

다수의 2차원 영상으로 3차원 물체 복원 방법

장우석* · 호요성**

1. 서 론

다시점 2차원 영상으로부터 3차원 물체를 복원하는 작업은 컴퓨터 비전 분야에서 활발히 연구되고 있다. 그 주요 목적은 물체의 3차원 기하학 정보와 색상 정보를 얻어내는 것이다 [1]. 일반적으로 3차원 물체를 복원하기 위해서는 다시점 영상과 카메라 매개변수를 이용하여 3차원 기하학 정보를 계산한다. 다시점 영상을 위한 카메라 매개변수는 카메라 보정 기술에 의해서 얻어진다 [2]. 3차원 점 (X)과 그것의 어떤 한 시점 위로의 투영된 점 (x) 사이의 관계는 다음과 같다 [3].

$$x = PX \quad (1)$$

여기서 P는 카메라 투영 행렬을 나타낸다. 각 시점은 각각 자신의 투영 행렬을 가지고 있다. 카메라 투영 행렬은 3차원 복원을 위해서 유용한 요소이다.

그림 1은 카메라 매개변수에 의해서 3차원 공

간의 한 점이 2차원 영상 평면으로 투영된 것을 보여준다. 그림 1로부터 두 시점에서의 일치하는 점들의 3차원 공간으로의 역투영된 직선의 교차점을 이용하여 물체의 3차원 점을 구할 수 있다는 것을 알 수 있다. 이러한 방법을 삼각화(triangulation) 방법이라고 한다. 이와 같이 3차원 복원 방법은 일반적으로 영상간의 매칭에 기초를 둔다. 이러한 접근은 특히 일치점을 찾는 문제를 단순화하는 기술에서 효율적이다. 그러나 일치점을 찾는 문제는 쉬운 문제가 아닐 뿐만 아니라, 복잡도도 상당히 높다. 게다가 시점들은 서로 가까워야한다는 제약사항이 있다. 이러한 단점을 피하기 위해서 본 논문에서는 삼각화와 다른 접근법을 사용하였다.

본 논문에서 제안하는 알고리즘은 실루엣 기반 복원(shape from silhouette)[4]과 복셀 컬러링(voxel coloring)[5]의 이론적인 기법을 결합하였

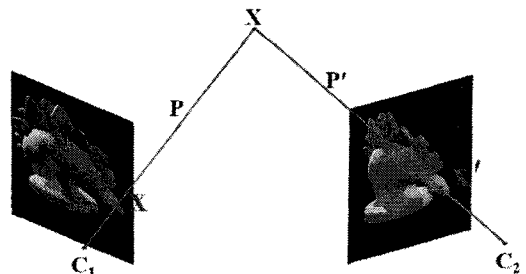


그림 1. 3차원 점의 2차원 평면으로의 투영

* 교신저자(Corresponding Author): 호요성, 주소: 광주광역시 북구 첨단과기로 261 (500-712), 전화: 062)970-2211, FAX: 062)970-3164, E-mail: hoyo@gist.ac.kr

* 광주과학기술원 정보통신공학과 (E-mail: jws@gist.ac.kr)

** 광주과학기술원 정보통신공학과

※ 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음(NIPA-2009-(C1090-0902-0017)).

다. 이는 초기의 3차원 물체의 윤곽을 구하고 이를 좀 더 정확하게 보정하기 위한 것이다. 가시성(visibility) 검사는 불필요하게 물체의 내부까지 표현되는 것을 막기 위해 사용된다. 물체의 내부까지의 표현은 3차원 복원의 복잡도를 높일 뿐만 아니라, 3차원 모델의 정확도를 떨어뜨린다. 이는 3차원 물체의 한 점과 각 시점의 카메라 중심 사이의 직선의 방정식을 이용하여 수행된다. 색상값은 카메라 위치에 따라서 정해진다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 제안하는 방법에 이용되는 일반적인 3차원 복원 방법의 이론에 대해서 소개하고, III장에서는 제안하는 3차원 물체 복원 방법에 대해서 설명한다. 그리고 IV장에서 실험 결과를 분석한 뒤에, V장에서 본 논문을 마무리한다.

