

카메라 교정 오차에 강인한 3차원 직선 경로 추종을 위한 전환 비주얼 서보잉 기법

A Switched Visual Servoing Technique Robust to Camera Calibration Errors for Reaching the Desired Location Following a Straight Line in 3-D Space

김도형¹·정명진²

Kim, Do Hyoung¹·Chung, Myung Jin²

Abstract The problem of establishing the servo system to reach the desired location keeping all features in the field of view and following a straight line is considered. In addition, robustness of camera calibration parameters is considered in this paper. The proposed approach is based on switching from position-based visual servoing (PBVS) to image-based visual servoing (IBVS) and allows the camera path to follow a straight line. To achieve the objective, a pose estimation method is required; the camera's target pose is estimated from the obtained images without the knowledge of the object. A switched control law moves the camera equipped to a robot end-effector near the desired location following a straight line in Cartesian space and then positions it to the desired pose with robustness to camera calibration error. Finally simulation results show the feasibility of the proposed visual servoing technique.

Keywords: Visual Servoing, Pose Estimation, View Synthesis

1. 서론

로봇이나 주변 환경의 불완전한 정보를 가지고도 로봇 움직임의 정확성을 얻기 위하여 시각센서의 피드백을 쓰는 것은 효율적이다. 더욱이, 시각 센서의 정보는 다른 하드웨어와는 다르게 많은 정보를 함축하고 있어서 소프트웨어 알고리즘을 통해 시각 정보를 어떻게 처리하느냐에 따라 그 기능을 변화시킬 수 있다. 이러한 시각 정보의 유용성 때문에 근세기에 시각 정보 처리에 대한 많은 연구가 진행되어 왔고, 아직도 미결된 많은 문제점들이 끊임없이 제시되고 해결되고 있다[1].

시각 기반의 로봇 제어는 크게 세 그룹으로 나눌 수 있다. 영상 기반 (image-based) 제어 시스템, 위치 기반 (position-based) 제어 시스템, 하이브리드 (hybrid) 제어 시스템으로 구분된다[2,3]. 위치 기반 비주얼 서보잉 (position-based visual servoing; PBVS)에서는 3차원 공간에서 입출력 간의 오차를 정의함으로써 이를 최소화하도록 동작된다. 목표 위치는 로봇 시스템에 장착된 시

각 센서로부터 받아들인 물체의 영상 정보와 물체의 3차원 CAD 모델을 바탕으로 3차원 위치를 예측하게 된다. 이때, 카메라의 교정(calibration) 정보 역시 사용되는데, 예측된 3차원 목표 위치는 카메라의 교정 변수에 크게 의존하고 있어 카메라의 교정과정에 오차를 많이 포함하게 되면 목표 위치 역시 큰 오차를 가지게 되어 정확성을 떨어뜨리게 된다. 앞서 언급한 바와 같이 물체의 3차원 모델 정보까지 요구되므로 동적으로 변화하는 환경과 다양한 물체에는 부적합하게 되어 적응성 (flexibility)이나 자발성(autonomous ability) 측면에서 그 활용도가 낮다.

영상 기반 비주얼 서보잉(image-based visual servoing; IBVS)에서는 3차원 목표 위치 대신에 목표 영상이 주어져 현재 위치에서의 영상과 목표 영상에서 발생하는 오차를 최소화하도록 동작된다. 영상 기반의 특징점들의 위치오차로부터 시스템을 제어하기 때문에 물체에 대한 3차원 CAD 모델이 요구되지 않는다는 장점이 있고, 카메라의 교정(calibration) 오차에 강인하게 반응하게 된다. 하지만, 영상의 오차를 최소화 하는 방향으로 시스템이 제어되기 때문에 3차원 공간에서의 로봇 움직

¹ 한국과학기술원 전자전산학과 박사과정

² 한국과학기술원 전자전산학과 교수

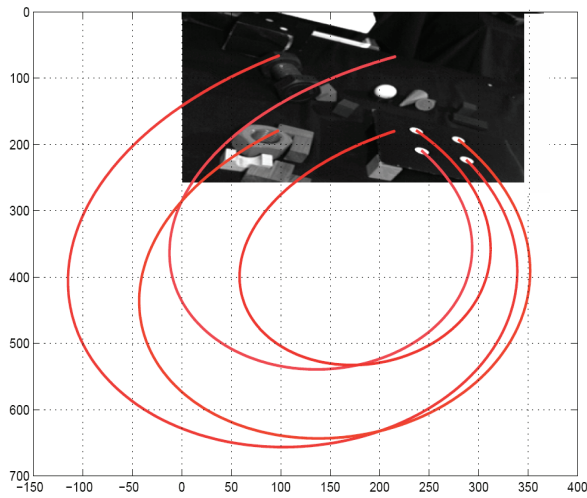


그림 1. 카메라의 시계 제한에 의해 발생하는 문제의 한 모의 실험 예제 [7]. 영상 특징점들의 경로가 영상의 경계를 크게 벗어난다.

