

산업용 로보트를 이용한 비쥬얼 서보

°원 영범, 박찬웅, 이쾌희
서강대학교 전자공학과

Image-based visual servo with Industrial robot

°Young Bum Pyon, Chan Eung Park and Kuae Hi Lee
Dept. of Electronics Eng., Sogang Univ.

Abstract

Visual feedback has traditionally been used in robot manipulator control to a limited extent. However in varying environment, visual data is needed to control the manipulator to complete the desired task. In this paper we present a method of manipulator control scheme called image-based visual servo. In this scheme we use image data as feedback to control robot manipulator. And we also show the experimental results with an industrial robot manipulator.

1. 서론

로보트 시스템에 요구되는 프로그램은 그 작업 환경에 따라 순차적 방식과 자동 방식 나눌 수 있다.[1] 순차적 로보트 시스템은 로보트가 정확한 위치로 움직여서 대상체를 다루어야 하므로 작업환경이 완벽하게 구성되어 있어야 한다. 만약 대상체의 위치가 조금만 변한다면 로보트의 작업은 실패하게 된다. 반면에 자동 로보트 시스템은 대상체의 위치와 자세를 찾아내어 작업을 수행하도록 프로그램 되므로 대상체의 위치의 변화가 문제를 일으키지 않게 된다.

카메라를 이용하는 자동 로보트 시스템에서는 대상체의 위치와 자세를 알아내어야 하는데, 정보 추출 방법에 따라 스테레오 정합 기법과 calibration 기법으로 나눌 수 있다. 스테레오 정합 기법에 관한 연구는 많이 진행되어 왔는데, 가장 널리 이용되고 있는 방법은 두 대의 카메라를 이용하는 스테레오 정합 기법(binocular stereo)과 한 대의 움직이는 카메라를 이용하는 스테레오 정합 기법(active monocular stereo)으로 나눌 수 있다. 두 대의 카메라를 이용하는 스테레오 기법에서는 시차를 가지고 있는 좌, 우 영상에서의 동일 점을 찾아내는 정합에 의해 깊이 정보를 알아내고, 한 대의 움직이는 카메라를 사용하는 스테레오 기법에서는 카메라를 움직여 시차가 있는 영상을 취득, 역시 동일점을 찾아 내어 깊이 정보를 알아내게 된다. 이러한 스테레오 정합 기법에서는 두 영상사이의 동일점을 찾는 정합 과정에서 많은 시간이

요구되므로 일반적인 로보트 매니퓰레이터의 제어에는 부적합하다. 이러한 이유에서 본 논문에서는 영상으로부터 대상체의 위치와 자세정보를 얻기 위해 대상체의 크기와 형태 정보를 이용하는 calibration 기법[2,3]을 사용한다.

시각 센서를 이용하여 로보트를 제어하는 visual servo 기법은 그 방식에 따라 위치 기반 visual servo(position based visual servo)와 영상 기반 visual servo(image based visual servo)로 나눌 수 있다. 위치 기반 visual servo(그림 1)에서는 로보트의 위치 오차의 채환을 이용하여 로보트를 제어하게 되는데 비해, 영상 기반 visual servo(그림 2)에서는 영상 평면 위에서의 대상체의 위치오차를 채환 하여 로보트를 제어하게 된다.[4]

일반적으로 카메라를 이용하여 영상을 취득하는 경우, 렌즈의 왜곡으로 인한 측정 잡음과 컴퓨터의 기억 장치를 이용하는 데서 발생하는 양자화 오차 등으로 인해 영상위의 대상체의 위치는 오차를 포함하게 된다. 따라서 이 오차를 포함하고 있는 영상으로부터 삼차원 정보를 추출하여 그대로 이용하는 위치기반 visual servo는 이 오차로 인한 영향을 받게 된다. 반면, 영상 기반 visual servo의 경우는 영상위에서의 대상체의 위치를 채환의 대상으로 하고 이를 최소화하는 방향으로 로보트를 움직이게 되므로 이 오차의 영향을 받지 않게 된다.[1]

