

영상기반 모델러의 설계 및 구현

박종승^o, 이민석, 구분기, 오광만
ETRI 가상현실연구개발센터 영상기반렌더링팀

Design and Implementation of an Image-based Modeler

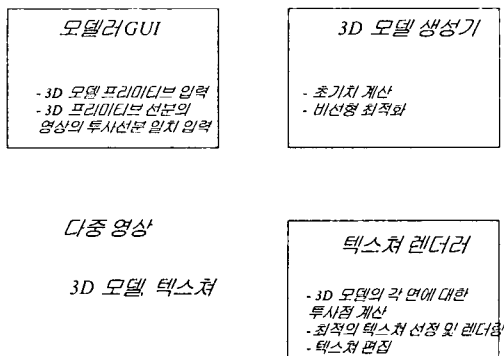
Jong-Seung Park, Min-Suk Lee, Bon-Ki Koo, Kwang-Man Oh
Image-based Rendering Team, Virtual Reality Research Center, ETRI

요 약

다중영상으로부터 3D 모델을 생성하고 각 면의 texture를 렌더링하는 영상기반 modeler를 설계 및 구현하였다. 본 논문에서는 영상기반 modeler의 개발 과정에서의 기술적인 문제점과 해결책을 기술한다.

1. 서론

본 논문에서는 다중영상으로부터 3차원 모델을 생성하고 texture를 렌더링하는 ETRI 영상기반 모델링 시스템의 모델링 기법에 대해서 설명한다. ETRI 영상기반 modeler는 웹 기반 가상공간 구축을 위한 저작도구의 성격으로 개발되고 있다.



[그림 1] 영상기반 modeler의 흐름도

세계적으로 여러 상용제품이 출시되고 있는 실

정이다. 대표적인 제품으로 Evans & Sutherland의 RAPIDsite™

[4], KAON의 HyperReality™[5], MetaCreations사의 Canoma™[6], EOS Systems사의 PhotoModeler™[7] 등이 있다. 이들 제품은 *사용의 편리성과 모델의 정확성의 trade-off*에 따라서 특징이 조금씩 다르다. 사용의 편리성과 모델의 정확성의 두 가지 요소를 동시에 만족시키는 정도가 제품의 질을 결정한다. 본 논문에서 기술하는 시스템의 개발방향은 두 가지 요소를 동시에 만족시키는 것이다.

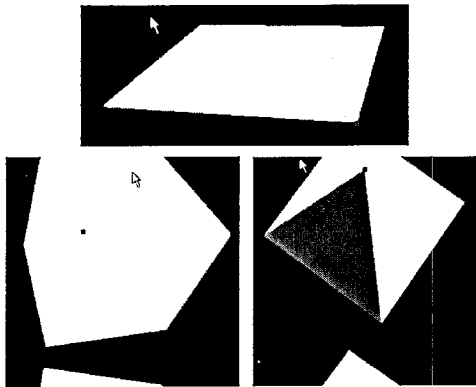
영상기반 modeler의 전체 시스템 흐름도는

[그림 1]과 같다. modeler GUI는 사용의 편리성에 중점을 둔다. 이를 위해서 3D primitive들을 이용하여 쉽고 빠르게 3D 모델을 구축하도록 한다. 사용자는 3D primitive를 선택한 후, 주어진 다중영상에서 해당하는 선분을 지정한다.

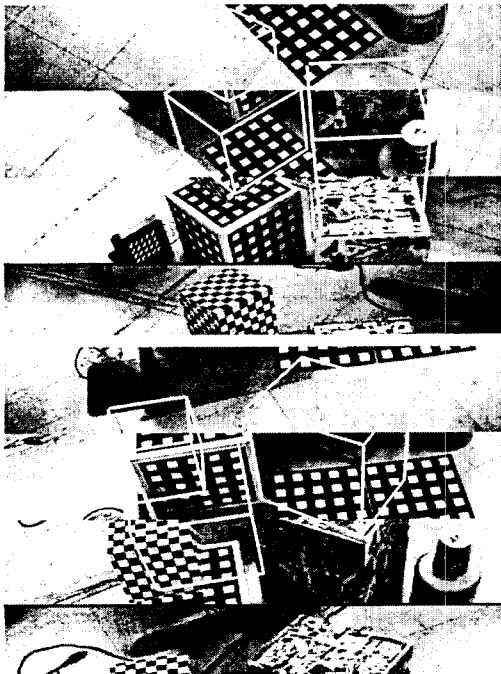
3D 모델 생성기에서는 주어진 영상 선분으로부터 실제의 3D 선분을 추정한다. 매우 복잡한 수치적인 계산 절차가 요구된다. 이 모듈은 모델의 정확성에 중점을 둔다.

