

로봇 팔에 부착된 카메라를 이용한 3차원 측정방법

이 상 용*

한 민 홍**

*포항공과대학 산업공학과

**고려대학교 산업공학과

Axial Motion Stereo Method

Sang-Yong Rhee

Min-Hong Han

Department of Industrial Engineering

Department of Industrial Engineering

Pohang Institute of Science and Technology

Korea University

ABSTRACT

This paper describes a method of extracting the 3-D coordinates of feature points of an object from two images taken by one camera. The first image is from a CCD camera before approaching the object and the second image is from same camera after approaching the object along the optical axis.

In the two images, the feature points appear at different position on the screen due to image enlargement. From the change of positions of feature points their world coordinates are calculated. In this paper, the correspondence problem is solved by image shrinking and correlation.

1. 서론

3차원 정보를 얻기 위한 일반적인 camera modeling은 두 대의 카메라를 인간의 시각 체계처럼 나란히 놓는 것으로 이러한 camera modeling을 lateral stereo model이라고 한다. 이 모델은 두 대의 카메라에 동시에 보이는 물체를 삼각측량을 이용하여 거리를 계산하는데, 어떤 장애물에 가리거나 혹은 두 카메라 사이의 간격이 크므로 인하여 물체가 카메라 한 대에만 보이는 경우가 종종 발생한다.

이러한 문제는 한 대의 카메라를 사용한다면 해결이 가능하다. 즉, 물체에 접근하기 전에 한 장의 영상 (back image)를 얻고, 카메라의 시축을 따라 이동한 후에 또 하나의 영상 (front image)을 얻음으로서 3차원 정보를 추출할 수 있다. 이러한 camera modeling을 axial motion model이라고 하는데 [1], 이 분야에 대한 연구는 그리 많지 않다.

William [6]은 도로를 따라서 이동하면서 도로주변의 실제 이미지 (real world scene) 두 장으로부터 알고자 하는 물체까지의 거리를 계산하는 방법을 제시하였고, Itoh et al. [3]는 직육면체를 대상물체로 하여 각 꼭지점까지의 3차원 좌표를 계산하였다. 이들의 연구들은 correspondence problem에 대해서는 고려하고 있지 않다. Jain et al. [4]는 correspondence problem을 풀기위하여 image를 complex logarithmic space로 mapping하고 이 space의 수평선을 따라서 탐색하는 방법을 사용하였으며, Alvertos et al. [1]는 radiance를 사용하여 correspondence problem을 해결하고자 하였다.

본 연구에서는 correspondence problem을 해결하기 위하여 normalized correlation을 이용하는 방법을 제시한다. 거리를 알고자 하는 물체 특징점의 back image에서의 image point 와 상관계수가 가장 높은 화소를 front image에서 찾아내는 것이다. 본 연구에서 카메라는 로봇 팔에 장착되어 이동되는데, 로봇팔이 장착된 카메라의 시축 (optical axis) 방향으로 정확하게 움직인다고 가정한다.

본 연구는 다음과 같이 구성된다. 2장에서는 axial motion model에서 3차원 정보를 추출하는 과정을 소개하고, 3장에서는 실험내용을 제시하며, 4장에서는 결론및 추후 연구방향을 제시한다.

2. Axial motion stereo camera model

Axial motion model은 depth-in-motion model이라고도 하며 [6], 한 대의 카메라를 이용하여 첫번째 영상 (back image)은 물체에 접근하기 전에, 두번째 영상 (front image)은 카메라가 optical axis를 따라서 물체에 접근하고난 후에 얻는다. 이 때에 얻어진 두 개의 영상을 비교하여 보면, 물

체 특징점들의 image point 위치가 바뀌어 있음을 알 수 있고, 이 image point에서의 위치의 변화 정도를 이용하여 물체 특징점의 3차원 좌표를 계산하게 된다.