2. 3차원 복원 방법들

2.1 실루엣 기반 복원 (shape from silhouette)

실루엣 기반 복원은 입력 영상으로부터 시각체(visual hull)를 만들어 3차원 물체의 모양을 복원하는 방법이다. 시각체는 3차원 물체를 근사한 것이다. 이 근사한 모양을 얻기 위해서 먼저 입력 영상에서 물체와 배경을 분리한 실루엣 영상을 만들어야 한다. 실루엣 영상에서 실루엣의 윤곽선을 추출한 후, 카메라 보정 정보를 사용하여 윤곽선의 모든 점을 3차원 공간으로 역투영한다. 이 역투영된 실루엣 윤곽선들은 시각 원뿔(visual cone)을 만들고, 시각 원뿔들의 교차가 시각체를 생성한다. 시각체와 같이 실루엣 기반 복원 방법은 구현 방법이 간단하여 다수의 2차원 영상으로부터 3차원 물체 모양을 복원하기 위해서 많이 사용되고 있다 [6].

2.2 복셀 컬러링(voxel coloring)

복셀 컬러링 방법은 색상 일치성에 기초를 둔다. 서로 다른 시점으로부터 보이는 가려지지 않는 점의 색상이 같다면, 그 3차원 점은 3차원 물체의 표면이라고 간주된다. 그 반대의 경우는 3차원 점이 빈 공간에 있다고 추정된다. 복셀 컬러링은 이러한 원리를 통하여 3차원 물체를 복원한다.

3. 3차원 물체 복원 알고리즘

3.1 초기의 3차원 물체 윤곽 탐지

3차원 물체의 모양을 근사적으로 표현하기 위해서 복셀 구조를 적용한다 [7]. 3차원 물체의 윤곽을 찾기 위해서 각 복셀을 각 카메라의 영상 평면에 투영을 한다. 만약 복셀이 2차원 영상의 물체 영역에 투영이 된다면, 이 복셀은 3차원 물체에 속한 복셀이라고 간주한다. 그림 3은 복셀 표현과 3차원 공간의 복셀로부터 2차원 영상 평면에 투영하는 과정을 보여준다. 목표하려는 물체의 3차원 윤곽이 이러한 과정을 거쳐서 얻어진다.

그러나, 이 과정에서는 3차원 물체의 오목한 부분을 인식할 수 없을 뿐만 아니라, 물체의 오목한

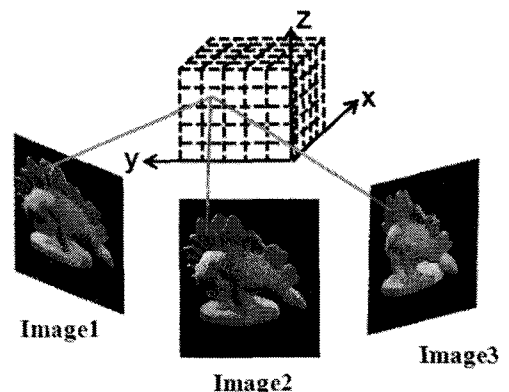


그림 2. 초기 3차원 윤곽을 위한 복셀 구조

부분이 없을지라도 유한한 카메라를 가지고는 정확하게 3차원 물체의 윤곽을 얻을 수 없다. 그러므로 이 결과를 실제 3차원 물체의 윤곽으로 사용하기는 어렵다. 그러나 이 결과에 또 다른 과정을 더해서 정확한 3차원 물체 복원을 위한 초기 3차원 물체의 윤곽으로 사용할 수 있다.

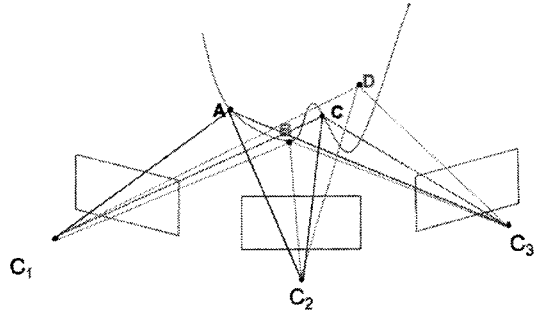


그림 3. 가시성 검사

3.2 가시성 검사

3차원 물체 복원에 있어서 가시성 검사는 여러 과정에서 큰 이득이 된다. 가시성 검사를 통해서 3차원 복원 과정의 복잡도를 낮출 수 있고, 3차원 물체의 정확도도 증가시킬 수 있다. 이러한 이유로 인해, 3차원 물체의 초기 윤곽에 있는 3차원 점과 각 카메라의 중심 사이의 직선을 만든다. 이 직선의 방정식은 다음과 같이 얻어진다.

