

## 카메라 파라미터 보정이 필요없는 물체의 3차원 구조 복원

\*추창우, 표순형, 박태준, 최병태, \*\*정순기  
\*한국전자통신연구원 가상현실연구부  
\*\*경북대학교 컴퓨터공학과

### Calibration-free 3D Structure Recovery

\*Chang Woo Chu, Soon Hyung Pyo, Tae-Joon Park, Byung Tae Choi, \*\*Soon Ki Jung  
\*Electronics and Telecommunications Research Institute  
E-mail : {cwchu, shpyo, ttjpark, btchoi}@etri.re.kr  
\*\*Department of Computer Engineering, KyungPook National University  
E-mail : skjung@knu.ac.kr

#### 요 약

컴퓨터 그래픽스 기술을 이용한 응용분야가 증가함에 따라 물체의 사실적인 모델에 대한 요구가 증가되고 있다. 모델링 시간이 많이 필요한 기존의 3차원 모델링 툴 외에 최근에 사용자의 스케치에 기반한 모델링 방법과 영상기반 모델링, 3차원 스캐너가 발표되었지만, 정확성이 떨어지거나 고가의 장비를 필요로 하는 단점이 있다. 본 논문에서는 정육면체 프레임과 광평면(light plane) 프로젝터, 카메라를 이용한 물체의 3차원 구조 복원 시스템을 제안하고, 실험을 통하여 모델링의 정확도를 분석한다.

#### 1. 서론

물체의 3차원 모델링은 가상 시뮬레이션, 인간-컴퓨터 상호작용(human-computer interaction) 등과 같이 컴퓨터 그래픽스 기술을 이용한 응용작업의 첫 번째 과정이다. 그러나, 캐드(CAD)와 같은 도구를 이용한 물체의 모델링 작업은 숙련된 전문가조차도 많은 시간을 들여야 하는 힘든 작업이다. 이러한 모델링 작업을 쉽게 하기 위하여 컴퓨터 그래픽스와 컴퓨터 비전 분야에서 많은 연구 결과들이 나왔으며, 대표적으로 스케치 기반 모델링[5, 9], 영상 기반 모델링[4, 6, 8], 3차원 스캐너[11, 12] 등이 있다.

스케치 기반 모델링[5, 9] 방법은 사용자의 입력 획(stroke)을 이용하여 물체 또는 환경을 모델링한다. Igarashi[5]는 사용자가 그린 물체의 윤곽선(silhouette)을 부풀려서 삼각형 메시로 이루어진 물체를 생성하고, 절단(cut), 내밀기(extrusion) 등의 조작을 통하여 물체를 모델링할 수 있는 Teddy라는 시스템을 제안하였다. Tolba[9]는 사용자가 원근 표시판(perspective guide)을 바탕으로 원근감을 살려 스케치한 각각의 획을 단위구에 매핑하여 파노라믹 영상

처럼 돌려볼수 있게 하는 시스템을 제안하였다. 사용자의 스케치에 기반한 모델링 방법은 사용자의 생각을 빠르고 간단하게 모델링할 수 있다는 장점이 있지만, 정확하지 않다는 단점이 있다. 이러한 기법은 프로토타입(prototype) 제작에 유용하게 쓰일 수 있다.

이와는 반대로 실물을 바탕으로 한 모델링 방법으로 영상 기반 모델링과 3차원 스캐너가 있다. 영상 기반 모델링은 사영기하학(projective geometry)[3, 7] 이론을 바탕으로 한 장[2, 4, 6] 혹은 여러 장[8]의 영상으로부터 물체의 3차원 구조를 복원하는 방법이다. Leibowitz[6]는 소실점(vanishing point)과 미리 알고 있는 각, 길이의 비 등을 이용하여 한 장의 영상으로부터 물체의 3차원 구조를 복원하는 방법을 제안하였다. 그 외에 영상에 나타나는 서로 직교하는 방향으로의 소실점을 이용하여 카메라를 보정(calibration)하고, 계산되어진 카메라 파라미터를 이용하여 물체를 모델링하는 방법이 있다. Guillou[4]는 2개의 소실점을 사용하여 카메라 보정을 하였고, Cipolla[2]는 3개의 소실점을 이용하였다. 이러한 방법들은 직교성을 바탕으로 하므로 건축물과 같은 물체의 모델링에는 적합하