2.1 Axial motion model을 이용한 3차원 정보의 추출

그림 1에서 두 대의 카메라 CF와 CB는 Z₀ 방향으로 Δz만큼 떨어져 있다. 이 두 카메라 CF와 CB의 화면좌표계를 각각 (X_s^F, Y_s^F)와 (X_s^B, Y_s^B)라고 하고, 각 좌표계의 원점을 각각 O_s^F와 O_s^B라고 하자. 이때에 3차원 공간 상의 한 점 p^F(x_e^F, y_e^F, z_e^F)가 back image에 투영된 좌표를 (x_s^B, y_s^B), front image에 투영된 좌표를 (x_s^F, y_s^F)라고 하고, 이 공간 상의 점의 3차원 좌표를 계산하기 위하여 기호를 다음과 같이 정의한다.

- h : 점 p^F와 optical axis와의 거리
- l₁ : back image에 점 p^F가 투영된 화소 p_s^B(x_s^B, y_s^B)의 image center O_s^B로부터의 거리:
- l₂ : front image에 점 p^F가 투영된 화소 p_s^F(x_s^F, y_s^F)의 image center O_s^F로부터의 거리:
- d : focal length:
- Δz : 카메라가 이동한 거리:

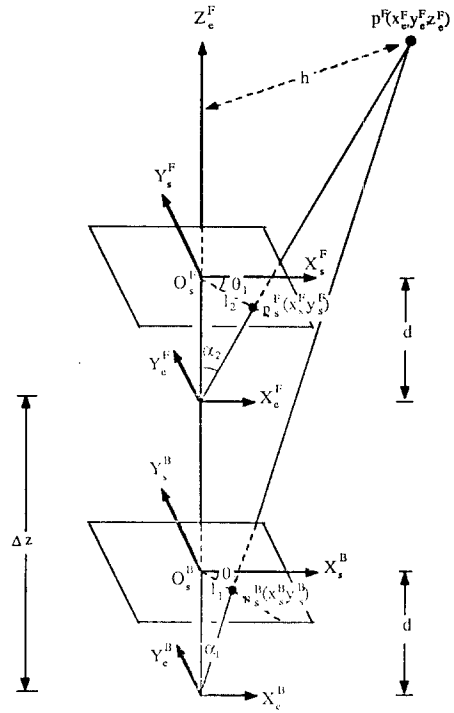


그림 1. Axial motion stereo camera model

그림 1에서 다음의 식이 성립된다.

$$h = (\Delta z + z_e^F) \tan(\alpha_1) \quad (1)$$

$$h = z_e^F \tan(\alpha_2) \quad (2)$$

여기에서 α_1 과 α_2 는 각각 $\tan^{-1}(l_1/d)$, $\tan^{-1}(l_2/d)$ 이다. 이제 z_e^F 는 다음과 같이 구할 수 있다.

$$\begin{aligned} z_e^F &= \frac{\Delta z \cdot \tan(\alpha_1)}{\tan(\alpha_2) - \tan(\alpha_1)} \\ &= \frac{\Delta z \cdot l_1}{l_2 - l_1} \end{aligned} \quad (3)$$

이 값을 이용하여 eye coordinate system의 좌표값 x_e^F, y_e^F 는 다음과 같이 구할 수 있다.

$$x_e^F = h \cos(\theta_1) \quad (4)$$

$$y_e^F = h \sin(\theta_1) \quad (5)$$

단, $\theta_1 = \tan^{-1}(x_B/y_B) = \tan^{-1}(x_F/y_F)$ 이다.

2.2 Correspondence problem의 해결 방법

만일 back image에서의 물체 특징점의 image point의 위치와 front image에서의 위치를 알고 있다면, 앞 절에서 유도된 식에 의해서 물체 특징점의 3차원 좌표값을 계산할 수 있다. 그러나 실제로는 back image, front image 두 개의 영상에서 대응이 되는 물체 특징점의 image point, 즉 conjugate pair를 알아내어야 하는 과정이 선행되어야 한다.

이러한 것을 correspondence problem이라고 하는데, 본 연구에서는 이 correspondence problem을 해결하기 위하여 상관계수를 이용한다. 즉, back image의 한 화소와 front image의 한 화소가 conjugate pair인가를 판단하기 위하여 그 화소들의 주위에 있는 화소들을 비교한다는 것이다. 그 두 화소 주위에 있는 화소들의 밝기가 같다면 이 두 개의 화소를 conjugate pair라고 할 수 있다.

