

논문 2012-49SC-1-1

색상과 깊이 카메라를 이용한 3차원 영상 구성

(3D Image Construction Using Color and Depth Cameras)

정 하 형*, 김 태 연*, 유 준**

(Ha Hyoung Jung, Tae Yeon Kim, and Joon Lyou)

요 약

본 논문에서는 색상과 깊이 카메라로 구성된 복합형 카메라 시스템을 이용하여 3차원 영상을 구성하는 방법을 제안한다. 복합형 카메라 시스템은 3차원 영상 구성에 대하여 각 카메라가 갖는 단점을 상호 보완할 수 있다. 3차원 영상 구성에 앞서 개별적인 카메라의 내부변수 및 외부변수를 추출하며 이를 이용하여 영상 정합에 필요한 두 카메라 사이의 관계를 얻는다. 또한, 깊이 카메라의 출력과 거리 관계를 실험적으로 유도하며 정합된 영상에 대하여 좌표 변환을 통해 3차원 영상을 구성한다. 제안된 기법을 구현하기 위해 마이크로소프트사의 복합형 카메라를 이용하여 실험을 수행하며 3차원 영상 및 실제 측정치와의 비교 결과를 제시함으로써 성능을 검증한다.

Abstract

This paper presents a method for 3D image construction using the hybrid (color and depth) camera system, in which the drawbacks of each camera can be compensated for. Prior to an image generation, intrinsic parameters and extrinsic parameters of each camera are extracted through experiments. The geometry between two cameras is established with these parameters so as to match the color and depth images. After the preprocessing step, the relation between depth information and distance is derived experimentally as a simple linear function, and 3D image is constructed by coordinate transformations of the matched images. The present scheme has been realized using the Microsoft hybrid camera system named Kinect, and experimental results of 3D image and the distance measurements are given to evaluate the method.

Keywords : 3D image construction, hybrid camera, depth camera, image calibration, image geometry

I. 서 론

최근 영상처리에 기반한 여러 분야에서 3차원 영상의 활용에 대한 관심이 높아지고 있는 가운데, 이를 구성하는 방법에 대한 연구가 급증하고 있다. 3차원 영상은 기존의 2차원 영상에 깊이 혹은 거리 정보가 추가된 형태이며 이러한 영상 정보는 크게 접촉식과 비접촉식 방식으로 획득할 수 있다.

접촉식 방식은 주로 3차원 좌표 측정기(Coordinate Measuring Machine, CMM) 등을 이용하여 대상 사물

의 표면을 이동하면서 3차원 공간 좌표를 검출 하는 방식으로 측정 정확도가 높은 장점이 있다. 하지만 측정 속도가 느리고 측정 대상의 크기 및 위치에 따라 응용이 제한된다. 이에 비해 비접촉식 방식은 사용되는 센서의 종류와 측정 기법에 따라 차이가 있지만 대체로 측정 대상의 크기 및 위치에 관계없이 고속으로 깊이 및 거리 정보를 획득할 수 있다.^[1] 비접촉식 방식 중 레이저 거리계와 카메라를 이용한 방식은 수평 거리를 측정하는 레이저 거리계를 기구적으로 회전시켜 3차원 공간에 대한 거리를 획득하고 이를 출력 영상과 융합하는 방식으로써 측정 속도 및 정확도가 뛰어나고 기법이 간단하다. 그러나 정확한 3차원 영상 구성을 위해 반복적인 실험이 요구되며 회전 기구에 의해 부피가 커지고 고장에 취약한 단점이 있다.^[2-3]

* 학생회원, ** 평생회원-교신저자, 충남대학교
전자공학과
(Dept. of Electronics Engineering, Chungnam
National University)
접수일자: 2011년11월5일, 수정완료일 2011년12월14일

