

물체의 3차원 모델데이터 획득 및 VRML 표현

엄성훈 · 이종수
컴퓨터정보통신공학부

<요약>

본 논문에서는 Porsche Toy Car의 2차원 영상들로부터 3차원 모델 데이터를 획득하는 실험과 이 실험을 통해 획득한 DXF 파일을 VRML 2.0 버전 파일로의 변환에 대해 기술하고 있다. 2차원 영상들에 의한 3차원 모델링 기술은 물체에 관한 모든 정보를 입력 영상에 의존하기 때문에 최적의 영상획득을 위한 카메라 위치 설정 기술을 먼저 다룬다. 영상들은 획득시의 카메라 위치 정보와 함께 3차원 모델링 도구인 PhotoModeler에 입력되며 PhotoModeler는 사용자가 지정하는 특징점들을 사용하여 물체의 3차원 정보를 추출한다. PhotoModeler로 3차원 정보를 추출하는 과정과 관련된 기술들과 이론들도 요약 설명하고, Porsche Toy Car의 영상들로부터 3차원 모델 데이터를 추출하는 실험과정도 기술한다. 획득한 3차원 모델 데이터가 일반적으로 DXF 파일 형식인데 이를 VRML 2.0 버전 파일 형식으로 변환시키는 기법도 제시한다.

Three Dimensional Model Data Acquisition and VRML Representation

Seong-Hoon Eom · Jong-Soo Lee
Division of Computer Engineering & Information Technology

<Abstract>

This paper describes three dimensional model data acquisition experiments from multiple images and the technique converting the DXF output to VRML 2.0 format. In the experiment, digital images captured by a digital camera are processed by PhotoModeler to acquire 3D model data. Camera positions are considered first, for better image acquisition. Next, principles and theories related to 3D modeling from images are studied to have better experimental results. A porsche toy car is used as a 3D object in the experiments. Finally,

a technique converting a DXF outputfile into VRML 2.0 file is shown.

I. 서 론

물체를 3차원 모델링하여 표현하고자 하는 기술에 대하여 많은 연구가 진행되고 있는데 그 기술은 두 가지로 나눌 수 있다. 하나는 사용자의 주관에 의해 물체의 모든 정보를 모델링하는 것이며, 다른 하나는 물체의 영상들에 근거하여 모델링한다. 전자의 경우 3차원 모델 제작 도구인 AutoCAD 또는 3D Studio가 사용된다. 이 제작도구로 물체를 모델링할 경우 실제 물체와 유사한 효과를 나타낼 수 있으나 실사로 나타낼 수는 없다. 이와 같은 단점을 해결하기 위해 영상들을 사용하여 물체를 모델링한다. 최근에 이러한 기술들을 제안하는 연구논문이 많이 발표되고 있다.

본 논문의 구성은 서론에 이어 II장에서 모델의 영상 획득을 위한 카메라 위치 설정과 획득 영상들로부터 3차원 모델 데이터를 추출하는 원리에 대해 알아본다. III장에서는 VRML 2.0 파일의 형식을 분석하고 변환 기법을 제시한다. IV장에서는 PhotoModeler를 사용하여 획득 영상들로부터 모델의 3차원 데이터를 추출하는 실험 과정과 VRML 2.0 파일형식으로의 변환 결과 및 오차분석을 기술한다. 마지막으로 V장의 결론으로 끝맺음한다.

II. 최적영상 및 물체 모델링

영상기반 모델링은 물체 정보를 영상들로부터 얻기 때문에 물체의 모델링을 정확히 하게하는 영상 획득이 필요하다. 최적의 영상 획득 조건에 대해 알아보고 이를 Porsche Toy Car에 적용하여 본다. 다음에 영상기반 모델링의 기본 이론에 대해 살펴본다.

2.1 최적 영상과 카메라 설정

최적 영상을 위해서는 인접하는 영상과의 각도를 90° 에 가깝도록 하여야 한다. 적어도 4개의 영상들이 필요하고, 이 영상들에는 물체의 모든 특징점(feature point)들이 나타나야 한다. 특징점들이 이웃 영상과 최소한 한면 이상 공통적으로 존재할 수 있도록 카메라 위치를 설정해야 한다.

Porsche Toy Car와 같은 박스 형태의 물체를 모델링하기 위해서는 그림 1과 같이 4개의 영상들을 사용한다. 각각의 카메라 위치는 자동차를 위에서 보았을 때를 나타낸다. 측면에서 보았을 때, 카메라의 위치는 모두 지면에서 약 45° 정도 위쪽에서 물체를 향하고 있다.