지만, 곡면을 가진 물체나 나무와 같이 복잡한 물체의 모델링에는 적합하지 않다. Pollefe[8]는 여러 장의 영상에서 코너점을 대응시키고, 대응관계를 이용하여 조밀한 깊이 영상(depth map)을 만들어서 삼각형 메쉬 구조의 모델을 얻는 방법을 제안하였다. 영상기반 모델링 기법들은 간단한 도구를 사용한다는 장점이 있지만, 영상에 포함되어 있는 잡음으로 인해 정확도가 떨어지는 단점이 있다.

3차원 스캐너는 능동적 비전(active vision)[10] 기술의 광삼각법(optical triangulation)을 이용한 것으로, 수동적 비전(passive vision)[10] 기술을 이용한 영상기반 모델링 방법에 비해 정확하다. 그러나, 카메라 파라미터를 정확하게 보정해야 하며, 카메라와 광원의 위치 관계를 정확하게 알아야 한다. 한번의 스캐닝으로는 물체 전체를 모델링할 수 없기 때문에 여러 방향에서 스캐닝해야 하고, 각각의 범위 영상(range image)을 하나로 정합(registration)하는 과정이 필요하다. 이때, 여러 방향에서 스캐닝을 하는 과정에서 카메라와 광원의 움직임이 정확하게 제어되어야 하므로 고가의 장치가 필요하게 된다.

한편 Bouguet[1]은 카메라와 탁상용 램프, 연필 등의 간단한 도구를 사용하는 방법을 제안하였다. 이 방법은 영상의 모든 화소에 대하여 연필의 그림자에 포함되는 시간을 구하고, 램프와 그림자가 이루는 그림자 평면(shadow plane)과 COP에서 화소로 역투영된 직선이 만나는 점을 계산하여 범위 영상을 얻어낸다. 그러나, 영상에 나타나지 않는 부분과 물체에 의해 그림자가 생긴 부분에서는 정보를 추출할 수 없기 때문에 카메라와 램프를 이동시키면서 여러 번의 스캐닝을 수행하여야 한다. 따라서, 카메라나 광원을 이동시킬 때마다 위치를 보정해 주어야 하고, 모든 범위 영상을 하나의 영상으로 정합시켜야 한다.

본 논문에서는 정육면체 프레임과 광 평면 프로젝터(light plane projector), 카메라를 이용하여 물체의 기하 구조를 복원하는 새로운 시스템을 제안한다. 제안한 시스템에서는 영상에 나타나는 특징점을 이용하여 정육면체에 대한 물체의 상대적인 위치를 계산하기 때문에 카메라 파라미터 보정이 필요없고, 각각의 스캐닝된 정보는 하나의 기준 좌표계에 대한 상대적인 위치를 나타내기 때문에 정합과정도 필요하지 않은 장점이 있다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 시스템의 개요를 설명한다. 3장에서는 본 시스템에서 사용하는 3차원 좌표 계산 알고리즘을

서술하며, 4장에서는 실험결과를 보인다. 마지막으로 5장에서 결론을 내리고, 향후 과제를 제시한다.

## 2. 시스템 구성

본 장에서는 제안하는 시스템의 구성에 대해 설명한다. 제안하는 시스템은 그림 1과 같이 10개의 LED가 부착된 정육면체 프레임 및 디지털 카메라, 광 평면(light plane) 프로젝터를 사용한다. 정육면체 프레임의 각 꼭지점마다 1개의 LED를 부착하고, 좌표계의 z축과 원점을 표시하기 위해 보조 LED를 2개 더 부착한다. 보조 LED는 정육면체 모서리의 안쪽과 바깥쪽에 각각 1개씩 부착하여 카메라의 위치에 관계없이 항상 적어도 1개는 영상에 포착되도록 한다.