위의 방법과 달리 일반 카메라로부터 획득한 영상만을 이용하여 깊이 혹은 거리 정보를 추출하고 3차원 영상을 구성하는 방법이 연구되고 있는데, 대표적으로 스테레오(Stereo) 기법과 다중시점(Multi-view) 기법을 들 수 있다. 스테레오 기법은 인간의 눈과 같이 수평 정렬된 두 카메라로부터 얻은 영상의 시차(Disparity)를 이용하여 대상 환경 및 물체에 대한 깊이 정보를 추출하며, 다중시점 기법은 시점이 다른 다수의 영상에 대해 기하 해석을 적용하여 깊이 정보를 계산한다. 이러한 방법들은 앞서 언급한 부가적인 센서와 기구가 필요 없으므로 실용적이다. 하지만 영상처리만을 이용한 3차원 영상 구성 과정은 높은 연산 능력을 요구한다.^[4~7]

한편, 대상에 인위적으로 에너지를 투사하고 이의 변화를 감지하여 3차원 형상을 복원할 수 있는 깊이 카메라가 이용되고 있다. 깊이 카메라는 크게 초음파를 투사하여 대상 물체에 반사되어 돌아오는 시간을 계산함으로써 영상의 깊이를 측정하는 시간지연 방식과 일정한 패턴의 적외선을 대상 물체에 조사하고 패턴의 변화를 이용하여 깊이를 측정하는 구조광방식이 있는데, 이들은 모두 일반적인 카메라에 비해 해상도가 낮고 색상을 표현할 수 없으므로 3차원 영상 구성을 위해 단독으로 사용될 수 없다.^[8~10]

본 논문에서는 일반 색상 카메라와 구조광 방식의 깊이 카메라를 결합한 복합형 카메라를 이용하여 3차원 영상을 구성한다. 복합형 카메라는 3차원 영상 구성에 있어 일반 카메라와 깊이 카메라의 단점을 상호 보완할 수 있는데, 이를 위해 먼저 각각의 카메라로부터 내부변수 및 외부변수를 추출하고 두 카메라 사이의 관계를 알아냄으로써 서로 다른 출력 영상을 정합한다. 또한, 깊이 카메라의 출력을 실제 거리 정보로 변환하기 위한 실험적 방법을 제안하며 이로부터 깊이와 거리의 비선형 관계를 유도하고 정합된 영상을 3차원 영상으로 구성한다. 제안된 방식의 성능을 입증하기 위해 복합형 카메라를 이용하여 실험을 수행하며 전 수행 과정에서의 결과 및 실제 측정치와의 비교 결과를 제시한다.

이하 본 논문의 구성은 다음과 같다. II장에서는 복합형 카메라의 보정에 대하여 언급하고 III장에서는 3차원 영상 구성을 위한 깊이와 거리 관계 유도와 좌표변환에 관하여 기술한다. 그리고 IV장에서는 실험 결과를 제시하며 결론 및 추후 연구내용은 V장에서 언급한다.

II. 복합형 카메라 보정

그림 1은 마이크로소프트사의 복합형 카메라 시스템으로써 동일한 수평축에 일반 컬러 카메라와 적외선 영상센서 및 적외선 조사기로 구성된 구조광 방식의 깊이 카메라가 정렬되어 있으며, 내부에 일정하게 조사된 적외선 패턴의 구조적 변화를 처리하여 깊이 정보를 추출하는 처리장치가 포함되어 있다.

본 장에서는 이러한 복합형 카메라를 이용하여 3차원 영상을 구성하기 위해 서로 다른 특성을 갖는 두 카메라로부터 내부변수 및 외부변수를 추출하고 두 카메라의 관계를 획득하는 전처리 과정을 기술하였다.

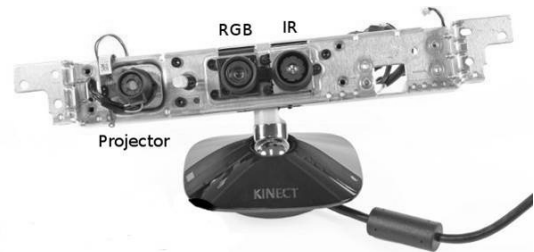


그림 1. 마이크로소프트 복합형 카메라 시스템
Fig. 1. Microsoft hybrid camera system.