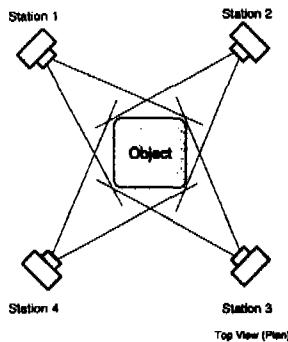


그림 1. 모델에 대한 카메라 위치 설정

물체의 한 면만을 모델링할 경우 카메라의 위치는 그림 2와 같으며, 적어도 3개의 영상이 필요하다. 측면에서 보았을 때, 각각의 카메라 위치는 지면에서 수평이며 같은 높이를 유지한다.

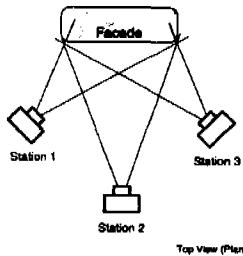


그림 2. 정면에 대한 카메라 위치 설정

물체의 특정부분이 방해물로 인하여 공통된 부분을 찾을 수 없는 카메라 위치가 그림 3에 나타나 있다. 각각의 카메라 위치는 지면에서 수평이며 같은 높이를 유지한다. 이와 같은 물체에서는 뒷면의 방해물 때문에 Station 1과 4의 영상에서 공통적으로 겹치는 부분을 찾을 수 없다. 따라서 Station 2와 3의 영상만을 가지고 물체의 3차원 정보를 추출하기 때문에 모델과의 정확도가 떨어진다. 이러한 문제를 해결하기 위해서는 Station 1과 4 사이의 중간 윗부분에 다섯 번째 카메라를 위치함으로서 Station 1과 4 사이의 공통면을 만들 수 있다.

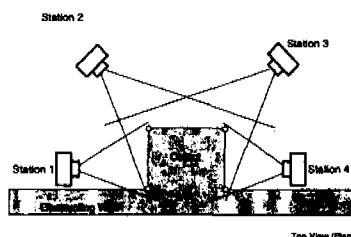


그림 3. 공통면이 없는 카메라 위치의 문제

2.2 물체 모델링

최적의 카메라 위치에서 획득된 물체의 영상들로부터 물체의 3차원 모델이 만들어질 수 있다.^{[1][5][6][8]} 스테레오 일치^{[2][3]}에 의해 3차원 모델구조가 결정될 수 있고, 물체의 면들은 영상들에 의해 표현될 수 있다.

스테레오 일치로 에지 일치 기법이 널리 사용된다. 이 기법은 영상들에 일치되는 에지들을 표시하고 이들과 그림 4(a)와 같이 투영되는 모델 에지와의 변위함을 목적함수로 설정, 최소화 시킨다. 그림 4(b)는 투영된 모델 에지와 관측된 영상에서의 에지를 보여준다. 이 두 에지의 차이가 하나의 Error이고 모든 Error의 합이 목적함수이다.

$$O = \sum \text{Error} \quad (1)$$

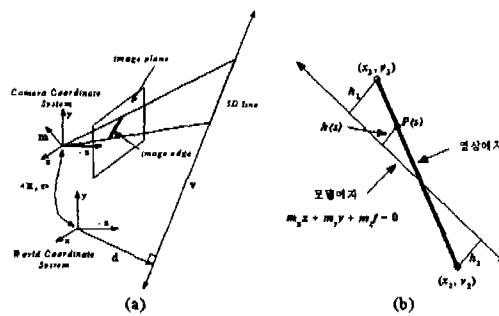


그림 4. (a) 카메라 평면에 투영되는 직선
(b) 영상 에지와 모델 에지

그림 4(b)에서 영상 에지의 길이 l 은 아래와 같으며 모델 에지를 $m_x x + m_y y + m_z z = 0$ 이라 할 때 점 $p(s)$ 에서 모델 에지에 이르는 최단거리 $h(s)$ 는 아래의 식으로 구해진다.

$$l = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2)$$

$$h(s) = h_1 + s \frac{h_2 - h_1}{l} \quad (3)$$

$$h_1 = \frac{m_x x_1 + m_y y_1 + m_z z_1}{\sqrt{m_x^2 + m_y^2}} \quad (4)$$

$$h_2 = \frac{m_x x_2 + m_y y_2 + m_z z_2}{\sqrt{m_x^2 + m_y^2}} \quad (5)$$

따라서 변위 Error는 다음과 같은 적분에 의해 구현된다.

$$Error = \int_0^l h^2(s)ds = \frac{l}{3}(h_1^2 + h_1h_2 + h_2^2) = m^T(A^TBA)m \quad (6)$$

$$A = \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{pmatrix}$$

$$B = \frac{l}{3(m_x^2 + m_y^2)} \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

목적함수 O 는 여러들의 합이며 이의 최소화는 모델과 카메라의 변수들에 대한 Hessian과 gradient를 반복 적용하는 Newton-Raphson 기법이 사용된다.