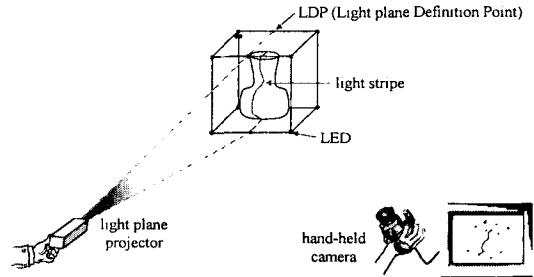


그림 1 시스템 구성도

영상을 획득할 때, 광 평면 프로젝터로부터 발하는 빛은 항상 정육면체의 평행한 3개의 모서리에서 만나도록 한다. 이때 3개의 점으로 정의되는 평면은 프로젝터로부터 발하는 평면과 같은 평면이므로, 이 3개의 점을 LDP(Light plane Definition Points)라고 한다. 만약 광 평면 프로젝터와 카메라가 평행한 방향으로 향하고 있다면, 스트라이프(stripe)는 직선으로 나타나므로, 스트라이프가 물체의 외형이 되기 위해서는 그림 1과 같이 카메라는 광 평면 프로젝터와 떨어진 위치에 있어야 한다. 그림 2는 마니킹을 정육면체 프레임 내부에 놓고 촬영한 영상의 예이다.

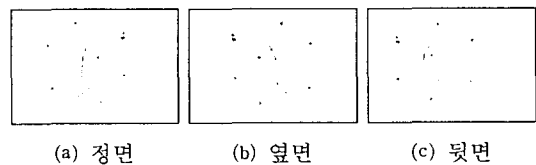


그림 2 입력 영상의 예

### 3. 3차원 좌표 계산 알고리즘

제안한 시스템으로 물체의 3차원 구조를 복원하는 세부 과정은 그림 3과 같은 단계를 거친다. 획득한 각각의 영상에서 LED에 좌표를 할당하고, LDP의 3차원 좌표를 계산한 후, 스트라이프 상의 점의 3차원 좌표를 계산한다. LDP의 3차원 좌표를 계산하기 위해서 입력 영상에서 LED를 판별하여 정육면체 프레임을 복원한 후 각 꼭지점에 정규화된 3차원 좌표를 할당한다.

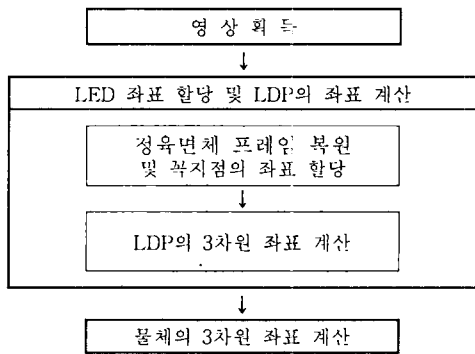


그림 3 시스템 흐름도

본 장에서는 LDP와 스트라이프 상의 점들의 3차원 좌표를 계산하는 방법에 대해 서술한다.

#### 3.1 LDP의 3차원 좌표 계산

광 평면 프로젝터에서 나온 빛은 정육면체의 모서리와 세 개의 LDP에서 만나고, 이 세 개의 점으로 광 평면을 정의할 수 있다. 스트라이프 상의 점  $P_s$ 는 그림 4와 같이 세 점  $P_u, P_v, P_w$ 로 정의되는 광 평면 위에 있다.

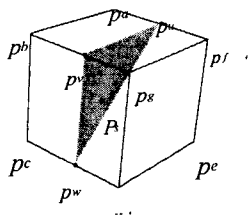


그림 4 정육면체 프레임과 광 평면의 관계

정육면체의 각 꼭지점에 3차원 좌표가 할당되어 있으므로, LDP의 3차원 좌표는 두 꼭지점을 보간(interpolation)하여 구한다. 예를 들어,  $P_u$ 는  $P_a$ 와

$P_f$ 를 이용하여 구한다. 즉, 보간을 위해  $\overline{P_a P_f}$ 와  $\overline{P_u P_u}$ 의 길이 비를 계산하여야 한다. 그러나, 일반적인 핀홀(pinhole) 카메라는 실세계를 원근투영(perspective projection)을 통하여 영상을 포착하므로, 실세계의 평행성(parallelism)과 길이를 그대로 반영하지 못한다[3, 7, 8]. 따라서, 영상에서의  $\overline{p_a p_f}$ 와  $\overline{p_u p_u}$ 의 길이 비는 실제 3차원 상의 길이 비를 나타내지 않으므로 그대로 적용할 수 없다.