Texture 렌더러는 3D 모델의 각 면에 대한 texture를 생성한다. 시각적으로 가장 좋은 texture를 다중영상으로부터 선택 및 편집한다.

2. 3D 모델의 생성



[그림 2] 3D primitive들; plate, cube, pyramid



[그림 3] 영상 선분들은 GUI를 통한 사용자의 수동 입력으로 얻는다.

다중 영상에서 3D 모델을 복원하기 위해서는 복잡한 컴퓨터 비전의 기법이 필요하다. 영상에서의 특징값으로서 선분을 사용한다.

[그림 2]와 같이 미리 정의된 3D primitive들을 선택한다. 선택된 primitive들의 선분에 해당하는 영상 선분들을 [그림 3]에서와 같이 사용자의 수동입력으로 얻는다.

선분으로부터 S&M (Structure and Motion)을 얻는 방법은 Taylor와 Kriegman의 방법[1]을 따른다. 자세한 기법은 다음 절에 기술되어 있다.

2.1 선분을 이용한 3차원 복원 - Taylor와 Kriegman의 방법[1]

다중 영상에서 선분으로부터 S&M (Structure and Motion)을 얻는 방법은 Taylor와 Kriegman의 방법[1]을 따른다. m 개의 영상이 있고 이들이 각각 n 개의 직선을 포함하고 있다고 하자. 영상을 위한 인덱스로 j ($1 \leq j \leq m$)를 사용하고, 직선을 위한 인덱스로 i ($1 \leq i \leq n$)를 사용하자. 영상의 생성 과정은 $F(p_i, q_i)$ 로 표현된다. 여기서 p_i 는 3차원 직선 i 의 위치이고 q_i 는 영상 j 의 카메라 위치와 방향이다. u_{ij} 를 직선 i 를 영상 j 로 투사한 것의 측정치라고 하자. 목적함수는 $F(p_i, q_i)$ 와 u_{ij} 의 disparity의 합으로 둘 수 있다. 3차원 복원의 목표는 주어진 영상 측정치로부터 목적함수를 최소화하는 p_i 와 q_i 를 찾는 것이다. 직선 i 의 측정된 선분 $\langle (x_1, y_1), \langle (x_2, y_2) \rangle$ 과 예측된 선분 m 의 에러 Err_i 는 두 선분의 disparity로 정의된다:

$$Err_i = m^T (A^T B A) m, \text{ 여기서,}$$

$$m = (m_x, m_y, m_z)^T, l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

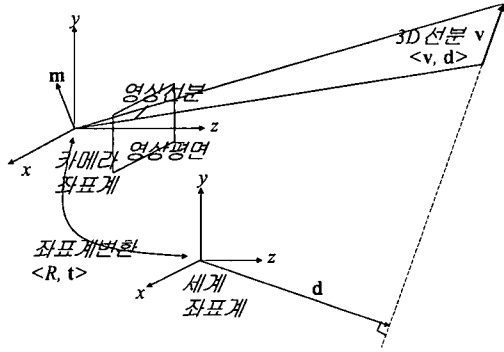
$$A = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix}, B = \frac{l}{3(m_x^2 + m_y^2)} \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

이다. 즉 목적함수는

$$O = \sum_{i=1}^n Err_i$$

이다. 위의 함수는 미지수에 대해 비선형으로서, 초기값에 의존되는 국부최소치가 해가 될 수 있다. 올바른 해를 얻기 위해서는 적당한 초기치를 구해야 한다.

카메라 위치 q_i 는 이동 벡터 t_i 와 회전 행렬 R_i 로 표현된다. 문제의 간결성을 위해서 기준좌표를 첫번째 카메라 위치로 둔다. 또한 카메라의 초점거리를 모두 1로 둔다.



[그림 4] 한 3차원 직선은 $\langle v, d \rangle$ 로 표현된다.

직선은 직선의 방향을 나타내는 단위 벡터 v 와 원점과 가장 가까이 위치한 직선상의 점에 대한 위치 벡터 d 의 쌍 $\langle v, d \rangle$ 으로 표현한다.

