 있다면 정확한 움직임을 정의할 수 없으므로, 수신단에서는 주어진 정보로 부호화한 원래의 모델을 복원하기 어렵다.

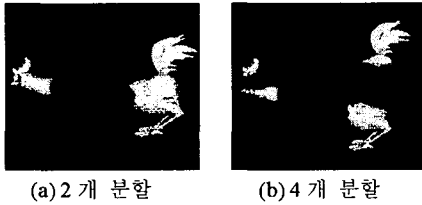


그림 10. Spatial Clustering 방법을 적용한 분할

그림 11은 첫번째 프레임에 Topology 특성을 적용하여 31개의 단위로 분할한 결과이다. 이 방법도, Spatial Clustering 방법과 마찬가지로, 시간적으로 어떤 의미있는 단위로 분할하기 어렵다.



그림 11. Topology 특성을 이용한 분할 (회색 명암)

그림 12는 첫번째 프레임과 10 번째 프레임 사이의 움직임 정보를 구 좌표계로 변환하고, 비슷한 움직임을 갖는 부분을 4개와 8개의 그룹으로 묶은 결과이다.

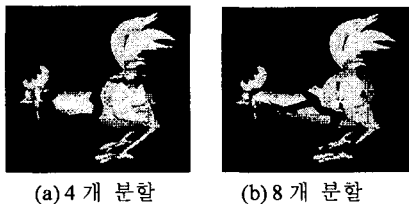


그림 12. 움직임에 기반한 분할

CHICKEN 모델은 그림 1에 나타난 것처럼 머리를 위로 올리면서 한쪽 다리를 들어올리는 운동을 한다. 이때 움직임 벡터의 분포는 그림 5에 나타내었다. 그림 13에서 보듯이, 4개의 부분으로 나누었을 경우에는 머리, 목, 몸통, 꼬리 부분으로 나뉘어지며, 8개로 분할할 경우에는 몸통, 머리, 다리 부분이 좀더 자세하게 분할됨을 알 수 있다.

4.2 모델 보간 결과

본 실험에서는 움직임에 기반하여 첫번째 프레임의 모델을 32개로 분할하고, 각 영역에 대한 대표 움직임 벡터를 구해서 10 번째 프레임에 해당하는 모델을 복원하였다. 그림 13은 앞에서 제안한 방법을 사용하여 얻은 결과이다. 원래의 모델과 비교하여 복원 모델에 대해 시각적으로 큰 차이가 없음을 알 수 있다.



(a) 원래의 모델 (b) 복원한 모델

그림 13. 보간법을 사용해 복원한 모델

5. 결론

본 논문에서는 구 좌표계에서 움직임에 기반하여 삼차원 애니메이션 모델을 분할하고 보간하는 방법을 제안하였다. 실험 결과에서 보인 것과 같이, 낮은 비트율에서도 원래 모델과 비교하여 시각적인 큰 차이없이 효과적으로 삼차원 애니메이션 모델을 복원할 수 있다.

6. 감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속망네트워크 연구센터(UFON)를 통한 한국과학재단 우수연구센터(ERC)와 교육부 두뇌한국 21(BK21) 정보기술사업단의 지원에 의한 것입니다.

7. 참고문헌

- [1] MPEG-4 SNHC, "MPEG-4 SNHC Verification Model 9.0," ISO/IEC JTC1/SC29WG11 MPEG98/M3809, July 1998.
- [2] J.H. Ahn and Y.S. Ho, "An Efficient Geometry Compression Method for 3D Objects in The Spherical Coordinate System," Int'l Conference on Image Processing (ICIP), pp. II.482-486, Oct. 1999.
- [3] G. Taubin, W. Horn, F. Lazarus and J. Rossignac, "Geometric Coding and VRML," ISO/IEC JTC1/SC29/WG11 MPEG98/M3061, Feb. 1998.