임은 고려할 수 없다는 단점이 있다. 예를 들어, 목표 영상과 현재 영상과의 오차로부터 선형적인 경로가 생기도록 제어를 한다고 하더라도 3차원 공간상에서는 로봇이 선형적인 움직임을 갖지 않고 불필요한 움직임을 갖게 될 수 있다.

이러한 기존의 비주얼 서보잉의 문제점을 해결하고자 여러 가지 하이브리드 비주얼 서보잉(hybrid visual servoing; HVS) 방법들이 1990년대 후반부터 제안되었다. 하이브리드 비주얼 서보잉은 분할된 IBVS (partitioned IBVS)로 불리기도 한다. 그 이유는 기존의 하이브리드 기법은 IBVS에 근간을 두고 현재 영상과 목표 영상, 두 장의 영상을 가지고 부분적이거나 3차원 공간의 정보를 추출하여 이를 제어루틴에 삽입하여 3차원 상의 제어가 가능하도록 하는 것이 일반적인 방법이다. 추출된 3차원 정보를 이용하여 어떻게 제어루틴에 삽입하느냐에 따라서 여러 하이브리드 방법이 존재하게 된다[4,5,6].

현존하는 비주얼 서보잉 알고리즘들을 분석해 보면 IBVS이 카메라 교정 에러에 가장 강인한 특성을 보여주며 PBVS 기법이 카메라 교정 파라미터에 의존한 물체의 위치를 추정하기 때문에 가장 민감한 특성을 가지고 있다. HVS 역시 일부분이나마 영상 데이터로부터 3차원 정보를 추출하여 사용하기 때문에 IBVS에 비해서는 카메라 교정 에러에 민감한 반응을 보인다. 영상의 특징점들이 카메라의 시계를 벗어나서 발생하는 문제점을 가장 잘 다룰 수 있는 기법은 역시 IBVS이다. IBVS 기법은 앞서 설명한 바와 같이 현재 영상과 목표 영상의 특징점 위치 차이로부터 발생한 오차로부터 제어 입력을

생성하기 때문에 영상의 특징점들이 카메라의 시계를 벗어나는 것을 어느 정도 방지할 수 있다. 하지만 IBVS 일 경우라도 영상 자코비안(image Jacobian)의 추정 오차가 크게 되면 영상 공간에서 조차 직선 경로를 형성하지 않기 때문에 카메라의 시계를 벗어나는 경우가 발생할 수도 있다. PBVS와 HVS는 영상 오차만으로 제어 입력을 생성하지 않기 때문에 특히 카메라 시계 제한 문제점이 빈번하게 나타난다. IBVS과 같이 영상의 오차만으로 제어 입력을 생성할 경우에는 영상의 변환은 직선 경로를 생성하며 수렴해 나가는 경우라도 실제 3차원 상에서의 로봇의 움직임을 살펴보면 요구되지 않은 움직임이나 불필요한 움직임을 보이게 된다. 이러한 3차원 상의 로봇 움직임 제어를 위해서는 PBVS 기법이 가장 효과적이다.

앞에서 살펴본 바와 같이 기존의 비주얼 서보잉 방법에는 각기 장단점이 존재하게 된다. 본 논문에서는 기존의 비주얼 서보잉 방법의 장점만을 취하기 위해서 PBVS과 IBVS를 순차적으로 전환하여 카메라 교정 오차에 강인하며 영상의 특징점들을 카메라 시계 안에 유지시키며 3차원 상에서의 직선경로를 추종하는 방법을 제안하였다. 또한, 제안된 기법이 카메라 교정 오차에 강인하게 제어가 가능할 뿐만 아니라 영상 특징점들이 영상의 경계를 벗어나지 않으면서 3차원 상에서 목표 위치까지 직선 경로의 움직임으로 수렴하는 과정을 모의실험을 통해 보인다.

2. 기존 비주얼 서보잉의 문제점

비주얼 서보잉의 방법들은 앞에 구분했던 바와 같이 제어에 사용되는 입출력의 형태에 따라 나뉘어 진다. 다양한 입출력에 따라 로봇의 움직임 형태는 다르게 나타날 지라도, 일반적으로 시각 정보를 제어루틴 상에 피드백 되어 사용하기 때문에 최종 위치에 도달하는 수렴성과 영상 측정 오차에 강인한 특성을 보이게 된다. 하지만, 수렴성과 안정성 문제가 발생하는 경우도 있다.