2.2.1 Normalized Correlation

상관계수는 이미지의 한 부분과 모델, 템플릿(template), 또는 커널(kernel)이라고 불리는 기준패턴(reference pattern)과의 유사성의 정도를 나타내는 척도로

서 인식되어져 왔다. 유사성의 정도를 측정하여 기준패턴과 가장 유사한 부분을 우리는 찾아낼 수 있다. 여기에서 가장 유사하다는 의미는 넓은 의미에서 가장 상관계수가 높은 곳을 의미하며, 좁은 의미에서는 주위 화소들의 상관계수를 보다 높은 곳을 의미한다.

Back image의 model window의 화소값들을 M_i 라고 하고, front image window의 화소값들을 T_i 라고 할 때, 상관계수 r 은 다음과 같이 계산된다.

$$r = \frac{[N \sum T_i M_i - \sum T_i \cdot \sum M_i]}{\sqrt{[N \sum T_i^2 - (\sum T_i)^2] [N \sum M_i^2 - (\sum M_i)^2]}} \quad (6)$$

이때의 r 의 값은 항상 $-1 \leq r \leq 1$ 이다. 여기에서 1이란 값은 back image의 모델 패턴과 front image의 window의 패턴의 완전한 일치 (perfect match)를 의미한다.

그런데 유사한 정도를 나타내기 위해서 상관계수 r 대신에 r^2 을 사용하기도 한다. 여기는 두가지 이유가 있는데 첫째는 분모의 제곱근을 구하는 것보다 분자를 제공하는 것이 계산속도가 빠르기 때문이다. 그러나 이렇게 상관계수 r 을 무조건 제공하는 것은 부호에 대한 정보를 잃어버리기 때문에 좋은 방법은 아니다. 또하나의 장점은 r^2 은 크게 의미가 없는 상관계수가 낮은 영역의 값을 더욱 줄여준다는 것이다. 본 연구에서는 r^2 의 장점을 살리면서 부호에 대한 정보를 무시하지 않는

$$\max(r, 0)^2 \quad (7)$$

을 사용하는데 이것을 normalized correlation이라고 부른다.

2.2.2 대응점 (corresponding point)의 탐색

카메라가 어떤 scene의 영상을 멀리서 얻고, optical axis를 따라 접근한 후에 또 하나의 영상을 얻어 비교하여 보면, 화면상에서 scene의 요소들이 이동해 있다는 사실을 알 수 있다. Scene의 요소들의 이동 정도는 scene을 구성하고 있는 요소들과 카메라 사이의 거리, 카메라가 이동한 거리, scene을 구성하는 요소의 화면상의 위치에 따라 달라지는데, 카메라에 가까이에 있는 scene 요소들은 카메라로부터 멀리 있는 scene 요소들보다 상대적으로 많이 움직이고, 또 카메라의 이동거리가 클수록, scene을 구성하는 요소가 화면의 가장자리에 위치할수록 위치 변동이 커진다. 이러한 현상을 이용하여 우리는 관찰자로부터 scene을 구성하고 있는 각 요소들까지의 거리를 계산할 수 있다.

카메라가 어떤 scene에 접근하면 화면에 나타난 scene의 모든 요소들은 하나의 점으로부터 밖으로 향해 움직인다 [그림 2]. 이 점을 확장 초점 (focus of expansion : FOE)라고 부르는데, 이 FOE는 3차원 정보를 추출하는데 중요한 역할을 한다. 카메라의 움직임이 회전이동 없이 위치이동 뿐이며, 그 위치이동의 방향을 알고 있다면, FOE의 화면상의 위치 계산이 가능하다.

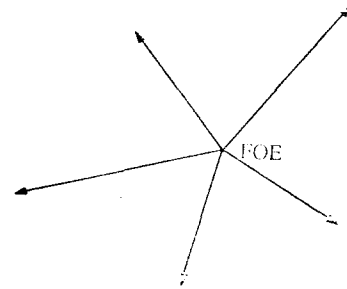


그림 2. 확장 초점

만일 카메라가 시속방향으로 이동한다면, FOE의 위치는 image center와 동일하며, scene을 구성하는 요소들은 FOE를 중심으로 방사선 방향 (radial direction)으로 이동한다는 사실을 알 수 있다. 즉, back image의 어떤 화소와 대응되는 front image의 화소는 front image의 방사선 방향으로 위치 이동하여 있다는 것이다.