3차원 직선 i 와 영상 j 의 카메라 중심이 이루는 평면의 노말을 m_{ij} 으로 하자. 실제로 세계 좌표를 영상 j 의 카메라 좌표로 변환은 R_j, t_j 로 이루어 진다고 하자. 직선 i 는 카메라 좌표로 $\langle v_i, d_i \rangle$ 로 표현되고 세계 좌표로는 $\langle v, d \rangle$ 로 표현된다고 하자. 그러면 카메라 좌표계로 표현되는 m_{ij} 은 $m_{ij} = v_i \times d_i$ 이다. 세계좌표로 표현되는 직선의 벡터를 이용하면

$$m_{ij} = R_j(v_i \times (d_i - t_j))$$

이다. 따라서,

$$m_{ij}^T R_j v_i = 0 \quad (1)$$

$$m_{ij}^T R_j (d_i - t_j) = 0 \quad (2)$$

의 두 제약식을 얻을 수 있다.

영상의 측정치 u_{ij} 를 $\langle (x_1, y_1), (x_2, y_2) \rangle$ 라고 하자. 카메라 중심과 측정 에지를 지나는 평면의 normal은

$$m'_{ij} = (x_1, y_1, 1)^T \times (x_2, y_2, 1)^T$$

이다.

식 (1)로부터 목적함수를 고안한다.

$$O_1 = \sum_{j=1}^m \sum_{i=1}^n (m'_{ij}{}^T R_j v_i)^2$$

즉, 카메라 위치와 직선의 위치에 무관하게 카메라 방향과 직선의 방향을 추정할 수 있다. 위의 목적함수는 $O_1 = \sum_{i=1}^n O_{1i}$, 여기서

$$O_{1i} = \sum_{j=1}^m (m'_{ij}{}^T R_j v_i)^2$$

으로 분해가능하다. 각 항이 서로 다른 v_i 를 가지므로 각 O_{1i} 의 최소해를 결정하면 O_1 의 최소해를 얻는다. O_{1i} 의 해는 SVD[2]로 얻을 수 있다.

일단 R_j, v_i 가 얻어지면, 식 (2)를 이용하여, d_i, t_j 를 추정한다.

$$O_2 = \sum_{j=1}^m \sum_{i=1}^n (m'_{ij}{}^T R_j (d_i - t_j))^2$$

O_{1i} 와 O_2 의 최소화는 LM 비선형 최적화 방법[2]을 사용한다.

과정 1: 목적함수 O_1 를 사용하여 카메라 방향과 직선의 방향을 추정한다.

과정 2: 목적함수 O_2 를 사용하여 카메라 위치와 직선의 위치를 추정한다.

과정 3: 목적함수 O 를 사용하여 비선형 최적화한다.

과정 1에서 올바른 해를 구하기 위한 방법으로 Taylor와 Kriegman[1]는 다양한 회전행렬을 임의로 생성하고 이들을 시도해보는 방법을 사용하였다. Debevec 등[3]은 건물 등의 구조물의 알려진 수직 및 수평 방향 직선을 이용하여 직선의 방향을 고정시킨 뒤에 회전행렬만을 먼저 고려하였다.

Taylor와 Kriegman의 방법에서는 수많은 임의의 회전행렬을 초기값으로 생성함에도 불구하고 올바른 해를 찾지 못하는 경우가 발생할 수 있다. Debevec 등의 경우에는 알려진 방향의 직선을 가정하는 것이 일반적인 영상에서는 invalid하다.

우리는 주어진 세계좌표계에 대해서 적당한 카메라 위치와 방향을 GUI를 통한 그래픽 입력으로 제공한다. 이 힌트는 탐색 영역을 크게 줄일 수 있고 국부 최소값의 위험을 제거할 수 있다.