$$X = [X - C] * s + C \tag{2}$$

벡터 $X = [xyz]^T$ 는 3차원 물체의 한 점을 나타내고, C 는 카메라 중심의 3차원 위치, s 는 임의의 스칼라 값을 나타낸다. 이 직선의 방정식은 가시성 검사를 위해 사용된다. 만약 3차원 물체에서의 한 점과 모든 카메라 중심 사이에서 연결된 직선이 3차원 물체의 다른 어떤 점을 포함한다면, 이 점은 가시성이 없는 것으로 간주한다. 따라서 이 점을 3차원 물체의 표면의 후보에서 배제시킨다. 초기의 3차원 물체 윤곽 위의 모든 점에서 이러한 과정을 적용하면, 합리적으로 가시성을 검사할 수 있다. 가시성 검사는 뒤에서 소개할 물체 다듬기와 색상값 사상 과정에서 유용하게 쓰인다. 그림 3은 가시성 검사의 과정을 보여준다. 그림 3에서 A, B, C와 D는 3차원 물체의 점들이다. 점 A가 3차원 물체의 가시성을 가지는지 아닌지를 검사하기 위해서 점 A에 연결된 직선을 조사한다. 3차원 물체의 한 점 A와 한 카메라 중심인 C_1 을 연결

하는 직선 AC_1 이 3차원 물체의 어느 점도 만나지 않는다. 따라서 점 A는 3차원 물체의 보이는 점으로 간주된다. 점 B와 점 C가 가지는 직행의 직선의 개수도 하나 이상이므로 점 B와 점 C도 3차원 물체의 보이는 점으로 간주된다. 그러나 직선 DC_1 , 직선 DC_2 와 직선 DC_3 는 3차원 물체 위의 점들을 만난다. 그러므로 점 D는 보이지 않는 점이라는 것을 알 수 있다.

3.3 3차원 물체 다듬기

이전의 과정까지는 3차원 물체의 오목한 부분을 포함해서 정확한 물체의 윤곽을 얻지 못했다. 따라서 3차원 물체의 오목한 부분과 유한한 카메라로 인하여 생긴 정확하지 못한 물체의 윤곽을 다듬어서 3차원 물체 윤곽의 정확도를 높이는 과정이 필요하다. 그림 4는 3차원 물체 다듬기의 과정을 보여준다. 다듬기 과정을 간단히 설명하면 다음과 같다. 초기의 3차원 물체의 윤곽 탐지 과정을 통해서 이미 실제 물체를 포함하는 물체의 분량을 가지고 있다. 가시성 정보를 이용하여 가시성이 있는 물체의 한 점 위의 복셀을 선택한다. 이 복셀을 보이는 입력 영상 위로 투영한다. 각 시점의 색상값이 다른 시점에서의 색상값과 다르다면 다듬기 과정을 수행한다 [8]. 이러한 과정을 더 이상 다듬기를 할 수 없을 때까지 반복하고,

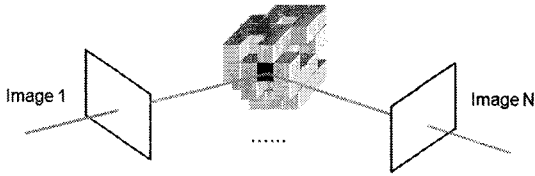


그림 4. 물체의 표면 찾기

가시성 정보를 갱신한다.

물체 다듬기는 RGB 거리를 이용한 색상값 차이를 이용하여 처리된다. 그림 5는 RGB 영역에서의 RGB 거리를 구하는 방법을 보여준다.