1. 내부변수 추출

일반적으로 영상은 3차원 환경이 2차원 평면에 투영된 것으로 그림 2와 같은 핀홀 카메라 모델을 이용하여 관계를 표현할 수 있다.^[5] 카메라 좌표계에서 표현된 3차원 공간상의 임의의 점 $P(X,Y,Z)$ 는 영상평면 좌표계의 점 $p(u,v)$ 로 대응되며 영상평면 좌표계는 카메라 좌표계와 초점거리 f 만큼 떨어져 있다.

이에 따라 점 P 에서 점 p 로의 투영 변환(Projective transform)은 식 (1)와 같이 나타낼 수 있다.

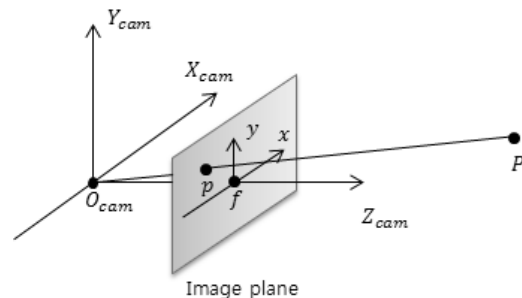


그림 2. 핀홀 카메라 모델
Fig. 2. Pinhole camera model.

$$\frac{Z}{f} \begin{pmatrix} u \\ v \\ f \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (1)$$

하지만 실제 카메라 좌표계의 중심과 영상평면 좌표계의 중심은 일치하지 않으므로 중심 위치의 차이를 고려하여 식 (1)은 식 (2)과 같이 나타낼 수 있다.

$$Z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (2)$$

$$A = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{Intrinsic parameter})$$

여기서 f_x, f_y 는 x 축과 y 축의 카메라 초점거리, c_x, c_y 는 중심 위치의 차이를 나타내며 이들로 구성된 행렬 A 는 카메라의 내부변수로 정의된다. 이러한 내부변수를 추출하기 위해 역시 다수의 격자무늬판 영상을 이용하였으며 표 1은 실험적으로 획득한 각 카메라의 내부변수이다.

표 1. 각 카메라의 내부변수

Table 1. Intrinsic parameters of each camera.

일반 카메라		깊이 카메라	
f_x	533.360	f_x	596.659
f_y	532.686	f_y	595.564
c_x	301.052	c_x	311.531
c_y	248.348	c_y	241.726

2. 외부변수 추출

내부변수로부터 카메라 좌표계와 영상평면 좌표계의 관계를 획득할 수 있지만 3차원 공간상의 대상의 위치는 실세계 좌표계로 표현되며 이것은 그림 3과 같이 카

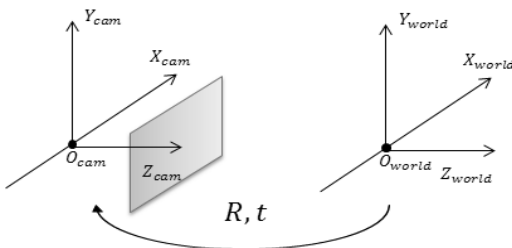


그림 3. 카메라 좌표계와 실세계 좌표계 관계
Fig. 3. Camera and world coordinates system.

메라 좌표계와 일치하지 않고 방향 또한 같지 않다.^[9]

R 을 3×3 크기의 회전행렬, t 를 3×1 크기의 이동벡터로 정의하면 카메라 좌표계의 점 $P(X, Y, Z)$ 와 실세계 좌표계의 점 $P_w(X_w, Y_w, Z_w)$ 의 관계는 식 (3)과 같이 표현될 수 있으며 이때 행렬 $[R|t]$ 를 카메라의 외부변수라고 한다.

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = [R|t] \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (3)$$

$$[R|t] = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \quad (\text{Extrinsic parameter})$$

따라서 실세계 좌표계의 점 P_w 는 식 (4)와 같이 영상평면 좌표계로 대응된다.

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} [R|t] \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (4)$$

3. 카메라 관계 추출

두 카메라는 동일한 수평축에 정렬되어 있지만 렌즈 화각과 같은 물리적인 특성이 서로 다르므로 3차원 영상 구성에 앞서 두 영상을 정합하기 위해 그림4 과 같이 카메라 사이의 관계를 추출해야 한다.