따라서 우리가 3차원 좌표를 알고자 하는 물체 특징점의 conjugate pair를 찾는 방법론은 다음과 같이 요약될 수 있다. 만일 back image에서 3차원 좌표를 알고자 하는 물체 특징점의 위치 ps^B 를 알고 있다면, front image에서 방사선 방향으로 ps^B 보다 바깥쪽에 있는 한 화소를 선정하여, 이 화소가 ps^B 의 대응점이라는 가설을 세우고 normalized correlation을 계산하여, 이제까지 계산된 normalized correlation 보다 크다면 가설을 채택하고, 반대의 경우라면 가설을 기각하되, 방사선 방향의 다른 화소를 선택하여 동일한 절차를 반복한다는 것이다. front image에서 대응점을 선택할때, 본 연구에서는 보다 효율적인 탐색을 위하여 line search method를 이용한다.

이러한 방법론 아래서 유의해야 될 것 중에서 하나는 front image에서는 back image의 model window와 동일한 크기의 window를 사용해서는 안된다는 것인데, 그 이유는 scene의 구성요소들의 이미지가 확대되었기 때문이다. 따라서 front image에서 사용하는 window의 크기는 back image의 model window의 크기보다 커야 한다. front image에서의

window 크기는 다음 식에 의해서 계산된다.

$$\begin{aligned} \text{window size of front image} \\ = [\text{window size of back image} \times l_2/l_1] \end{aligned} \quad (8)$$

여기에서 [x]는 x와 같거나 혹은 x보다 작은 최대 정수를 의미한다.

어떤 두 자료의 상관계수를 구하기 위해서는 두 자료의 수가 같아야 하는데, 두 window의 크기가 다르면 window안의 화소의 갯수도 다르다. 따라서 front image의 확장된 window의 자료의 갯수를 model window와 자료의 갯수를 같게 해주기 위하여 본 연구에서는 front image의 window를 동일한 크기의 영역으로 분할하였다. 예를 들어 model window의 크기가 9 x 9 이고, 위의 식(8)에 의해서 계산된 결과가 15 x 15 라면, 15 x 15의 window를 가로 세로를 각각 9등분하였다.

이 경우에 그림 3처럼 한 영역은 여러개의 화소에 걸쳐서 형성되는데 이 영역에 대한 gray level은 이 영역이 걸쳐 있는 여러 화소의 가중평균치를 사용하였다. 즉, 한 영역에 속한 각 화소의 면적과 그 화소의 gray level을 곱하여 합을 구하고 면적으로 그 합을 나누었다.

$$g'_{ij} = \frac{\sum_{i,j \in A} a_{ij} g_{ij}}{\sum_{i,j \in A} a_{ij}} \quad (9)$$

단, g'_{ij} 는 새로운 영역 A의 gray level이고, g_{ij} 는 현재 각 화소의 gray level이며, a_{ij} 는 새로 정의된 영역 A에 포함된 현재 각 화소의 면적이다.

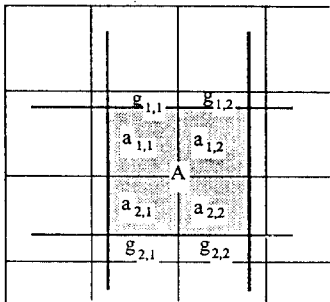


그림 3. 새롭게 정의된 영역

위의 절차는 다음과 같이 요약된다.

단계 1 : back image에서 $p_s F$ 를 선택하여 l_1 을 계산하고 model window의 화소들의 gray level을 읽음.

단계 2 : 카메라의 시축 방향으로 Δz 만큼 이동한 후,

front image를 읽음.

단계 3 : front image에서 방사선 방향의 한 화소를 대응점이라고 가정, l_2 를 계산하고 window의 크기를 결정 한 후 window내의 화소들의 gray level을 읽음.