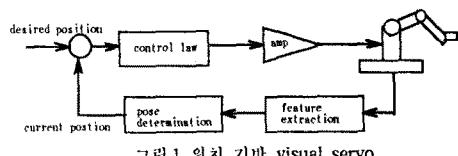


그림.1 위치 기반 visual servo

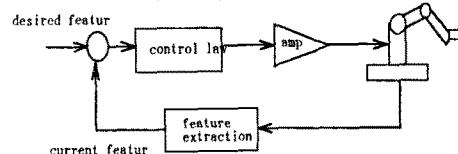


그림.2 영상 기반 visual servo

2. 대상체의 위치와 자세 추출

대상체로부터 3차원 정보를 얻어내는 방법으로 본 논문에서는 Tsai, Lenz가 제안한 calibration 기법을 이용하였다. 우선 영상 평면과 대상체와의 관계와의 관계는 투시 모델을 적용하여 식(1)과 같이 나타낼 수 있다.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix} \quad (1)$$

여기서 (x, y) 는 영상 평면위의 좌표, f 는 렌즈의 촛점 거리이고, X, Y, Z 는 카메라 좌표계에서의 대상체의 3차원 위치를 나타낸다. 대상체 좌표계에서의 임의의 한 점은 대상체 좌표계에서 ${}^o p_i = [{}^o X_i \ {}^o Y_i \ {}^o Z_i]^T$ 로 표현된다. 이 값을 안다고 한다면, 이 점은 카메라 좌표계에서는 식(2)와 같이 표현 할 수 있다.

$${}^c P_i = {}^c R_i {}^o P_i + {}^c P_o \quad (2)$$

이를 매트릭스 형태로 나타내면 식(3)과 같이 된다.

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} {}^o X_i \\ {}^o Y_i \\ {}^o Z_i \end{bmatrix} + \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3)$$

식(3)을 식(1)에 대입하여 풀면 식(4)와 같은 관계를 얻을 수 있다.

$$f(R_{11} {}^o X_i + R_{12} {}^o Y_i + R_{13} {}^o Z_i + X) = x_i (R_{11} {}^o X_i + R_{12} {}^o Y_i + R_{13} {}^o Z_i + Z) \quad (4)$$

$$f(R_{21} {}^o X_i + R_{22} {}^o Y_i + R_{23} {}^o Z_i + Y) = y_i (R_{21} {}^o X_i + R_{22} {}^o Y_i + R_{23} {}^o Z_i + Z) \quad (4)$$

이를 매트릭스 형식으로 나타내면 식(5)와 같이 된다.

$$M_i r = s_i \quad (5)$$

여기서, 각각은 M_i, s_i, r 각각은 아래와 같다.

$$M_i = \begin{bmatrix} fX_i & fY_i & 0 & 0 & -x_i X & -x_i Y & f & 0 \\ 0 & 0 & fX_i & fY_i & -x_i X & -y_i Y & 0 & f \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$r = \begin{bmatrix} R_{11} & R_{12} & R_{13} & R_{21} & R_{22} & R_{23} & X & Y \\ Z & Z & Z & Z & Z & Z & Z & Z \end{bmatrix}$$

$$s_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

대상체의 네 점에 대해 각각 식(5)와 같은 방정식을 세우고, 식(6)과 같이 나타내면, 최소좌승법과 ${}^c R_o$ 의 orthonormality를 이용하면, r 의 모든 성분을 구할 수 있게 된다. 이렇게 해서 영상으로 부터 대상체의 위치와 자세 정보를 얻게 된다.

$$Mr = s \quad (6)$$

여기서, $M = [M_1 \cdots M_N]^T$, $s = [s_1 \cdots s_N]$ 를 나타낸다.

3. 영상 자코비안

카메라의 움직임은, 영상 자코비안(image jacobian)이라 불리는 자코비안 매트릭스에 의해, 영상위의 대상체의 움직임으로 표현되어 질 수 있다.[4]

카메라 좌표계의 Z축이 광축과 같은 방향으로 고정되어 있다고 하면, 강체의 움직임은 선형 속도 성분과 좌표계의 원점을 지나는 축에 대한 각 속도 성분으로 나타낼 수 있다.[2][5] 카메라의 움직임으로 인해 발생한 대상체의 움직임 속도는 식(7)과 같이 카메라의 선형 속도 T 와 각 속도 R 로 나타낼 수 있다.

$$\frac{dP}{dt} = -T - R \times P \quad (7)$$

여기서 식(1)의 양변을 시간에 대해 미분한 식을 대입하여 정리하면 식(8)을 얻을 수 있다.

$$\begin{bmatrix} x_i \\ y_i \\ z_i \\ v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} J_{image}(x_1, y_1, Z_1) \\ \vdots \\ J_{image}(x_N, y_N, Z_N) \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$$

$$J_{image}(x, y, Z) = \begin{bmatrix} -\frac{f}{Z} & 0 & -\frac{x}{Z} & \frac{xy}{f} & -\frac{(x^2 + f^2)}{f^2} & y \\ 0 & -\frac{f}{Z} & -\frac{y}{Z} & -\frac{(y^2 + f^2)}{f^2} & -\frac{xy}{f} & -x \end{bmatrix} \quad (8)$$

v_x, v_y, v_z 는 각각 카메라 좌표계에서 본 카메라의 선 속도와 각 속도를 나타내고, x_i, y_i 는 영상 평면위에서의 특징점의 위치를 나타낸다. 이 매트릭스를 영상 자코비안이라 하는데 이를 통해 영상 평면위의 대상체의 움직임과 카메라의 움직임 사이의 관계를 알 수 있게 된다.