이를 위해 역시 격자무늬판 영상을 이용하여 동일한 표본에 대한 각 카메라의 외부변수를 추출하고 식 (5)을 이용하여 카메라 관계를 표현하는 회전행렬과 이동 벡터를 구할 수 있다.

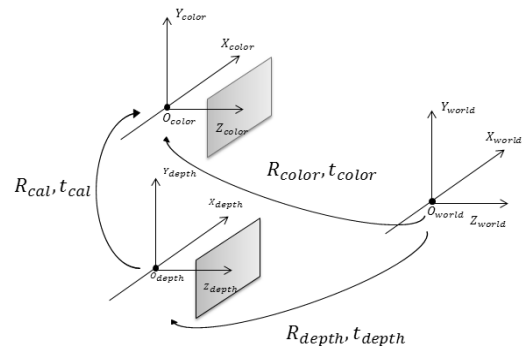


그림 4. 일반 카메라와 깊이 카메라 관계
Fig. 4. Color and depth camera coordinates system.

$$\begin{aligned} R_{cal} &= R_{depth}^T R_{color} \\ t_{cal} &= t_{depth} - t_{color} \end{aligned} \quad (5)$$

여기서 R_{color} , t_{color} 과 R_{depth} , t_{depth} 는 각각 일반 컬러 카메라와 깊이 카메라의 회전행렬 및 이동벡터를 나타내고 R_{cal} , t_{cal} 은 깊이 카메라를 기준으로 한 회전행렬 및 이동벡터를 나타낸다.

한편, 카메라의 외부변수는 영상에 따라 다르므로 이를 이용하여 계산된 두 카메라의 관계 역시 고정된 값이 아니다. 이에 따라 본 논문에서는 다양한 시점의 영상 및 표본을 통해 관계를 구하고 이들의 평균을 취해 다음과 같이 카메라 관계를 획득하였다. 이는 두 카메라의 물리적 특성 자체는 대상 영상에 따라 변하지 않는다는 사실에 기인한다.

$$\begin{aligned} R_{cal} &= \begin{bmatrix} 1 & -0.0044 & -0.0087 \\ 0.0044 & 1 & 0.0048 \\ 0.0087 & -0.0048 & 1 \end{bmatrix} \\ t_{cal} &= [-0.0264 \quad 0.0001 \quad -0.0034] \end{aligned} \quad (6)$$

그리고 그림 5는 지금까지의 기술한 복합형 카메라 보정 과정의 수행 전·후 정합된 영상을 나타낸다.

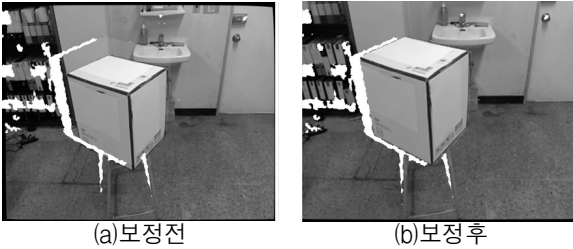


그림 5. 각 카메라 정합 영상
Fig. 5. Matched images of each camera.

III. 3차원 영상 구성

본 장에서는 깊이 카메라의 출력과 실제 거리 정보의 관계를 유도하는 실험적 방법에 대하여 기술하며 아울러 정합된 영상으로부터 3차원 영상을 구성하기 위해 깊이 카메라 좌표계를 기준으로 하는 두 영상의 좌표 변환에 대하여 기술하였다.

1. 깊이와 거리 관계 유도

깊이 카메라의 출력은 각 화소에 대하여 11비트의 이

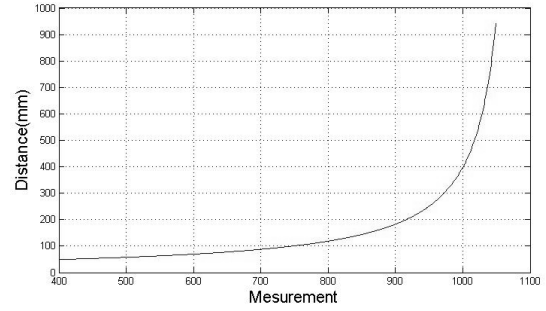


그림 6. 깊이 대 거리 그래프
Fig. 6. Graph between depth information and distance.