2.1 카메라의 시계 제한

비주얼 서보잉 기법에서 전형적인 문제점 중의 하나는 카메라의 시계 제한에 의해 발생하는 문제점이다. 아이-인-핸드(eye-in-hand) 구성을 가지는 비주얼 서보잉의 가장 기본적인 관측 데이터는 로봇 머니플레이터에 장착된 카메라를 통해서 얻어지는 물체 영상의 특징점들이다. 이러한 영상 특징점들을 목표 위치까지 카메라를 조작하는 도중에 카메라의 시계 제한에 의해서 영상의 경계 밖으로 나가 관측이 불가능해지는 경우가 발생한다.

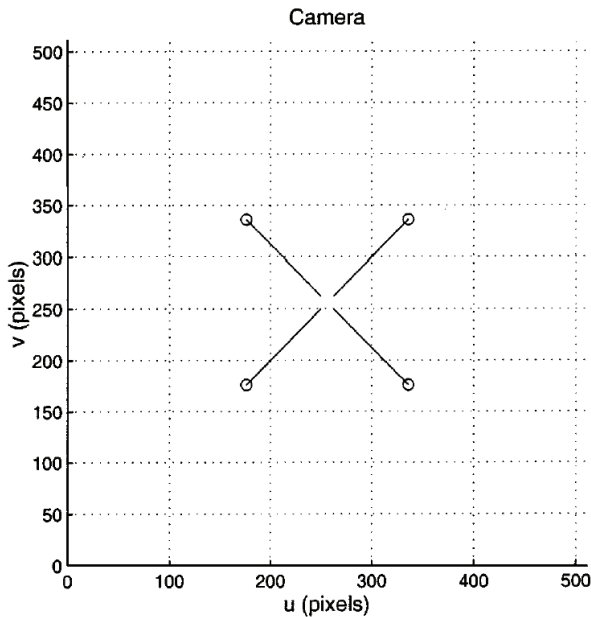


그림 2. 기존의 IBVS 수행시 발생하는 작업 특이점 문제로 카메라의 광축을 회전축으로 180° 회전하는 것이 목표 위치로 설정된 경우의 모의 실험 예제[10].

다. 이는 관측 정보를 추출할 수 없게 되어 결국 비주얼 서보잉의 실패를 야기하게 된다. 특히, PBVS 기법에서는 영상 공간상 제어가 없기 때문에 3차원 상에서 서보잉하는 도중에 물체의 영상 특징점들이 카메라의 시계를 벗어날 수 있게 된다(그림 1 참조). IBVS에서도 카메라의 교정(calibration)변수들이 부정확하면 현재 영상에서 목표 영상으로 직선 경로가 아닌 곡선 경로로 수렴하게 되는데 이 경우에 카메라의 시계를 벗어나는 경우가 발생할 수 있다.

이러한 문제점을 해결하기 위해 기존의 제안된 해결 방법으로는 영상의 경계 부근에 척력 전위장(repulsive potential field)을 형성하고 목표 위치에는 인력 전위장(attractive potential field)을 형성하여 제어하는 방법이 있다[7]. 척력 전위장에 의해 로봇은 영상의 경계에서 안쪽으로 움직이는 힘을 받게 되어 목표 위치의 영상 특징점들이 카메라의 시계 안에 유지되도록 한다. 하지만, 척력 전위장이 영상의 경계에 접근하는 움직임을 성공적으로 회피할지라도 카메라의 광축 방향으로 비효율적인 움직임을 야기시킬 수 있다. Corke와 Hutchinson도 이 문제를 영상의 경계를 장애물로 고려하여 회피하도록 하는 문제로 고려하여 전위장을 이용한 기법으로 해결하였다[6].

최근에는 이러한 문제점을 해결하기 위해 IBVS와

PBVS 기법을 영상 경계를 기준으로 전환하여 사용하는 방법이 제안되었다[8,9]. 이는 PBVS를 수행하는 도중에 영상 특징점들이 영상 경계 부근에 근접하면 IBVS로 전환하여 영상 특징점들이 다시 안쪽으로 움직이도록 제어 입력을 만들고 일정한 영역 안으로 영상 특징점들이 모두 들어오면 다시 PBVS를 수행하도록 전환하는 반복적 수행 방법이다[8]. 하지만 이 방법은 제어 신호에 있어서 심한 떨림 현상(chattering)이 존재하게 되어 제어 측면에서 비효율적이게 된다. 또 다른 전환 방법으로는 전위장 기반의 움직임 경로를 2차원 영상이나 3차원 상에서 발생할 수 있는 여러 가지 제약에 따라 IBVS와 PBVS로 전환하면서 생성한 후에 그 경로를 추종하도록 제어하는 방법으로 문제를 해결하고 있다[9].