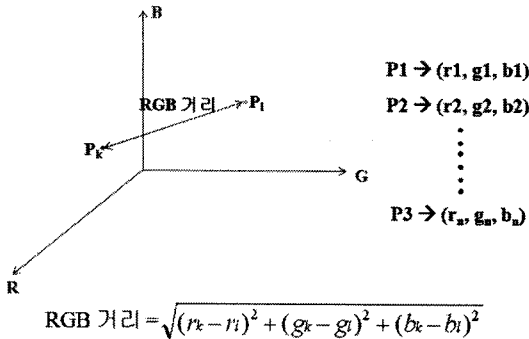


그림 5. RGB 거리

다듬기 과정을 위해서 RGB 거리의 문턱값을 정해야한다. 단순한 복셀 단위의 물체 다듬기에서는 어떠한 RGB 거리의 문턱값을 사용해도 물체의 정확한 윤곽을 얻어내기가 쉽지 않다. 그 이유는 복셀의 중심점만 복셀의 대표값으로 간주하여 물체 다듬기를 시도하기 때문이다. 실제로 3차원 공간에서의 하나의 복셀은 2차원 영상의 여러 화소에 영향을 받는다. 그림 6은 RGB 거리의 문턱값에 따른 물체 다듬기 결과를 보여준다. 문턱값이 작을 때는 정확히 3차원 물체가 얻어지지 않는 것을 알 수 있다. 이러한 잘못된 결과는 2차원 영상의 해상도가 복셀의 해상도보다 크기 때문에 발생한다. 하지만 복셀 해상도를 2차원 영상의 해상도 이상으로 설정하는 것은 구현 과정에 복잡도



그림 6. 문턱값에 따른 물체 다듬기 결과

를 증가시킨다.

따라서 복셀의 값을 대표할 수 있는 값을 얻기 위한 표본 추출법이 필요하다. 그림 7은 복셀의 중심값만 2차원 영상에 투영될 때와 복셀 전체가 투영되는 경우를 보여준다. 그림 7과 같이 하나의 복셀은 2차원 영상의 여러 화소에 걸쳐 있으므로 표본 추출법을 이용하여 합리적인 복셀의 값을 얻을 수 있도록 해야 한다.

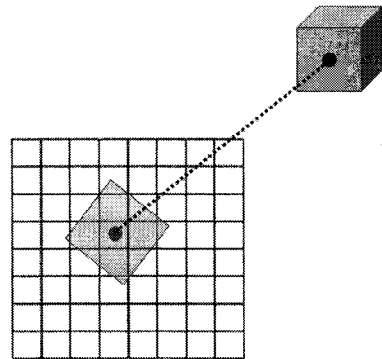


그림 7. 표본 추출법을 위한 복셀 투영

앞에서 설명했듯이, 하나의 복셀이 하나의 화소보다 더 넓은 범위로 투영될 때 단지 복셀 중심에 있는 하나의 중심점이 아닌 복셀이 투영된 모든 점의 정보가 고려되어야 한다. 이를 위해서 다중 벡터 교차 상호 관계 기준(multi-vector cross-correlation criterion) 방법을 이용하였다 [9]. 이 방법을 통해서 복셀이 투영된 모든 화소들 중에서 복셀을 대표할 만한 화소만 추출해낼 수 있다. 방법은 다음과 같다. 하나의 복셀이 투영된 2차원 영상의 모든 화소에 대해서 각각 벡터를 하나씩

만든다. 각 벡터는 R, G, B 세 가지 요소로 이루어져 있다.

$$v_i = [R_i \ G_i \ B_i]^T \quad i = 1, 2, 3, 4, 5, \dots, n \quad (3)$$

n개의 화소값을 (3)과 같이 벡터로 만든 후, 그들의 유사도를 측정한다. 만약 n=2라면 매우 잘 알려진 정규화된 상호관계(normalized cross-correlation)가 기준이 될 수 있다. 정규화된 벡터 n_i 를 정의하면 다음과 같다.

$$n_i = \frac{v_i - m_i}{\|v_i - m_i\|} \quad i \in \{1, 2, \dots, n\} \quad (4)$$

여기서 m_i 는 v_i 의 각 요소들인 R_i, G_i, B_i 의 평균으로 구성된 벡터이다. 또한 자기 자신의 벡터를 제외한 다른 벡터들의 평균을 구한 부분 평균 합 벡터 s_i 를 다음과 같이 정의한다.