단계 4 : front image window의 영역 재분할.

단계 5 : normalized correlation를 구하고 최대 normalized correlation와 비교.

만일 최대 normalized correlation보다 크면, 최대 normalized correlation을 수정하고 단계 3으로, 작다면 그냥 단계 3으로.

3. 실험 및 결과

Axial motion stereo camera model은 3차원 좌표값을 필요로 하는 로봇에게 응용할 수 있다. 그림 4는 Eshed Robotec 회사의 5축 로봇인 Scorbot ER VII에 PULNIX TM-545 CCD 카메라를 장착한 것이다.

이미지 디지털라이저 (image digitizer)는 Imaging Technology Inc.의 ITEX PCplus를 사용하였는데 이것은 512 x 480의 화소 배열을 갖는다. 그림 5의 a와 b는 각각 back image와 front image이다. 그림 6-a의 동그라미 안의 '+' 기호는 back image에서 3차원 좌표값을 알고자 하는 물체 특징점의 image point를 표시한 것이고, 그림 6-b의 동그라미 안은 그림 6-a에서 선택된 image point의 대응되는 화소를 찾기위해 front image에서 radial 방향으로 탐색해갈 때 normalized correlation의 변화를 막대그래프의 형태로 front image상에 표시한 것이다. 그림에서 normalized correlation의 변화가 unimodal인 것을 관찰할 수 있는데 이것은 본 연구에서 탐색의 효율을 높이기 위해 사용한 line search method의 타당성을 제공한다. 그림 6-c는 normalized correlation을 이용하여 대응되는 화소를 찾아낸 결과를 front image에 '+'로 표시한 것이다.

실험결과는 물체 특징점의 실제거리가 110cm정도이고, 카메라 이동 거리 Δz 가 20cm일 때 오차가 7%이내였고, Δz 가 30cm일 때에 오차가 5%이내였다.

4. 결론 및 추후 연구방향

본 연구에서는 어떤 물체가 두 대의 카메라에 동시에 보이지 않는 기존의 lateral model의 문제점을 해결하기 위하여, 로봇트의 팔에 장착된 한대의 카메라에서 얻어진 서로 다른 두개의 이미지를 이용하여 물체 특징점의 3차원 위치를