III. VRML 표현

3.1 VRML 2.0 형식

하나의 실린더를 3차원 공간에 표현하는 간단한 scene의 VRML^[9] 2.0 파일은 아래와 같다.

```
#VRML V2.0 utf8
Shape {
    appearance Appearance {
        material Material {
            diffuseColor 0.5 0.3 1.0
            shininess .5
        }
    }
    geometry Cylinder {
        radius 1
        height 3
    }
}
```

위 소스 코드를 도식화하면 그림 5와 같다.

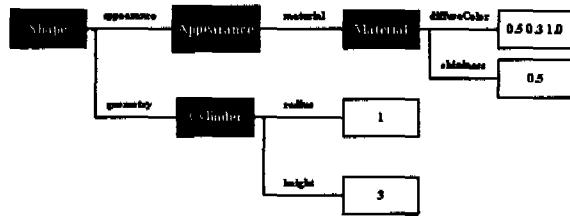


그림 5. VRML의 scene 구조

VRML 2.0의 기본 요소는 노드(node)이다. 그림 5에서 Shape, Appearance, Material, Cylinder 등이 하나의 노드이다. 노드는 대문자로 시작하며 중괄호(()) 안에 내용이 들어간다. 노드는 속성을 표현하는 필드(field)로 이뤄지며, appearance, material, diffuseColor, shininess, geometry, radius, height 등과 같이 소문자로 시작하는 것이 필드이다. 각 필드는 데이터형이 있어 그에 따른 데이터값만 가질 수 있다. 그림 5에서 반전된 박스는 노드, 이름이 주어진 선은 필드를 의미하며 반전되지 않은 박스는 필드값을 나타낸다.

3.2 VRML 2.0 변환

회복한 물체의 3차원 데이터를 VRML 2.0 형식으로 변환시킴으로서 웹상에서 물체의 3차원 정보를 표현할 수 있다.

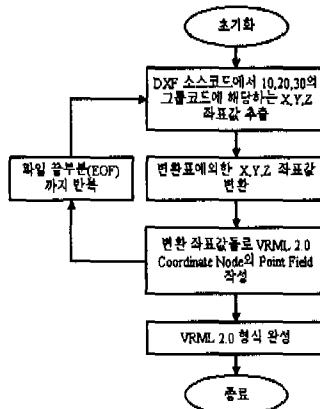


그림 6. VRML 2.0 변환

VRML 2.0 형식에서 물체 형상에 관한 정보는 Shape Node의 geometry 필드에 포함되어 있으며 그 필드는 IndexedFaceSet 노드로 정의되어 있다. 그리고 이것은 다시 coord와 coordIndex 필드로 나뉘어져 있다. 그림 6은 3차원 데이터를 VRML 2.0 형식으로 변환하는 알고리즘을 도식화 하였다.

IV. 실험 및 오차분석

4.1 실험과정 및 결과

이 장에서는 Porsche Toy Car에 대한 4개의 영상들을 사용하여 3차원 모델 데이터를 획득하는 실험 과정과 그 결과를 기술한다. 영상 획득은 Kodak DC20 Digital Camera를 사용하였다. 3차원 모델 데이터 획득 도구로 PhotoModeler를 사용하였으며, 486 PC에서 실험하였다. 영상 획득시 카메라의 위치는 그림 7과 같으며, 모델을 윗면, 정면, 그리고 우측면에서 보았을 때의 카메라 위치를 나타낸다. 4개의 영상들에 대한 각각의 카메라 위치를 그림 7과 같은 방법으로 입력한다.

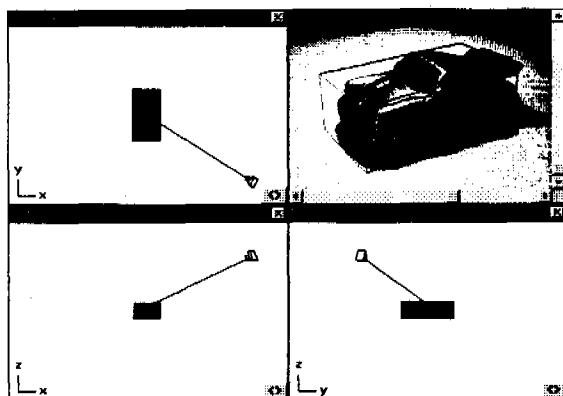


그림 7. 영상 획득시 카메라의 위치 설정

획득한 영상들에 설정한 특징점을 표시하고, 선으로 물체를 대략적으로 표현한 것이 그림 8이다. 두 영상에서 특징점을 사이를 연결한 점이 동일 위치를 가리키면 PhotoModeler는 이를 하나의 참조점(reference point)으로 인식한다.