$$s_i = \frac{1}{n-1} \sum_{j \neq i} n_j \quad (5)$$

모든 벡터와 부분 평균 합 벡터 s_i 와의 유사도를 계산하면 다음과 같다.

$$p_i = s_i \cdot n_i \quad \forall i \quad (6)$$

여기서 p_i 의 최대값 p_{max} 를 구하고, 다음을 만족하는 벡터의 화소만 복셀을 대표할 수 있는 표본으로 추출한다.

$$p_i > T \cdot p_{max}, \quad T \in (0, 1) \quad (7)$$

이렇게 화소의 표본이 추출되었다면 선택된 표본의 평균을 구해서 복셀의 대표값으로 설정한다.

물체의 다듬기를 할 때 어떤 점의 다듬기를 시도할 것인지 아닌지를 판단하기 위해서 그 점을 바라보는 시점들을 찾는다. 많은 시점들이 다듬기를 시도하는 점을 바라볼 때, 3차원 점이 물체의 실제 3차원 물체의 윤곽일지라도 한 시점이 다른 시점과 많이 떨어져있는 경우라면 RGB 거리가 멀어질 수 있다. 이러한 문제를 해결하기 위해서

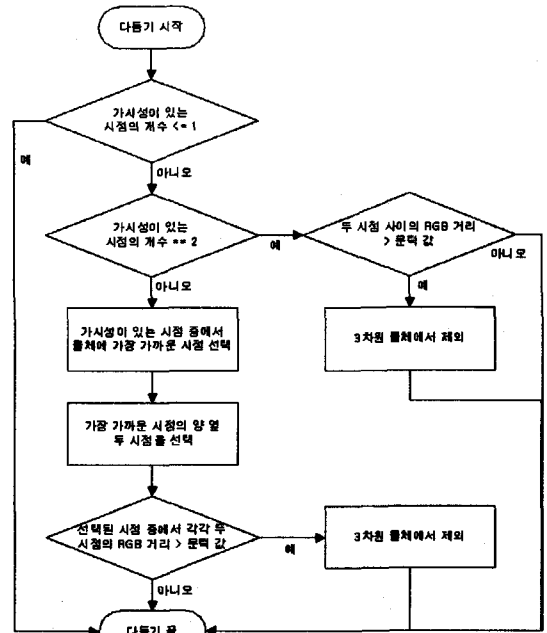


그림 8. 3차원 물체 다듬기 알고리즘

RGB 거리를 계산하기 위해서 최대 세 개의 시점만 선택을 한다. 그들 중의 하나는 3차원 표면의 점에서 가장 가까운 시점이고 나머지 둘은 가장 가까운 시점의 양 옆의 시점이다. 이는 시점과 3차원 물체 표면의 점의 거리가 가깝다면 보이는 시점에서의 색상값이 일치하는 3차원 점의 색상값을 더 잘 반영할 수 있다고 추정한 것이다. 그림 8은 3차원 물체 다듬기의 알고리즘을 보여준다.

이러한 다듬기 과정은 어떠한 두 시점의 RGB 거리가 정해진 문턱값보다 작을 때까지 계속 반복한다. 3차원 물체 다듬기 과정이 실행될 때마다 가시성 검사의 과정으로부터 가시성 정보는 갱신된다. 3차원 물체 다듬기 과정 후에 3차원 물체의 가시성 정보는 고정된다.

3.4 색상값 사상

가시성 검사를 통한 가시성 정보는 색상값 사

상에서 또한 유용하다. 이 과정에서 가시성이 있는 시점을 선택한다. 선택된 시점이 많다면 멀리 떨어진 시점들 사이에 일치하는 점들의 색상값의 차이가 많이 나므로 물체 다듬기의 과정처럼 최대 세 개의 시점을 선택한다. 이는 3차원 물체 표면에 더 가까이 있는 시점이 더 정확한 색상값을 반영한다는 사실을 이용한 것이다. 세 개의 시점은 색상값을 사상하는 3차원 물체의 점과 가장 가까운 시점과 그 양 옆의 시점이다. 시점이 선택된 후에 색상을 섞는 과정을 수행한다. 섞는 과정은 선택된 시점들과 3차원 점과의 거리에 반비례하여 구해진 식 (9)의 가중치 계수에 의해서 식 (8)로부터 결정된다.