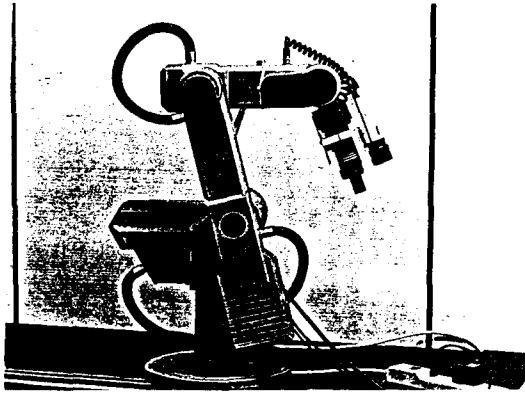


그림 4. Robot-vision system

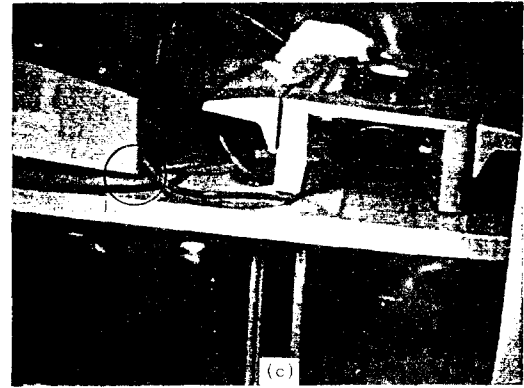
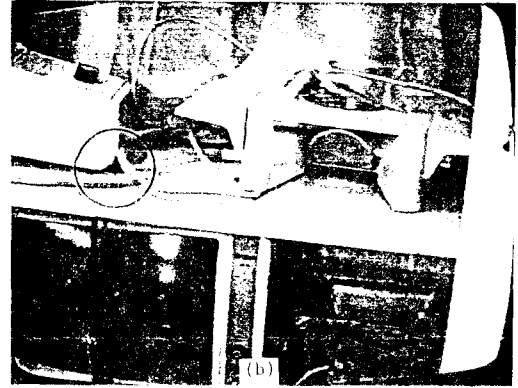
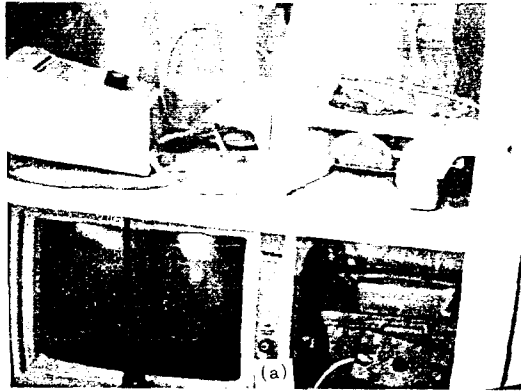
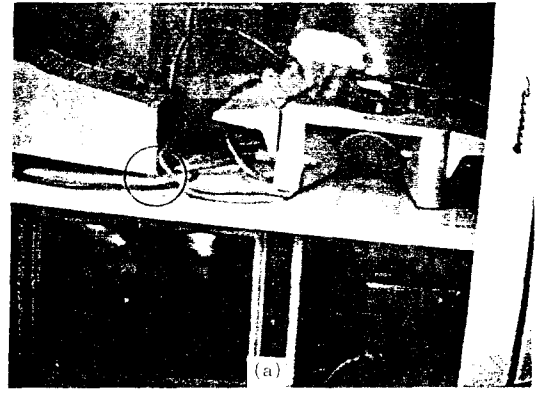


그림 5. (a) back image (b) front image

그림 6. (a) back image에의 선택된 large point, (b) normalized correlation의 binogram, (c) corresponding point

계산하는 robot-vision system을 구축하였다. 물체의 3차원 좌표를 얻기 위한 두개의 이미지중에서 한 이미지 (back image)는 물체에서 떨어진 상태에서 또다른 이미지 (front image)는 카메라의 optical axis를 따라 물체에 좀더 가까이 접근하여 얻었다. 두 개의 이미지를 이용하여 3차원 정보를

추출하기 위해서는 먼저 서로 대응되는 회소를 찾는 correspondence problem을 해결해야 하는데 본 연구에서는 normalized correlation을 이용하여 이 문제를 해결하였다.

본 연구의 다음 단계로는 axial motion model의 중요한

특징중에 하나인, 보는 위치를 어느 정도 범위내에서 자유롭게 선정할 수 있다는 점을 이용하여 복잡한 작업 환경, 예를 들어 여러개의 구조물 가운데에 놓여 있는 밸브를 조이거나 푸는 작업 혹은 조립작업이 비교적 복잡하여 여러 곳에서의 관찰이 필요한 작업 등에서는 여러 대의 로봇을 사용하여 눈으로서의 역할을 하는 로봇과 손으로서의 역할을 하는 로봇이 상호협력하여 작업을 진행하는 연구를 하고자 한다.

참 고 문 헌

- [1] N. Alvertos, D. Brazakovic, and R. C. Gonzalez, "Camera geometries for image matching in 3-D machine vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 9, pp. 897-917, 1989.
- [2] W. E. L. Grimson, "Aspects of a computational theory of human stereo vision," in Proc. Image Understanding Workshop, College Park, MD, pp. 128-149, 1980.
- [3] H. Itoh, A. Miyauchi, and S. Ozawa, "Distance measuring method using only simple vision constructed for moving robots," in IEEE Proc. 7th Int. Conf. Pattern Recognition, vol. 1, Montreal, P. Q., Canada, pp. 192-195, 1984.
- [4] R. Jain, S. L. Bartlett, and N. O'Brien, "Motion stereo using egomotion complex logarithmic mapping," IEEE Trans. PAMI, vol. PAMI-9, no. 3, pp. 356-369, 1987.
- [5] D. Marr, Vision, Freeman, San Fransico, 1982.
- [6] T. D. Williams, "depth from camera motion in a real world scene," IEEE Trans. PAMI, vol. PAMI-2, no. 6, pp. 511-516, 1980.