그림 9는 특징점을 선으로 연결하여 물체의 전체 윤곽을 나타낸다. 이웃하는 3개의 점을 연결하는 삼각형을 기본 단위로 하여 물체의 전체적 표면을 형성한다.

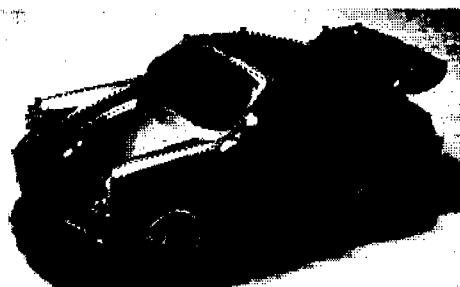


그림 8. 물체의 특징점 및 선 표시

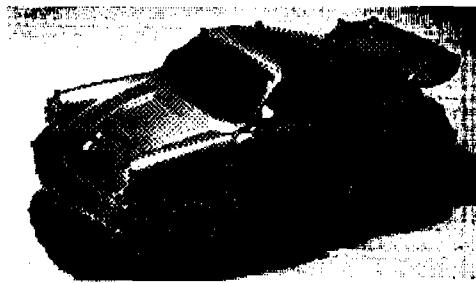


그림 9. 물체의 윤곽 및 표면 형성

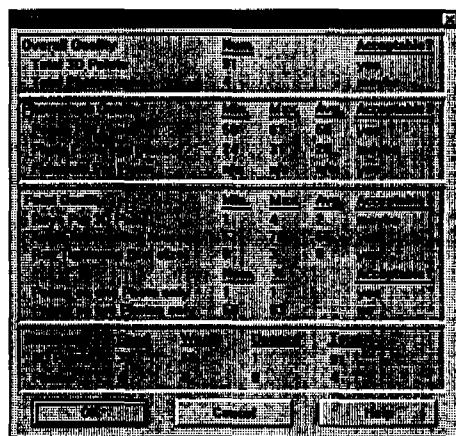


그림 10. 검사표

입력된 영상들의 카메라 위치, 그리고 지정된 특징점들에 대한 정보를 종합하여 그림 10과 같이 나타낸다. 4개의 영상에 사용된 특징점의 개수는 91개이다. 전체 영상에서 특징점들로 구성되는 표면 면적이 백분율로 나타낼 때 영상들은 최소 24%에서 최대 31%이다. 각각의 영상에서 기준점(fiducial point)은 없으며, Rays Per 3D Points는 특징점들이 영상들에 공통적으로 포함되는 영상 개수로, 최소 1개에서 최대 4개이다. Angle Intersection Average는 장난감 차에 적용되는 조명들 사이의 각도를 나타내며, 최소 57°, 최대 70°이다. Ray Tightness는 정확하지 못한 카메라의 위치로 인하여 장난감 차의 일부분에 조명이 적용되지 못하는 정도를 장난감 차의 크기에 대한 백분율로 3%이다. 특징점들과 검사표를 사용하여 PhotoModeler에서 출력한 DXF 파일을 VRML 2.0 파일로 변환 후에 본 모델의 3차원 그래픽스는 그림 11과 같다.



그림 11. VRML 2.0 파일 형식으로 본 모델의 3차원 그래픽스

4.2 오차분석

실험을 통하여 획득한 모델의 3차원 데이터를 실제 모델과 비교하여 정확도의 차이를 표 1, 표 2에 나타내었다. 기준이 되는 부분을 a로 정하고, a와 다른 부분의 상대적인 비를 표시하였다. 모델 자동차의 좌석에 앉은 사람을 기준으로 오른쪽을 우측면, 왼쪽을 좌측면으로 정하였다. 실제 모델에서 우측과 좌측의 비는 동일하다.