$$\begin{aligned} R &= \alpha * R1 + \beta * R2 + \gamma * R3 \\ G &= \alpha * G1 + \beta * G2 + \gamma * G3 \\ B &= \alpha * B1 + \beta * B2 + \gamma * B3 \end{aligned} \quad (8)$$

$$\begin{aligned} \alpha &= \frac{1/v_1}{1/v_1+1/v_2+1/v_3}, & \beta &= \frac{1/v_2}{1/v_1+1/v_2+1/v_3}, \\ \gamma &= \frac{1/v_3}{1/v_1+1/v_2+1/v_3} \end{aligned} \quad (9)$$

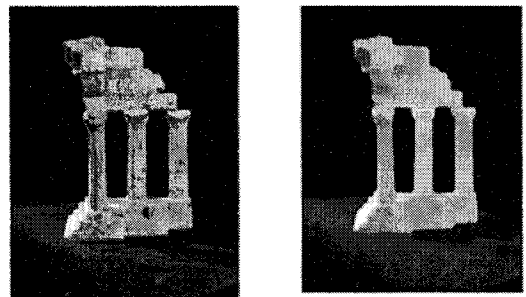
4. 실험 결과 및 분석

4.1 실험 조건

3차원 물체의 복원을 위해서 카메라 보정된 두 가지 다시점 영상을 실험에 사용하였다 [10]. 이들은 16장의 입력 영상을 사용한 “DinoSparse Ring” 다시점 영상과 10장의 입력 영상을 사용한 “TemplesparseRing” 다시점 영상이고, 카메라들을 3차원 물체 주위로 원을 그리면서 배치하여 획득하였다. 모든 다시점 영상의 크기는 640×480이고, 3차원 모델은 복셀 표현법을 사용하였다.

4.2 전처리 결과

이 논문에서 사용된 영상들의 배경은 검은색 계통이기 때문에 물체를 입력 영상들로부터 쉽게 분리할 수 있다. 그러나 물체에서 그림자가 드리워진 어두운 영역도 또한 검은색 계통이다. 따라서 이러한 영역도 배경으로 인식될 수 있다. 이 문제를 해결하기 위해서 팽창(dilation)과 침식(erosion)의 영상처리 기법을 사용하여 물체에 그림자가 드리워진 어두운 영역을 보완하였다. 그림 9는 전처리의 결과를 보여준다. 원본 영상에서 물체에 그림자가 드리워져 어두운 영역이 전처리된 영상에서는 밝은 영역으로 바뀐 것을 알 수 있다. 이로 인해서 영상의 배경과 물체를 구별할 수 있었다. 이 결과는 좀 더 정확한 3차원 물체의 초기 윤곽을 탐지하는데 사용되었다.



(a) 원본 영상 (b) 전처리된 영상

그림 9. 전처리 결과

4.3 3차원 복원의 결과

그림 10은 두 가지 다시점 영상의 입력 영상과 실험 결과를 보여준다. 실험 결과는 입력 영상과 같은 시점에서 획득한 영상이다. “DinoSparseRing”의 3차원 모델을 표현하기 위해서 복셀 해상도 184×220×94를 사용하였고, “TemplesparseRing”의 3차원 모델을 표현하기 위해서 복셀 해상도 258×402×96을 사용하였다.