본 논문에서는 정육면체 프레임의 평행한 모서리가 투영되어 만나는 소실점과 투영 기하(projective geometry)에서 불변(invariant)의 성질을 가진 교차율(cross-ratio)을 이용한다. 그림 5는 원근 투영시킨 정육면체의 소실점과 소실선(vanishing line)을 나타낸다. 어떤 평면에 대한 소실선은 그 평면과 평행한 모든 평행한 직선들이 만나는 소실점을 모은 것이다. 정육면체와 카메라의 위치에 따라 영상에 나타나는 소실점의 개수가 달라진다. 소실점이 2개 이하인 경우는 서로 평행한 모서리 쌍이 카메라의 영상 평면과 평행할 때 발생하며, 소실점은 무한대에 있다.

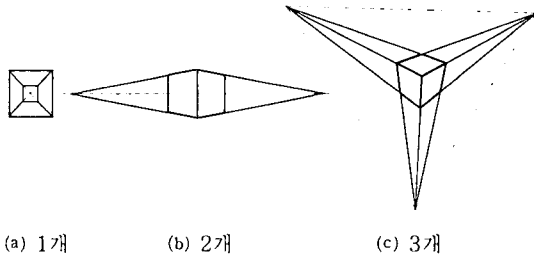


그림 5 원근 투영된 정육면체의 소실점 개수

LDP의 3차원 좌표는 교차율을 이용하여 계산한다. 교차율은 그림 6의  $p_v, p_a, p_u, p_f$ 와 같이 동일 직선 상의 네 점에 대해서 식 1과 같이 정의한다[3, 7, 8].

$$C_r(p_v, p_a; p_u, p_f) = \frac{|p_v p_u|}{|p_v p_f|} : \frac{|p_a p_u|}{|p_a p_f|} \quad \text{식 1}$$

식 1에서  $|\overline{p_v p_f}|$ 는 영상에서의  $p_v$ 과  $p_f$ 의 거리를 나타낸다. 식 1의 비율은 카메라의 위치에 관계없이 일정하게 유지될 뿐만 아니라, 실세계를 나타내는 유클리드 기하(Euclidean geometry)에서도 그대로 유지된다. 따라서, 실세계에서의  $|\overline{P_a P_f}|$ 에 대한  $|\overline{P_u P_u}|$ 의 비를 계산하기 위해 식 1을 그대로 사용하는다.

$$\frac{|P_a P_u|}{|P_a P_f|} = \frac{|p_{v_1} p_f|}{|p_{v_1} p_u|} \times \frac{|p_a p_u|}{|p_a p_f|} \times \frac{|P_{v_1} P_u|}{|P_{v_1} P_f|} \quad \text{식 2}$$

$$P_u = P_a + \frac{|P_a P_u|}{|P_a P_f|} (P_f - P_a) \quad \text{식 3}$$

식 2에서  $|P_a P_f|$ 는 3차원에서의  $P_a$ 와  $P_f$ 의 거리를 나타낸다. 그리고,  $P_{v_1}$ 은 무한대의 점이므로 정확한 좌표를 알 수 없기 때문에, 두 점 사이의 거리는 정확하게 알 수 없지만,  $|P_{v_1} P_f| \approx |P_{v_1} P_u|$ 라는 사실을 알 수 있다. 따라서, 영상에서의 교차율을 이용하여 식 2와 같이 3차원에서의 두 점간의 거리 비를 알 수 있다. 이것을 이용하여  $P_u$ 의 좌표는 식 3과 같이 계산할 수 있다. 나머지  $P_v$ 와  $P_w$ 에 대해서도 소실점  $p_{v_1}$ 을 지나는 직선 상의 네 점을 이용하여 같은 방식으로 3차원 좌표를 계산할 수 있다.

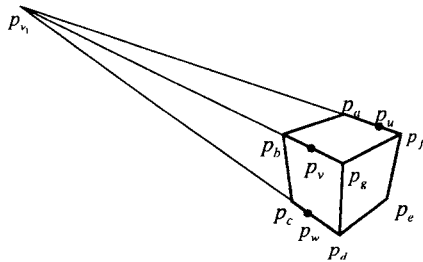


그림 6 교차율을 이용한 LDP의 3차원 좌표 계산

소실점이 무한대에 있는 경우, 즉, 그림 5의 (a), (b)와 같은 경우도 같은 방법으로 계산할 수 있다. 식 2에서  $|p_{v_1} p_f| \approx |p_{v_1} p_u|$ 이므로, 3차원 상의 거리의 비는 영상에서의 거리의 비와 같다. 따라서, 소실점이 3개인 경우보다 더 간단하게 LDP의 좌표를 계산할 수 있다.