2.2 작업 특이점 (task singularity) 문제

IBVS의 전형적으로 언급되는 문제점은 영상 기반으로 서보잉을 하기 때문에 3차원에서 발생하는 로봇에 장착된 카메라의 움직임이 불필요하게 매우 크다는 것이다. 서보잉 도중에 목표 위치의 수직 방향으로 멀어졌다가 다시 다가오는 것과 같은 요구되지 않은 움직임이 종종 발생한다. 이는 카메라의 영상 정보는 3차원 공간상의 정보가 2차원에 투영되어 있기 때문에 카메라 영상에서의 같은 위치의 점이라 할지라도 3차원 상에서는 매우 멀리 떨어져 있는 두 점일 수 있다. 이러한 불필요한 3차원 상의 매우 큰 움직임은 서보잉의 성능을 저하시키고 실제 로봇이 움직일 수 없는 경우도 발생한다.

이러한 작업 특이점 문제의 가장 대표적인 예는 움직임이 될 목표 위치가 카메라의 광축을 기준으로 180° 회전하도록 설정되어 있는 작업이다. 이러한 작업이 주어졌을 경우에 IBVS에서는 영상 특징점들의 현재 위치에서 목표 위치로 직선으로 이동하도록 제어를 하기 때문에 그림 2에서 보여주는 것과 같이 영상 특징점들이 가운데 한 점으로 모이는 직선 운동을 하게 되고 이는 카메라의 위치를 물체의 수직 방향으로 멀어지게 하는 움직임을 야기시킨다[10].

작업 특이점 문제를 해결하기 위해 제안되어 왔던 방법이 하이브리드 방법들로 영상 특징점들의 위치에 대한 오차와 영상으로부터 추출한 부분적인 3차원 정보를 이용한 오차를 동시에 고려하여 제어 입력을 생성하게 함으로써 문제를 해결하고 있다[4,5,6]. 하지만 이러한 하이브리드 방법들 역시 영상을 통해 3차원 정보를 추출하기 때문에 전형적인 IBVS에 비해 카메라 교정 오차에 민감하게 반응하는 문제점이 있다.

3. 제안된 전환 비주얼 서보잉

앞서 언급한 바와 같이 아이-인-핸드 비주얼 서보잉의 목적은 영상을 기반으로 도달하고자 하는 목표 위치까지 로봇에 부착된 카메라를 움직이는 것이다. 이는 영상의 특징점의 현재 상태와 요구된 상태 사이의 오차를 최소화하는 것에 의해 수행된다. 이번 장에서는 본 논문에서 제안된 비주얼 서보잉 방법에 대해 언급할 것이다. 제안된 방법은 목표 위치 추정과 전환 비주얼 서보잉 기법으로 구성되어 있다. 목표 위치 추정을 위해 물체를 주시하면서 능동적으로 카메라가 움직이며 가상의 영상을 합성한다. 그런 후에 가상의 카메라를 두 장 이상의 영상을 이용하여 주어진 목표 위치로 위치시킨다. 우리는 임의의 관측 위치에서의 보여 지는 새로운 영상을 두 장 이상의 기준 영상으로부터 등극선 기하학(Epipolar geometry)을 통해 합성한다. 생성된 가상의 영상은 실제 카메라에서 관측된 영상이 아닌 가상의 관측 위치에 대응되는 합성된 영상이다. 이렇게 합성된 영상을 가지고 IBVS 기법을 통해 목표 영상과 일치시켜 최종적인 3차원 목표 위치를 계산해 낸다.

전환 비주얼 서보잉 기법에서는 앞 단계에서 예측한 3차원 목표 위치에 실제 카메라를 PBVS 기법으로 우선 위치시키고, 좀 더 정밀하게 IBVS 기법으로 실제 목표 위치에 정확하게 움직이도록 한다. 이러한 전환 방법은 기존의 IBVS 기법과 PBVS 기법의 장점만을 가지게 된다. 목표 위치 추정 방법은 가상의 영상을 이용하여 3차원 목표 위치를 추정함으로써 기존의 다른 서보잉 기법에서 사용된 위치 추정 방법과는 다르게 물체의 3차원 정보가 필요하지 않고 목표 위치에서 얻은 목표 영상과 각각 다른 위치에서 얻은 두 장 이상의 영상으로 최종 카메라의 목표 위치를 추정할 수 있다. 전환 비주얼 서보잉 기법은 3차원 공간상에서 PBVS로 우선 예측된 목표 위치까지 도달하게 되므로 3차원 상에서 직선 경로로 이동할 수 있어 불필요한 로봇의 움직임을 없앨 수 있고, 그런 후에 IBVS로 좀 더 정밀하게 실제 목표 위치에 도달하도록 서보잉 한다. 이때, 물체를 카메라의 시계에서 놓치지 않기 위해서 기존의 PBVS 방법을 사용하지 않고 본 논문에서 제안한 변경된 PBVS 기법을 적용하여 이를 수행하게 된다. 결론적으로 제안된 전환 비주얼 서보잉은 PBVS의 장점인 3차원 공간상에서의 로봇 움직임을 제어할 수 있다는 점과 IBVS의 장점인 카메라의 교정 변수에 강인한 수렴성과 정확성을 보장할 수 있다는 점을 동시에 가질 수 있게 된다.