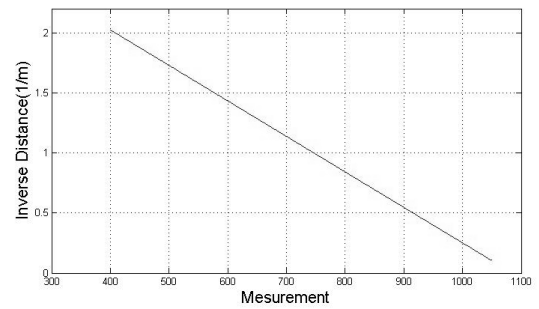


그림 7. 역함수 그래프
Fig. 7. Inverse distance graph.

산적인 값을 가지고 있다. 이로부터 대상과의 거리를 알기 위해 본 논문에서는 동일한 수평축에 레이저 거리계와 카메라를 배치하고 벽면과의 거리를 증가시키면서 데이터를 획득하였다.

그림 6은 실험적으로 획득한 깊이 대 거리 그래프로써 비선형성을 가지므로 그 관계를 함수로 정의하기 어렵다. 이에 깊이와 거리 관계를 분수함수 형태로 가정하고 분수함수의 역은 선형성을 나타낸다는 점에 기인하여 그림 7의 역함수 그래프를 근사화 하였다.

식 (7)은 유도하고자 하는 깊이 대 거리 관계로써 d 는 거리를 나타내고 m 은 11비트 깊이 값이다.

$$d = \frac{1}{a \cdot m + b} \quad (7)$$

그리고 이의 역은 식 (8)과 같으며 최소자승법을 이용하여 계수 a 와 b 를 구하였다.

$$\frac{1}{d} = a \cdot m + b \quad (8)$$

$$a = -0.002955, \quad b = 3.206$$

2. 깊이 영상의 좌표 변환

깊이 카메라 좌표계를 기준으로 하였을 때 깊이 영상의 좌표는 식 (9)과 같이 카메라 좌표계로 변환할 수 있다. 이는 식 (1)으로부터 유도되며 보정 과정에서 추출한 내부변수와 거리 정보를 이용하였다.

$$\begin{aligned} X_{depth} &= (u_{depth} - c_{x,depth}) \times \frac{Z}{f_{x,depth}} \\ Y_{depth} &= (v_{depth} - c_{y,depth}) \times \frac{Z}{f_{y,depth}} \\ Z_{depth} &= d \end{aligned} \tag{9}$$

3. 컬러 영상의 좌표 변환

일반 카메라 영상의 좌표는 마찬가지로 보정 과정에서 획득한 카메라 관계를 이용하여 식 (10)와 같이 변환 할 수 있다.

$$\begin{pmatrix} X_{color} \\ Y_{color} \\ Z_{color} \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_{depth} \\ Y_{depth} \\ Z_{depth} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{10}$$

여기서 $r_{ij}(i=1\sim3, j=1\sim3)$ 와 t_x, t_y, t_z 는 각각 회전행렬 및 이동벡터의 요소를 나타낸다. 그리고 컬러 영상의 좌표는 식 (11)과 같이 카메라 좌표계로 변환할 수 있다.

$$\begin{aligned} u_{color} &= X_{color} \times \frac{f_{x,color}}{Z_{color}} + c_{x,color} \\ v_{color} &= Y_{color} \times \frac{f_{y,color}}{Z_{color}} + c_{y,color} \end{aligned} \tag{11}$$

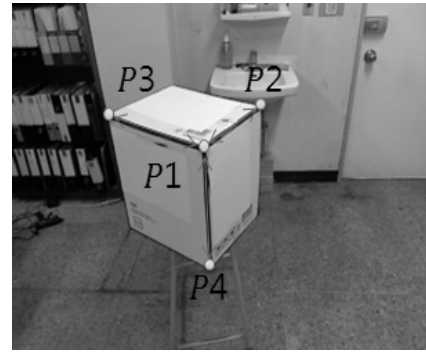
IV. 실험 및 고찰

실험은 동일한 실내 환경에서 수행하였으며 서로 다른 대상이 놓여 있는 방향에 대하여 각각 3차원 영상을 구성하고 영상 내에서 표본점을 추출하여 이들의 거리와 실제 측정치를 비교함으로써 결과를 검증하였다. 실험에 사용된 복합형 카메라의 깊이는 3장의 깊이와 거리 관계 그래프에서 3m 이상의 거리에 대해 분해능이 급격히 감소함을 확인하였으므로 실험 환경 내에서 카메라의 위치를 대상과 5m 범위 이내로 한정하였다.