### 3.2 물체의 3차원 좌표 계산

정육면체 프레임의 꼭지점 및 LDP의 좌표를 이용하여 스트라잎 상의 점의 좌표를 계산한다. 그림 7과 같이 스트라잎 상의 한 점  $p_s$ 는 공간상의 두 직선  $l'$ 과  $l''$ 의 교점이다. 이때,  $l'$ 은  $\overline{P_v P_w}$ 에 평행하면서  $P_s$ 를 지나는 직선이고,  $l''$ 은  $\overline{P_u P_v}$ 와 평행하면서  $P_s$ 를 지나는 직선이다. 공간상의 평행한 두 직선  $\overline{P_u P_v}$ 와  $l''$ 은 영상 평면에 원근 투영되면 더 이상 평행하지 않고 소실점  $p_{v_3}'$ 에서 만나게 되며,  $\overline{P_u P_v}$

와  $l''$ 도  $p_{v_3}'$ 에서 만나게 된다. 이때,  $\overline{P_u P_v}$ 는 수평 방향의 평면에 포함되므로  $\overline{P_u P_v}$ 와  $l''$ 의 소실점  $p_{v_3}'$ 는 그림 8과 같이  $p_{v_1}$ 과  $p_{v_2}$ 를 지나는 소실선에 포함되며, 마찬가지로,  $p_{v_3}'$ 는  $p_{v_1}$ 과  $p_{v_2}$ 를 지나는 소실선 상에 존재하게 된다.

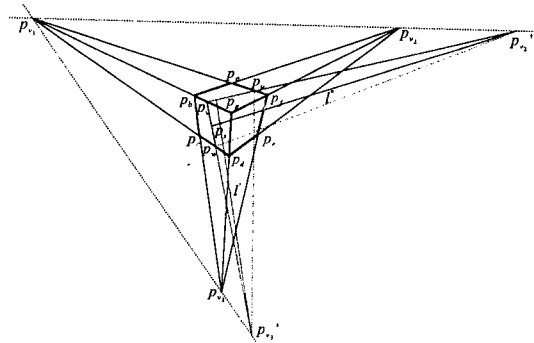


그림 7 광 평면과 정육면체 프레임의 관계

점  $p_s$ 의 3차원 좌표  $P_s$ 를 계산하는 것은 기본적으로 3.1절에서 설명한 교차율을 이용한 보간을 사용한다. 그림 8에서  $P_s$ 는  $l''$ 상의 네 점에 대한 교차율  $Cr(p_{v_1}', p_{s_1}; p_s, p_{s_2})$ 과  $Cr(P_{v_2}', P_{s_1}; P_s, P_{s_2})$ 를 이용하여 계산한다. 이때,  $P_{s_1}$ 과  $P_{s_2}$ 는 각각  $Cr(p_{v_1}', p_{e'}; p_{s_1}, p_u)$ 과  $Cr(P_{v_1}', P_{e'}; P_{s_1}, P_u)$ ,  $Cr(p_{v_1}', p_w; p_{s_1}, p_v)$ 과  $Cr(P_{v_1}', P_w; P_{s_1}, P_v)$ 을 이용하여 계산할 수 있다.

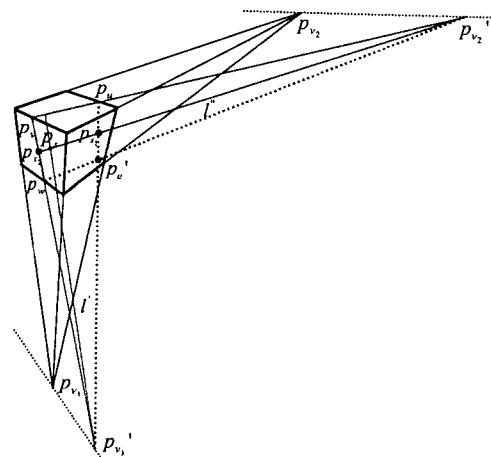


그림 8 스트라잎 상의 점의 3차원 좌표 계산

지금까지 서술한 바와 같이 카메라와 광원의 위치

에 관계없이 영상에 포착되는 정육면체 프레임의 소실점과 교차율을 이용하여 간단한 수식으로 물체 표면상의 3차원 좌표를 쉽게 구할 수 있다.