3.1 카메라 모델

본 논문에서는 카메라 모델이 3차원 상의 한 점 x 를 영상의 한 점 x 로 할당되는 $sx = PX$ 를 만족하는 일반적인 핀-홀(pin-hole) 카메라와 이에 따른 투영행렬 (Projection matrix) P 를 고려하고 있다. K 를 카메라 교정의 내부 파라미터를 나타내는 행렬이라고 하고 카메라 렌즈의 왜곡이 없다고 가정하면 정규화된 영상 좌표의 특징점인 $x_n = [xy1]^T$ 과 픽셀 영상 좌표의 특징점인 $x = [uv1]^T$ 사이의 카메라 투영 변환 행렬은 $x = Kx_n$ 과 같은 관계로 표현되고, 형태는 식(1)과 같다.

$$K = \begin{bmatrix} \alpha_x & \alpha_{xy} & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

여기서 α_x 와 α_y 는 각각 x , y 좌표 방향으로의 비례 정도를 나타내고, $(x_0, y_0)^T$ 는 기준 점(principal point)을 나타내며, α_{xy} 는 x 축과 y 축의 비틀어진 정도(skew)를 나타내는 변수이다. 일반적인 카메라에서 비틀어진 정도는 무시할 정도로 매우 작다.

따라서, 카메라 행렬 P 는 다음 식(2)와 같이 표현된다.

$$P = KT = \begin{bmatrix} \alpha_x & \alpha_{xy} & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \quad (2)$$

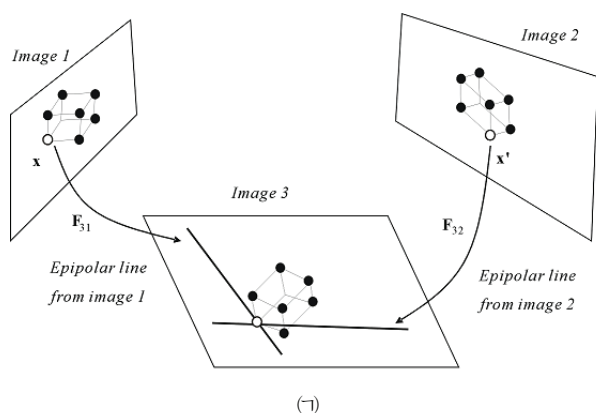
여기서 행렬 T 는 카메라 좌표계와 기준 좌표계 사이의 변환 행렬(transformation matrix)을 나타낸다.

3.2 위치 예측

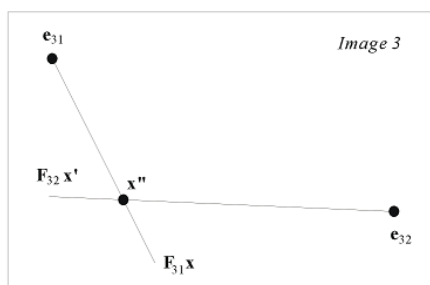
3.2.1 여러 영상으로부터의 새로운 영상 합성

여기서는 두 장 이상의 주어진 실제 영상으로부터 임의의 카메라 위치에서 보여 지는 가상의 새로운 영상을 합성하는 방법에 대해 언급하고자 한다. 생성된 영상은 실제 카메라에서 촬영된 영상이 아닌 가상의 관측 위치에 대응한다. 따라서 이렇게 새로이 생성된 영상을 가상 영상이라 하고 그 때의 3차원에 존재하는 카메라를 가상 카메라라고 하겠다.

새로운 영상 합성 방법은 여러 가지 접근이 있는데 크게 3차원 기반의 합성 방법과 2차원 기반의 합성 방법이 있는데 본 연구에서는 3차원 정보를 사용하지 않고 얻는 방법이므로 2차원 기반의 합성 방법 중의 등극선 기하학 (Epipolar geometry)을 사용한 영상 합성 방법을 사용하였다[11,12,13,14]. 두 장의 실제 영상에서 충분한 대응점들의 한 묶음이 주어지면 추가적인 환경에 대한



(a)



(b)

그림 3. 등극선 기하학 (Epipolar geometry)을 이용한 영상 특징점 생성 [11]. (a) 3차원 상의 점 x 의 투영된 영상은 첫 번째 영상과 두 번째 영상에서 같은 3차원상의 점이 투영된 것이므로 투영된 점들이, $x \leftrightarrow x'$, 대응 관계를 가지게 된다. 세 번째 영상에서의 x 의 투영 정보는 처음 두 장의 영상으로부터 얻어진 에피폴라(epipolar) 선들의, $F_{31}x$ 과 $F_{32}x'$, 교점으로 계산할 수 있다. (b) 세 번째 영상에서 관측된 에피폴라(epipole)와 계산된 영상 특징점인 x'' .