4. 세이 알고리듬

본 영상기반 비주얼 서보에서는 시각 채환시스템의 상태벡터를 $x = [x_1 \ y_1 \ \dots \ x_N \ y_N]^T$ 로, 세이 입력을 $u = [{}^c v_c \ {}^c \omega_c]^T$ 로 정의 하면 상태 공간 상에서 식(9)와 같이 표현할 수 있다.

$$\dot{x} = J^{+ image} u \quad (9)$$

여기서 $J^{+ image}$ 는, N개의 특징점에 대한 확장 영상 자코비안으로 식(10)과 같다.

$$J^{+ image} = \begin{bmatrix} J_{image}(x_1, y_1, Z_1) \\ \vdots \\ J_{image}(x_N, y_N, Z_N) \end{bmatrix} \quad (10)$$

위의 식은 이산 시스템에서는 식(11)과 같다.

$$x(k+1) = x(k) + J^{+ image} \Delta T u(k) \quad (11)$$

여기서 ΔT 는 샘플링 시간이고, $B(k) = J^{+ image} \Delta T$ 라고 하면 이 식은 (12)와 같이 나타낼 수 있다.

$$x(k+1) = x(k) + B(k)u(k) \quad (12)$$

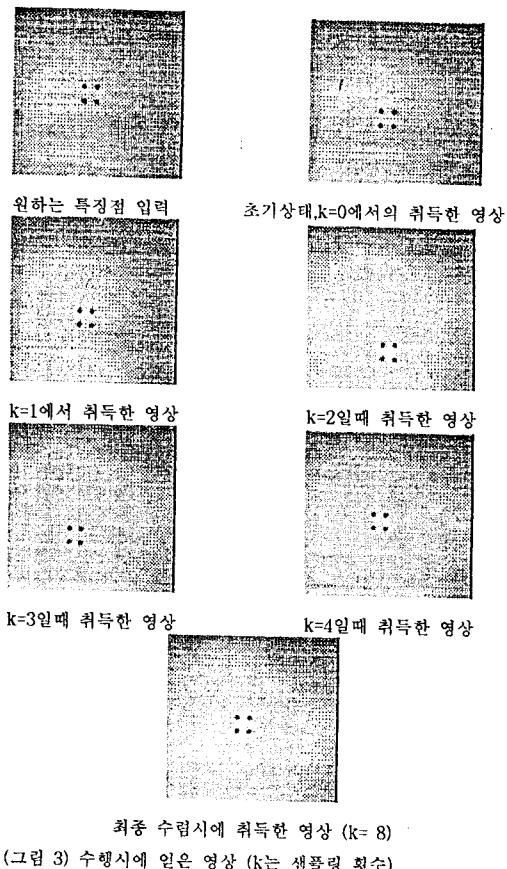
이러한 시스템에 대해 원하는 입력 $x_D(k+1)$ 과 $x(k+1)$ 사이의 오차를 최소화하는 최적제어 알고리듬을 사용한다. 최적 제어의 비용 함수는 식(13)과 같고, 이를 최소화시키는 최적의 입력은 식(14)와 같이 된다[6][7].

$$J(k+1) = [x(k) + B(k)u(k) - x_D(k+1)]^T [x(k) + B(k)u(k) - x_D(k+1)] \quad (13)$$

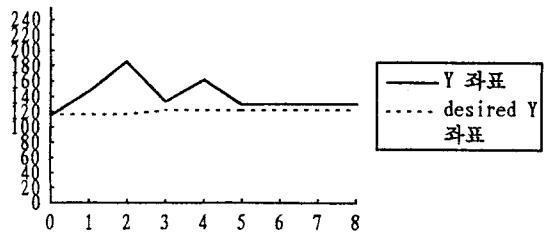
$$u(k) = -(B(k)^T B(k))^{-1} B^T (x(k) - x_D(k+1)) \quad (14)$$