#### 4. 실험 결과

본 논문에서는 물체의 3차원 좌표를 계산하기 위해서 광원과 카메라의 위치에 관계없이 영상의 소실점과 교차율을 이용하는 방법을 사용하였다. 또한, 본 시스템의 환경에서는 소실점을 이용하여 카메라 파라미터를 보정할 수 있다[2, 4, 6]. 보정된 카메라 파라미터를 이용하여 3차원의 선과 선, 선과 면의 교점을 계산하여 물체를 모델링할 수 있다.

본 장에서는 물체의 3차원 구조 복원 결과의 정확도를 분석하기 위하여 두 가지 실험을 하였다. 미리 알고 있는 물체를 모델링한 결과와 원래 물체의 유사 정도를 분석한다. 이를 위하여 형면을 가진 물체를 모델링하여 모델링된 평면이 얼마나 평면에 가까운가를 검사하였다. 그리고, 물체의 특징 한 점을 광 평면 프로젝터의 위치를 고정시키고 카메라를 임의의 위치로 움직여 가면서 계산한 3차원 좌표의 일관성(consistency)을 살펴보았다. 이 실험에서는 제안한 시스템에서 3차원 좌표를 계산할 수 있는 두 가지 방법으로 실험하여 결과를 비교한다. 실제 실험은 그림 9와 같이 크기를 알고 있는 상자를 정해진 위치에 놓고 수행하였다.

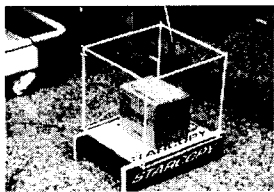


그림 9 실험 장면

##### 4.1 평면 검사

본 실험에서는 그림 9와 같은 환경에서 카메라의 위치를 고정시키고, 광 평면 프로젝터를 움직이면서 모델링 하였다. 그림 10은 획득한 영상의 예를 보여준다.



그림 10 평면 검사 입력 영상 예

그림 10에 나타난 것처럼 실험 환경에는 모두 3개의 평면이 있다. 그 중 바닥을 제외한 상자의 윗면과 수직면에 대해서 점들의 분포를 조사하였다. 실험에서 사용한 정육면체 프레임의 한 변의 길이는 26cm이며, 상자의 높이는 13.5cm이다. 그리고, 정육면체 프레임의 x축에 대해 8.6cm 거리에 상자를 위치시켰다. 실험 결과는 표 1에 나타나있으며, 그림 11은 계산한 점들을 렌더링한 결과이다.

	실제값 (높이/길이)	평균	표준 편차
윗면	$z = 0.519$	0.515	0.003684
수직면	$y = 0.331$	0.323	0.004140

표 1 윗면과 수직면의 평면 검사

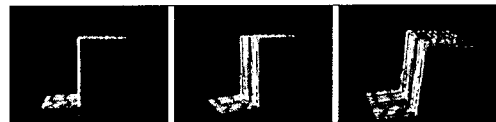


그림 11 평면의 렌더링 결과

##### 4.2 3차원 좌표의 일관성 검사

본 실험에서는 광 평면 프로젝터의 위치를 고정시키고, 카메라를 임의의 위치로 움직이면서 모델링한 결과가 어느 정도 일치하는지를 분석한다. 그림 12는 왼쪽으로 이동과 회전을 하면서 촬영한 영상의 예이다.