정보 없이 새로운 영상을 생성해 낼 수 있다. 이러한 합성 문제를 영상들 간의 기하학적인 특성을 나타내는 펀더멘털 행렬(fundamental matrix)만으로 해결할 수 있다.

주어진 두 장의 실제 영상과 우리가 합성하고자 하는 세 번째 영상 사이의 펀더멘털(fundamental)행렬인 F_{21} 와 F_{31} , F_{32} 를 알고 있다고 가정하고 두 장의 실제 영상에서 대응되는 점인 x 와 x' 가 있다고 하자. 우리가 해결하고자 하는 것은 세 번째 영상에서의 대응되는 점인 x'' 를 구하고자 하는 것이다. x'' 은 첫 번째 영상에서의 점 x 와 대응되기 때문에 등극선 기하학 (Epipolar geometry)의 특성에 의해 에피폴라(epipolar) 라인위의 한 점이라는 특징을 가지게 되고, F_{31} 을 이용하여 에피폴라(epipolar) 라인을 $F_{31}x$ 로 계산할 수 있다. 같은 방법으로, 두 번째 영상에서의 대응되는 점 x' 와 F_{32} 을 이용하

여 또 다른 에피폴라(epipolar) 라인을 $F_{32}x'$ 과 같이 구할 수 있다. 결국 구하고자 하는 세 번째 영상에서의 대응점인 x'' 은 두 라인위에 모두 존재하는 한 점이기에 때문에 두 에피폴라(epipolar) 라인의 교점이 바로 x'' 의 위치가 된다. 이는 다음 식(3)과 같이 간단히 계산할 수 있다.

$$x'' = (F_{31}x) \times (F_{32}x') \quad (3)$$

그림 3(a)에서는 등극선 기하학 (Epipolar geometry)을 이용하여 두 장의 영상으로부터 세 번째 영상의 특징점인 x'' 를 합성하는 방법에 대한 개념도를 보여주고 있다. 하지만 이러한 방법을 이용하여 새로운 영상을 합성할 때 조심해야할 점은 한 가지 특이한 불완전한 구조를 회피해야 한다는 것이다. 이 불완전함은 그림 3(b)에서 유추할 수가 있는데, 세 번째 영상에서의 대응점은 두 개의 에피폴라(epipolar) 라인의 교점으로 구해지게 되는데 이 때 사용되는 두 개의 에피폴라(epipolar) 라인이 일치하게 되거나 거의 일치할 정도로 기울기가 유사한 경우에 대응점을 정확하게 찾기 어렵다는 문제점이 있다. 결국 이러한 구조에서는 해결하고자 하는 문제가 악성 조건(ill-condition)이 된다는 것을 의미한다. 그러므로 이러한 불완전한 구조를 회피하기 위해서는 x'' 와 두 개의 에피폴라(epipole)인 e_{31} 와 e_{32} 가 세 번째 영상에서 서로 한 직선위에 있지 않아야 되고 이는 기하학적으로 세 개의 카메라의 중심인 C 와 C' , C'' 로 이루어지는 평면 위에 3차원 점인 x 가 존재하지 않아야 한다는 것을 의미한다. 세 개의 카메라의 중심으로 이루어진 한 평면을 일반적으로 트리포컬(trifocal) 평면이라고 부른다. 그리고 3차원 점인 x 가 트리포컬 평면위에 있지 않아도 그 평면에 가까우면 가까울수록 세 번째 영상에서의 대응점을 구하는 것이 부정확해진다[11,13].

3.2.2 가상 영상의 피드백 (Virtual IBVS)

여기서는 앞서 합성한 가상 영상을 이용하여 3차원 상의 카메라 목표 위치를 추정하는 방법에 대해 알아본다. 본 연구에서는 물체의 3차원 모델 정보 없이 위치를 추정하는데 그 목표가 있다. 본 논문에서 제안하는 가상 영상 피드백은 두 장의 실제 영상으로부터 합성된 가상 영상을 기존의 IBVS의 입력으로 주어 실제 목표 위치에서 촬영된 목표 영상과의 오차를 줄임으로써 3차원 목표 위치를 계산하는 것이다. IBVS의 특성에 의해 최종적으로 수렴된 가상 카메라의 위치는 추정하고자 하는 실제 카메라의 목표 위치와 일치하게 될 것이고, 추정 방법은 합성된 영상의 오차가 없다면 교정 파라미터들에 강인하게 동작할 것이다.