5. 실험 및 실험 결과

실험은 산업용 로보트 매니퓰레이터인 Fanuc Arc Mate를 이용하여 수행하고, 영상취득을 위하여 CCD camera를 사용하였다. 작업은 영상 평면위의 특징점의 위치를 원하는 입력으로 현재 취득한 영상에서의 특징점의 위치를 채환하여 영상 평면 위에서 원하는 입력값에 이를 때까지 로보트를 움직이는 것으로 정의했다. (그림 3)에서는 로보트가 작업 수행 시에 각 샘플 마다 취득한 영상을 보이고 있고, 각 특징점의 위치 변화 추이를 (그림 4,5,6)에 보았다. 계단형 입력에 대해 네 특징점의 위치 오차는 입력 특징점과 약 8 화소 정도가 되어 수렴함을 알 수 있었다.



(그림 3) 수행시에 얻은 영상 (k 는 샘플링 횟수)

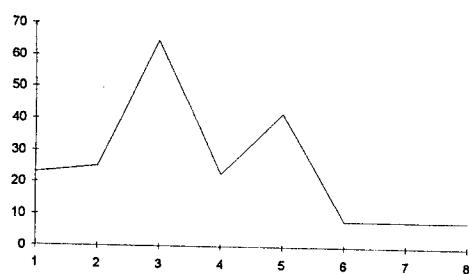


(그림 6) 영상 평면 위에서의 원하는 특정점의 y 좌표와 현재의 특정점의 y 좌표

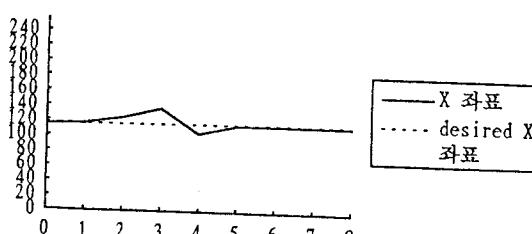
6. 결론

본 논문에서는 변화하는 작업환경 하에서 로보트 매니퓰레이터를 제어할 수 있는 프로그래밍 기법 중의 하나인 시각센서를 이용하여 대상체의 자세 및 위치 정보를 취득하여 로보트 매니퓰레이터를 직접 제어 하는 영상 기반 visual servo를 구현하고 그 실험 결과를 보았다. 이 visual servoing 시스템에서는 로보트가 움직이는 도중 영상 평면상에서 대상체를 놓치는 경우 작업을 더 이상 수행할 수 없게 되므로 이를 위한 알고리듬의 개발이 필요하다. 또한 빠른 영상 처리를 위한 병렬 구조의 하드웨어와 알고리듬의 개발이 필요하다. 본 연구에서는 매우 단순한 대상체를 실험 대상으로 삼았으나, 일반적으로 실제의 작업환경은 여러 복잡한 요소들을 포함하고 있을 것이므로, 3차원 물체 인식등의 알고리듬 개발이 필요하다.

7. 참고문헌



(그림 4) 각 각의 샘플링 시의 영상 상의 특징점의 오차 (비 특징 점간의 유클리디안 오차의 평균)



(그림 5) 영상 평면위에서의 원하는 특정 점의 x좌표 와 수행과정에서 얻은 특징점의 x 좌표

- [1] K. Hashimoto, T. Kimoto, T. Ebine and H. Kimura, "Manipulator Control with Image-Based Visual Servo," *IEEE Int. Conf. Robotics and Automation*, Sacramento, Calif., pp.2267-2272, 1991.
- [2] B. K. P. Horn, *Robot Vision*, Cambridge, Massachusetts, MIT Press, 1986.
- [3] K. S. Fu, R. C. Gonzalez and C. S. G. Lee: *Robotics: Control, Sensing, Vision, and Intelligence*, New York, McGraw-Hill, 1987.
- [4] Peter I. Corke, "Visual Control of Robot Manipulators - a Review," *Visual servoing*, pp.1-pp.31, 1993.
- [5] P. Anandan, J.R.Bergen, K.J. Hanna, "Hierarchical Model-Based Motion Estimation," *MotionAnalysis and Image Sequence Processing*, pp.1-22, 1993.
- [6] B. Nelson, N.P.Papanikolopoulos, and P.K.Khosla, "Visual Servoing for Robotic Assembly," *Visual servoing*, 1993.
- [7] Brian D.O.Anderson , John B. Moore, *Optimal Control: Linear Quadratic Methods*, Prentice-Hall, 1987.