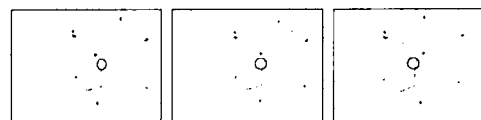


그림 12 카메라 이동 검사 입력 영상 예

표 2는 그림 12에 원으로 표시된 상자의 윗면과 수직면의 경계점을 두 가지 방법으로 계산한 3차원 좌표의 표준 편차를 나타낸다. 표 2에 나타난 바와 같이 두 방법 모두에서 임의의 카메라의 움직임에 대해 비교적 일정하게 좌표가 계산됨을 알 수 있다. 그러나, 한 장의 영상에서 정보를 추출한 후, 실제 계산 시간에서는 제안한 방법이  $1.292 \times 10^{-5}$ 초로 카메라 파라미터를 보정하는 방법의 수행시간  $2.521 \times 10^{-5}$ 초보다 훨씬 빠름을 알 수 있었다. 이는 매 프레임마다 카메라 파라미터를 보정해야 하기 때문이다. 또한, 소실점이 하나뿐인 경우에는 카메라 파라미터 보정이 불가능한

문제가 있다.

	x	y	z
제안한 방법	0.002865	0.003719	0.004652
카메라 파라미터 사용	0.002971	0.005221	0.004449

표 2 3차원 좌표의 일관성 비교

그림 13은 실험에 사용한 마네킹과 본 시스템을 사용하여 3차원 구조를 복원한 결과이다.

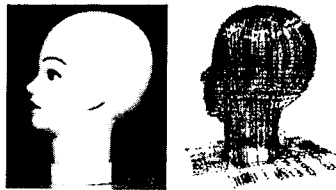


그림 13 마네킹을 모델링한 결과

#### 5. 결론 및 향후 연구 과제

본 논문에서는 정육면체 프레임과 광 평면 프로젝터, 카메라 등의 간단한 도구를 사용하여 물체를 모델링하는 새로운 시스템을 제안하였다. 카메라 파라미터의 정확한 보정과 카메라와 광원의 위치가 엄격하게 유지되어야 하는 기존의 시스템과는 달리 본 논문에서 제안한 시스템은 카메라 파라미터의 보정이 필요 없고, 광원과 카메라를 자유롭게 움직일 수 있는 장점이 있다. 또한, 계산된 3차원 좌표는 프레임에 대한 상대적인 위치이므로, 정합하는 과정을 거치지 않고도 점들의 집합으로 표현이 된다.

그러나, 제안한 방법은 3차원 좌표의 계산과 표면 재구성이 서로 유기적인 관계를 가지지 않기 때문에 사용자와의 상호작용이 부족한 단점이 있다. 즉, 획득한 3차원 좌표의 부분 결과로 표면을 재구성하여 사용자에게 보여줘서 사용자가 좀 더 세밀한 모델링이 필요한 부분을 선별할 수 있다면 보다 나은 모델을 얻을 수 있을 것이다.

#### [참고문헌]

[1] J. Y. Bouguet and P. Perona, "3D photography on your desk," *International Conference on Computer Vision*, IEEE, Bombay, India, pp. 43~50, 1998.

[2] R. Cipolla, T. Drummond and D. Robertson, "Camera calibration from vanishing points in images of architectural scenes," *British*

*Machine Vision Conference*, pp. 382-391, 1999.

[3] O. Faugeras, *Three-dimensional computer vision - a geometric viewpoint*, M.I.T Press, Cambridge, MA, 1993.

[4] E. Guillou, D. Meneveaux, E. Maisel and K. Bouatouch, "Using vanishing points for camera calibration and coarse 3D reconstruction from a single image", *The Visual Computer* 16:396-410, Springer, 2000.

[5] T. Igarashi, S. Matsuoka and H. Tanaka, "Teddy: A Sketching Interface for 3D Freeform Design," In proceedings of SIGGRAPH 1999.

[6] D. Leibowitz, A. Criminisi and A. Zisserman, "Creating architectural models from images," In proceedings of EUROGRAPHICS, vol. 18, pp. 39~50, 1999.

[7] J. L. Mundy and A. Zisserman, editors, *Geometric invariance in computer vision*, MIT Press, 1992.

[8] M. Pollefeys, *Obtaining 3d models with a hand-held camera*, courses on SIGGRAPH 2001.

[9] O. Tolba, J. Dorsey and L. McMillan, "A Projective Drawing System," *Interactive 3D Graphics*, 2001.

[10] E. Trucco and A. Verri, *Introductory techniques for 3-d computer vision*, Prentice Hall, 1998.

[11] Head & face color 3d scanner, by Cyberware, <http://www.cyberware.com/>.

[12] RealScan, by Real3D, <http://www.real3d.com/>.