본 논문에서는 이러한 방법을 가상의 영상이 IBVS의 입력으로 주어지므로 가상 IBVS(VIBVS)로 명명하였다. 가상 IBVS은 영상 특징점의 좌표가 입력 변수로 사용되고 합성된 가상 영상의 특징점과 목표 영상의 특징점 사이의 오차 변화와 가상 카메라의 3차원 위치 변화를 규명하는 영상 자코비안을 사용한다. 한 개의 영상 특징점의 변화에 대해 각각 x , y 위치 값에 따라 2개의 자코비안 식이 형성되고, m 개의 영상 특징점들이 있을 경우에 각 영상 특징점들에 의해 생성된 식들을 쌓음으로써 영상 제어를 위한 다음과 같은 식을 형성할 수 있다.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y}_1 \\ \vdots \\ \dot{x}_m \\ \dot{y}_m \end{bmatrix} = \begin{bmatrix} -1/Z_1 & 0 & x_1/Z_1 & x_1y_1 & -(1+x_1^2) & y_1 \\ 0 & -1/Z_1 & y_1/Z_1 & 1+y_1^2 & -x_1y_1 & -x_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1/Z_m & 0 & x_m/Z_m & x_my_m & -(1+x_m^2) & y_m \\ 0 & -1/Z_m & y_m/Z_m & 1+y_m^2 & -x_my_m & -x_m \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (4)$$

카메라 위치의 갱신은 비례 제어기(proportional controller)를 사용하여 식(5)에서 보여 주는 것처럼 기하급수적으로 감소하는 제어를 수행할 수 있다.

$$T = -\lambda J^+(s - s^*) \quad (5)$$

이때 W 는 상수로 이루어진 가중치 행렬이다. 갱신된 가상 카메라의 위치는 펀더멘털(fundamental) 행렬인 F 를 아래 식(6)을 통해서 변화된 위치에 맞추어 재계산할 수 있다.

$$\begin{aligned} E &= [t]_x \cdot R \\ F &= K^{-1} \cdot E \cdot K \end{aligned} \quad (6)$$

3.3 제안된 전환 비주얼 서보잉 기법

본 장에서는 실제 카메라를 제어하는 배치를 수행하기 위해 고안된 제어 벡터를 선택하는 것에 대해 거론할 것이다.

3.3.1 3차원 공간에서의 제어

3차원 공간에서 직접적으로 제어 벡터를 만드는 것은 현재 상태와 목표 상태의 위치 차이를 $[000]^T$ 로 되도록 하고 회전량 차이 역시 $[000]^T$ 이 되도록 하면 된다. 본 논문에서 사용된 3차원 공간에서의 제어는 목표 물체의 3차원 정보를 사용하지 않는다는 점을 제외하고는 기존의 PBVS와 유사하다. PBVS에서 사용되는 카메라의 목표 위치는 앞서 제안한 카메라 위치 추정 방법에 의해 미리 예측되어 있다. 따라서 현재와 목표 위치 사이의

오차가 3차원 상에서 정의가 되고 이 오차를 다음과 같은 식(7)과 같은 방법으로 카메라의 속도 성분을 선택해 주면, 목표 위치까지 도달할 수 있다.

$$\begin{cases} \omega = \lambda_r r \\ v = \lambda_t t \end{cases} \quad (7)$$

여기서 사용된 λ_r 과 $\lambda_t \in \mathbb{R}$ 은 양 수인 제어 이득 값이다.

본 논문에서는 기존의 PBVS의 제어 벡터를 그대로 사용하지 않고 앞서 언급한 기존 비주얼 서보잉 기법의 문제점을 해결하기 위해 각도의 속도 성분 갱신을 식(8)과 같이 수정한다.

$$\omega = \lambda_r r' = \lambda_r (\alpha r + (1-\alpha)r^{obj}) \quad (8)$$

여기서 r' 은 $r^{current} - r^{goal}$ 으로 구해지는 r 과 물체의 방향을 나타내는 r^{obj} 의 가중된 합으로 구해진다. 위치의 속도 성분 갱신은 현재 위치에서 목표 위치를 뺀, $t^{current} - t^{goal}$ 로 이루어진 위치 성분의 오차 t 에 의해 결정된다. 이렇게 각도의 속도 성분 갱신을 변경함으로써 기존의 PBVS에서 발생하는 카메라의 시계 제한에서 발생하는 문제점이 해결되며, 서보잉하는 동안에 계속해서 물체의 영상 특징점들을 카메라의 시계안에 보존하고자 하는 제어 성분이 발생하게 된다.

3.3.2 2차원 공간에서의 제어

2차원 공간에서의 카메라 제어는 3차원 공간에서의 PBVS 기법이 적용된 후에 카메라의 위치가 실제 카메라 목표 위치에 가까이 위치한 후에 수행된다. 2차원 공간에서의 제어는 좀 더 정밀한 카메라를 위치시키기 위한 것으로서 기존의 IBVS 기법의 장점인 카메라 교정 파라미터의 오차에 강인하게 동작하기 때문에 좀 더 정밀한 서보잉 결과를 얻을 수 있다.