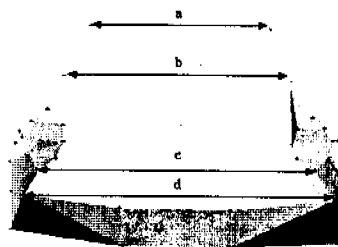


그림 12. 획득한 3차원 모델 데이터의 정면

표 1. 정면비교

비교부분 (a=1)	실제 모델비	정면획득 데이터비	정면 오차비율 (%)
b	1.20	1.26	5
c	1.14	1.53	34
d	1.30	1.72	32

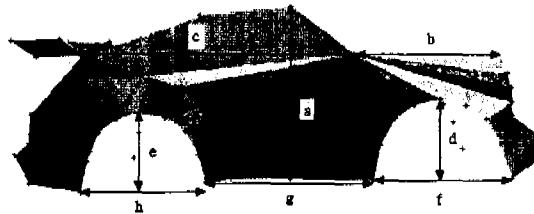


그림 13. 획득한 3차원 모델 데이터의 우측면

표 2. 측면비교

비교부분 (a=1)	실제 모델비	우측면 획득 데이터비	우측면 오차 비율 (%)	좌측면 획득 데이터비	좌측면 오차 비율 (%)
b	0.88	0.89	1.13	0.89	1.13
c	1.67	1.65	1.20	1.65	1.20
d	0.43	0.47	9.30	0.47	9.30
e	0.38	0.44	15.79	0.45	18.42
f	0.76	0.84	10.53	0.87	14.47
g	1.09	1.02	6.42	0.99	9.18
h	0.76	0.76	0	0.81	6.58

정면을 비교한 표 1에서 실제 모델비와 획득 데이터비가 차이가 남을 볼 수 있다. 정면을 나타내는 데이터에서는 원근법에 의한 표현이 이루어졌기 때문에 시점에 따른 데이터비가 실제 모델비와 차이가 난다. 측면을 비교한 표 2에서는 바퀴 부분을 제외하고 실제 모델비와 획득 데이터비가 거의 동일함을 알 수 있다.

VII. 결 론

본 논문에서는 먼저 Porsche Toy Car의 4개 영상들로부터 3차원 모델 데이터 획득실험에 대해서 기술하며, 다음에 DXF 파일로 출력되는 실험결과를 VRML 2.0 파일로 변환하는 기법에 대해 기술하고 있다.

실험에 사용된 도구는 PhotoModeler이며 이를 사용하여 Porsche Toy Car를 실사에 가깝게 모델링할 수 있었다. 모델링에 소요된 시간은 3시간 이내였다. 카메라의 적절한 위치 설정에 의해 획득한 최적 영상을 사용하였다. 그 결과 물체의 3차원 모델 데이터의 정확도가 증가됨을 알았다. 3차원 점 및 선들로 나열된 DXF 파일을 VRML 2.0 Shape 노드로 변환하여 물체의 형

태를 표현하도록 하였다. 표면의 재질감은 DXF 파일에 없어 각 표면은 단일색으로 표현하였다.

실제 물체와 획득한 모델 데이터를 비교하였다. 정면비교에서는 카메라에 가까운 데이터일수록 원근효과 때문에 오차가 증대되었으며, 측면비교에서는 오차가 균일하였다.

추후 연구할 사항으로는 표면재질의 표현 기술, 영상으로부터 숨겨진 부분의 보렐링 기술이다.

참고문헌

- [1] Olivier Faugeras, Stephane Laveau, Luc Robert, Gabriella Csurka, and Cyril Zeller. 3D reconstruction of urban scenes from sequences of images. Technical Report 2572, INRIA, June 1995.
- [2] Roger Tsai. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE Journal of Robotics and Automation, 3(4):323-344, August 1987.
- [3] Oliver Faugeras and Giorgio Toscani. The calibration problem for stereo. In Proceedings IEEE CVPR 86, pages 15-20, 1986.
- [4] Lance Williams and Eric Chen. View interpolation for image synthesis. In SIGGRAPH '93, 1993.
- [5] Leonard McMillan and Gary Bishop. Plenoptic modeling: An image-based rendering system. In SIGGRAPH '95, 1995.
- [6] R. Szeliski. Image mosaicing for tele-reality applications. In IEEE Computer Graphics and Applications, 1996.
- [7] Camillo J. Taylor and David J. Kriegman. Structure and motion from line segments in multiple images. IEEE Trans. Pattern Anal. Machine Intell., 17(11), November 1995.
- [8] Paul E. Debevec and Camillo J. Taylor and Jitendra Malik. Modeling and Rendering Architecture from Photographs: A hybrid geometry and image-based approach. In SIGGRAPH '96 Conference Proceedings, 1996.
- [9] The Virtual Reality Modeling Language Specification, Version 2.0, ISO/IEC CD 14772, August 4, 1996.