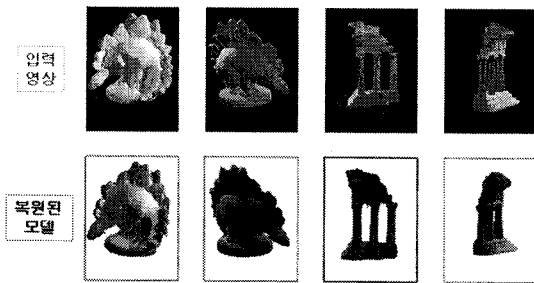


그림 10. 3차원 물체 복원 실험 결과

5. 결 론

본 논문에서는 다수의 2차원 영상을 이용한 3차원 물체 복원 방법을 제시하였다. 기본적으로 실루엣 기반 복원 방법을 사용하였고, 복셀 컬러링의 이론적 바탕을 이용하였다. 3차원 물체 복원을 위해 가시성 검사를 하였고, 이는 각 과정에서의 정확도와 효율성을 높여주었다. 또한 물체 다듬기 과정에서는 3차원 공간의 복셀을 2차원 영상에 투영하고, 투영된 모든 화소들의 상관관계를 이용하여 화소의 표본을 추출하였다. 추출된 모든 화소의 정보를 이용하여 좀 더 합리적인 복셀의 대표값을 설정할 수 있었다. 실험 결과로부터 제안하는 방법으로 3차원 물체의 모델을 복원하는 일이 유효함을 알 수 있었다.

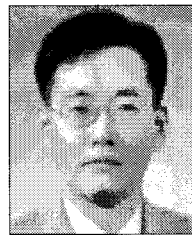
참 고 문 헌

- [1] M. Grum and A.G. Bor, "Multiple Image Disparity Correction for Scene Representation," IEEE International Conference on Image Processing, pp. 209-212, Oct. 2008.
- [2] J.M. Lavest, M. Viala, and M. Dhome, "Do we really need an accurate calibration pattern to achieve a reliable camera calibration?," 5th European Conf. on Computer Vision, Vol. 1, pp. 158-174, 1998.
- [3] Y. Lu, J.Z. Zhang, Q.M. Jonathan Wu, and L. Ze-Nian, "A Survey of Motion-Parallax-Based 3-D Reconstruction Algorithm," IEEE Transactions on System, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 34, No. 4, Nov. 2004.
- [4] L. Zhang, N. Snavely, B. Curless, and S.M. Seitz, "Spacetime Faces: High Resolution Capture for Modeling and Animation," ACM Transactions on Graphics, Vol. 23, No. 3, pp. 548-558, Aug. 2004.
- [5] W. Culbertson and T. Malzbender, "Generalized Voxel Coloring," Vision Algorithms: Theory and Practice, vol. 1883, pp. 67-74. Sept. 1999.
- [6] T. Matsuyama, X. Wu, T. Rakai, and S. Nobuhara, "Real-time 3D shape reconstruction, dynamic 3D mesh deformation, and high fidelity visualization for 3D video," Proc. of Computer Vision and Image Understanding, Vol. 96, Issue 3, pp. 393-434, July 2004.
- [7] K.M. Cheung, T. Kanade, J. Bouguet, and M. Holler, "A real time system for robust 3D voxel reconstruction of human motions," IEEE International Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 714-720, June 2000.
- [8] S. Seitz, and C. Dyer, "Photorealistic scene reconstruction by voxel coloring," International Journal of Computer Vision Vol. 35, No. 2, pp. 151-173, 1999.
- [9] C. Esteban and F. Schmitt, "Multi-Stereo 3D object reconstruction," International Symposium on 3D Processing, Visualization, and Transmission, pp. 159-166, 2002.
- [10] Middlebury "DinoSparseRing" data set and "TempleSparseRing" data set available at: <http://vision.middlebury.edu/mview/data/>.



장 우 석

- 2007년 2월 전남대학교 전자정보통신공학과 학사
- 2009년 8월 광주과학기술원 정보통신공학과 석사
- 2009년 9월 ~ 현재 광주과학기술원 정보통신공학과 박사 과정
- 관심 분야 : 디지털 영상처리, 컴퓨터 비전, 영상신호 처리 및 압축



호 요 성

- 1981년 2월 서울대학교 전자공학과 학사
- 1983년 2월 서울대학교 전자공학과 석사
- 1989년 12월 Univ. of California, Santa Barbara, Department of Electrical and Computer Engineering, 박사
- 1983년 3월 ~ 1995년 9월 한국전자통신연구소 선임연구원
- 1990년 1월 ~ 1993년 5월 미국 Philips 연구소, Senior Research Member
- 1995년 9월 ~ 현재 광주과학기술원 정보통신공학과 교수
- 관심분야 : 디지털 신호처리, 영상신호 처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV, MPEG 표준, 3차원 TV, 실감 방송