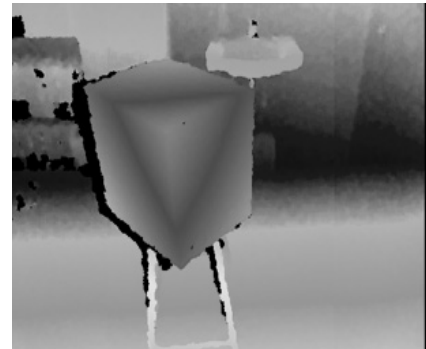
표 2. 영상 및 측정치 비교 결과 1

Table 2. Distance comparison result 1.

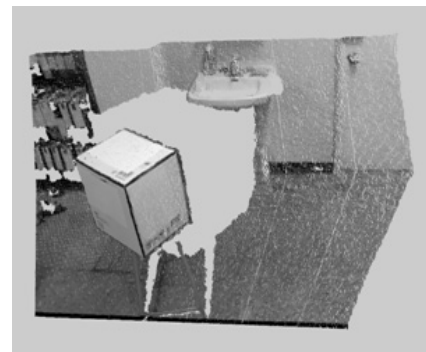
표본	실제거리	3차원 영상 거리			
		1	2	3	4
P1, P2	0.30	0.288	0.282	0.289	0.289
P1, P3	0.36	0.354	0.333	0.350	0.351
P1, P4	0.38	0.360	0.352	0.361	0.365



(a) 컬러영상



(b) 깊이영상



(c) 3차원 영상

그림 8. 실내환경 시험 결과

Fig. 8. Indoor experimental results.

그림 8은 일반 카메라 영상(a), 깊이 카메라 영상(b), 그리고 이를 바탕으로 구성한 3차원 영상(c)을 나타낸다. 깊이 영상은 거리가 가까울수록 붉은색, 멀어질수록 푸른색으로 표현하였으며 3차원 영상은 입체감을 위해

조감도로 나타내었다. 특히, 각 영상에서 박스 형태의 물체는 비교적 근거리에서 위치하며 뒷면의 경우 깊이 카메라의 적외선이 조사될 수 없으므로 3차원 영상에서 표현될 수 없었다.

표 2는 각각의 3차원 영상에서 특정한 점 ($P1 \sim P4$)을 선택하여 추출한 거리와 실제 대상으로부터 측정된 거리를 비교한 결과로써 두 실험 모두 최대 5cm 이내의 차이를 갖고 있음을 확인하였다.

V. 결론 및 추후 연구 내용

본 논문에서는 복합형 카메라를 이용한 3차원 영상 구성 방법을 제안하였다. 이러한 복합형 카메라는 일반 카메라와 깊이 카메라를 결합한 시스템으로써 3차원 영상 구성에 있어 일반 카메라와 깊이 카메라가 갖는 단점을 상호 보완할 수 있다. 이를 위해 개별적인 카메라 영상에 대한 내부변수와 외부변수를 추출하였으며 두 카메라 사이의 관계를 획득하여 서로 다른 특성의 영상을 정합하였다. 또한, 깊이 카메라의 출력과 거리 관계를 실험적인 방법을 통해 비교적 간단하고 정확한 선형 함수로 근사화 하였으며 이를 3차원 영상 구성에 적용하였다.