3. Texture 렌더링

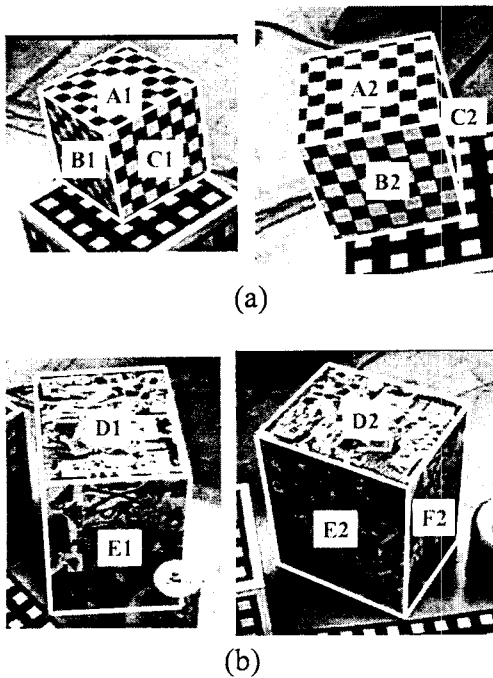
3차원 모델이 구해지면 각 면에 대한 texture를 결정한다. 3차원 모델의 각 면에 해당하는 부분의 영상에서의 영역을 알고 있다. 다중영상으로부터 중복되는 texture를 처리해야 하므로, 다음과 같은 기능의 모듈이 필요하다:

- texture의 품질을 결정하는 기준 주어진 texture에 대해서 얼마나 왜곡이 적고 넓게 모델의 면을 덮을 수 있는가의 기준으로서, 정사각형/정삼각형에서의 뒤틀림 정도에 반비례하고, 내부에 포함하는 화소수

에 비례하는 축도이다.

- 칼라 정규화 기능을 포함한 texture blending 기능. 각 texture의 칼라 히스토그램을 이용한다. 품질이 최상인 texture의 인접 영역의 히스토그램을 기준으로 변환 대상 texture의 히스토그램을 선형 변환한다.

- texture가 부재한 면에 대한 가상 texture 생성 기능. primitive의 대칭성을 이용하여 대칭면의 texture를 사용한다.

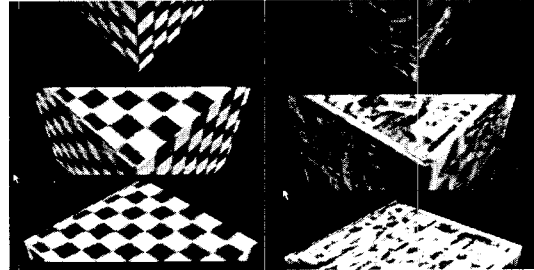


[그림 5] 여러가지 선정 기준에 의하여 다중영상으로부터 최상의 texture를 얻는다.

[그림 5]의 예를 보자. (a)의 경우에서, A면은 A1과 A2 중에서 품질이 우수한 것이 texture로 선정된다. B와 C의 경우도 동일하다. (b)의 경우에서, D면은 품질이 우수한 것이 texture로 선정된다. E면은 부분적으로 가려진 E1이 선정될 수도 있다. 이 경우, 후처리 texture 편집기의 사용이 요구된다. F면은 F2만 있으므로 이를 선정한다.

각 영상은 획득시의 조명조건에 따라서 칼라차를 가진다. texture가 모두 선정된 다음, 이를 사용하여 rendering을 하면 칼라차로 인해서 매끄럽지 못한 결과를 얻게 된다. 각 선정된 texture를 칼라 정규화 처리를 하여 이러한 문제를 해결한다.

Texture 렌더링 결과. [그림 6]에 있다. B면과 F면의 경우 texture의 품질이 떨어진다.



[그림 6] Texture 렌더링 결과

4. 결론

본 논문에서는 다중영상으로부터 3D 모델을 생성하고 모델의 면에 대한 texture를 생성하는 영상기반 modeler에 대한 기술적인 측면을 다루었다.

3D 모델 생성기의 초적 해를 찾는 과정은 여전히 개선의 여지가 있다. Texture 렌더링 과정에서 에지 매칭을 통한 여러 texture의 blending 연구가 필요하다.

[참고문헌]

- [1] Camillo J. Taylor and David J. Kriegman, Structure and Motion from Line Segments in Multiple Images, IEEE Trans. On PAMI, Vol. 17, No. 11, Nov. 1995, pp.1021-1032
- [2] 박종승, 한준희, 컴퓨터 비전에서의 최소화 기법의 고찰과 성능비교, 한국정보과학회 '98 가을 학술대회, 1998
- [3] E. Debevec and Camillo J. Taylor and Jitendra Malik, Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach, Proceedings of SIGGRAPH 96, pp. 11-20.
- [4] <http://www.es-rapidsite.com/>
- [5] <http://www.kaon.com/>
- [6] <http://www.metacreations.com/>
- [7] <http://www.photomodeler.com/>