또한, 본 논문에서는 마이크로소프트사에서 제작된 복합형 카메라를 이용하여 제안된 기법을 구현하였고 실내 환경에서의 실험을 통해 3차원 영상 결과 및 실제 측정치와의 비교 결과를 제시함으로써 기법의 성능을 검증하였다. 하지만 구성된 3차원 영상은 화소간 미세한 거리 차이와 깊이 카메라에 의해 조사된 적외선의 불균일함으로 인하여 질감이 낮을 뿐 만 아니라 점군(Point cloud) 형태로 표현되어 있어 데이터 크기가 매우 크다. 이에 따라 향후 형태학적 처리(Morphological processing) 및 데이터 축소를 위한 벡터 그래픽 처리에 관한 연구를 수행할 예정이다.

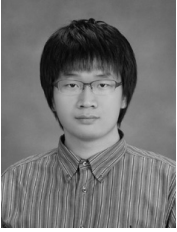
참 고 문 헌

- [1] 이승표, 강형주, 하성규, “3차원 좌표 측정기의 측정 성능에 대한 실험적 연구,” 한국정밀공학회지, 제26권, 제3호, 47-54쪽, 2009년 3월
- [2] 김정훈, 임선, 김봉선, 정일균, “학습 메카니즘을 적용한 3D LRF 수평 속도 제어 구현,” 대한전자공학회 2010년 하계종합학술대회, 1306-1308쪽, 2010년 6월
- [3] 정지훈, 안광호, 강정원, 김우현, 정명진, “무인 차

량의 자율 주행을 위한 2차원 레이저 거리 센서와 카메라를 이용한 입방형 격자 기반의 3차원 지형형상 복원,” 전자공학회논문지, SC편, 제45권, 6호, 26-34쪽, 2008년 11월

- [4] 이정수, 양승구, 김준성, “실시간 스테레오 비전 시스템을 위한 SAD 정합연산기 설계,” 대한전자공학회 논문지 CI편, 제45권, 제1호, 55-61쪽, 2008년 1월
- [5] 이남호, 박순용, 이승민, “스테레오 카메라를 이용한 3차원 정보의 가시화,” 전자공학회논문지 SC편, 제47권, 4호, 15-20쪽, 2010년 7월
- [6] 임영철, 이충희, 권순, 이종훈 “스테레오 비전에서 향상된 적응형 퍼지 칼만 필터를 이용한 거리 추정 기법,” 전자공학회논문지 SC편, 45권, 6호, 106-116쪽, 2008년 11월.
- [7] 유성구, 정길도, “단일 영상과 거리센서를 이용한 SLAM 시스템 구현,” 전자공학회논문지 SC편, 45권, 6호, 149-156쪽, 2008년 11월.
- [8] 김태준, 장은영 허남호, 김진웅, 유지상, “다시점 카메라 모델의 기하학적 특성을 이용한 가상시점 영상 생성 기법,” 한국통신학회논문지, 제34권, 제12호, 1147-1153쪽, 2009년 12월
- [9] 엄기문, 안충현, 이수인, 김강연, 이관행, “깊이정보 카메라 및 다시점 영상으로부터의 다중 깊이맵 융합기법,” 방송공학회논문지, 제9권, 제3호, 185-195쪽, 2004년 9월
- [10] Yan Cui, Schuon, S, Chan, D, Thrun, S.; Theobalt, C, “3D shape scanning with a time-of-flight camera” in Conf of IEEE CVPR, pp.1173-1180, June 2010.

— 저 자 소 개 —



정 하 형(학생회원)
 2010년 충남대학교
 전자전파정보통신공학과
 학사 졸업.
 2010년~현재 충남대학교 대학원
 전자공학과 석박사과정.
 <주관심분야 : 영상처리, 제어시
 스템, 자율이동로봇 등>



유 준(정회원)
 1978년 서울대학교 전자공학과
 학사 졸업.
 1984년 한국과학기술원 전기전자
 공학과 박사 졸업.
 1984년~현재 충남대학교 전자
 공학과 교수
 <주관심분야 : 산업공정제어, 센서신호처리, IT기
 반 로봇, 항법시스템>



김 태 연(학생회원)
 2006년 충남대학교 전자공학과
 학사 졸업.
 2006년~현재 충남대학교 대학원
 전자공학과 석박사과정.
 <주관심분야 : 센서신호처리, 이
 동로봇, 무인헬리콥터, 실내항법>