영상 자코비안은 영상 특징점의 변화와 카메라 위치의 변화의 관계를 나타내고 있으며 식(9)와 같이 정의된다.

$$L(s, Z) = \begin{bmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \\ 0 & -1/Zy & Z/1+y^2 & -xy & -x \end{bmatrix} \quad (9)$$

현재 영상 특징점의 위치, s 와 목표 영상 특징점의 위치, s^* 와의 오차를 e 로 관계식을 $e = C(s - s^*)$ 로 표현하면, 제어는 e 를 0으로 만드는 제어 문제로 변경되게 된다. 이때 가중치 행렬 C 는 $s = s^*$ 와 $Z = Z^*$ 를 목표 영상에서의 값으로 놓은 영상 자코비안의 의사 역변환인 $L(s^*, Z^*)^+$ 로 간주할 수 있다. 이때 Z^* 는 목표 카메라의 위치에서의 대략적인 물체까지의 거리이다. IBVS가 목

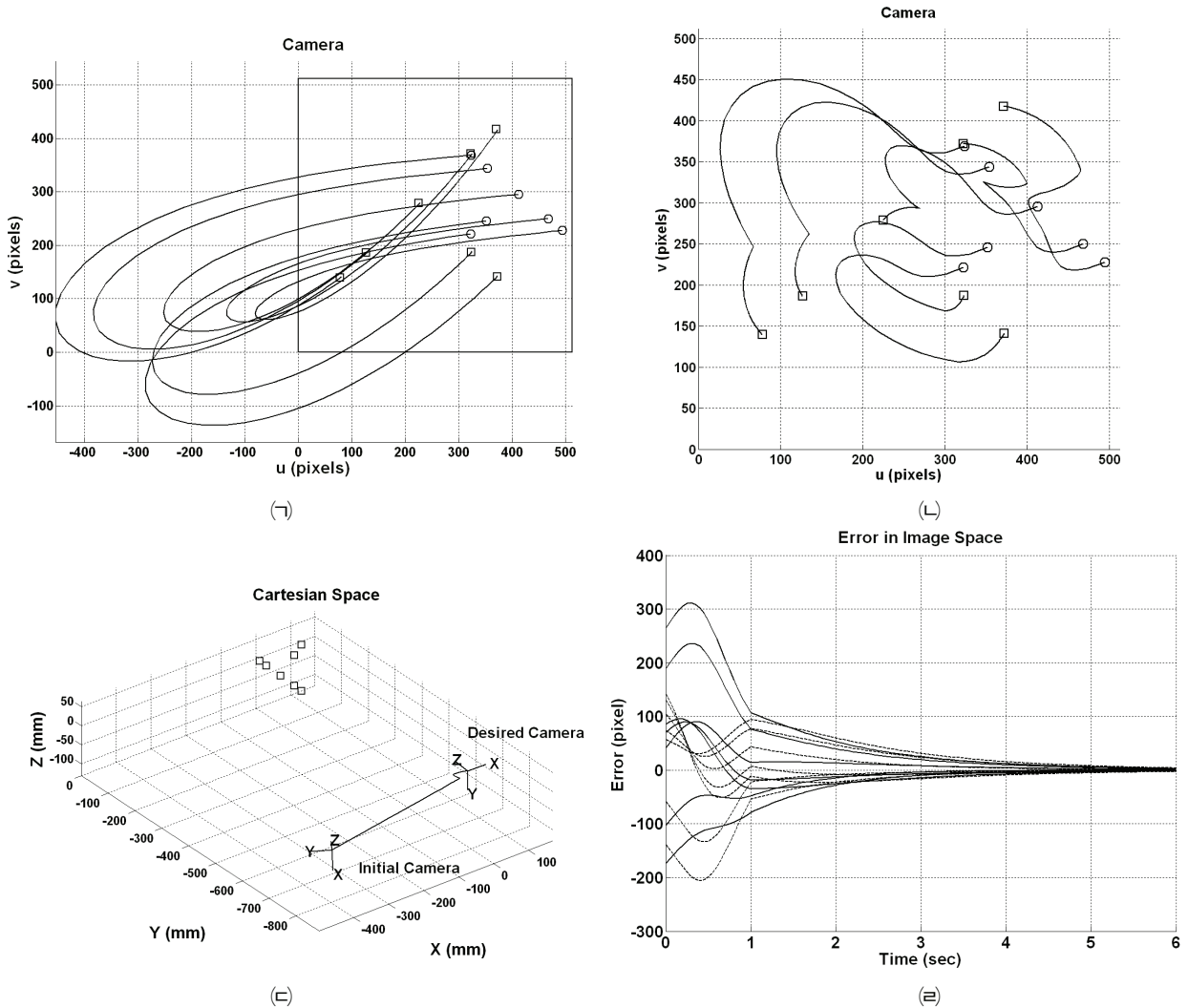


그림 4. 제안된 비주얼 서보 시스템의 카메라의 시계 제한에 의해 발생하는 문제를 해결하는 모의실험 결과. (a) (b) 영상 특징 점들의 경로로 원은 영상 특징점들의 처음 위치이고 사각형은 최종 위치를 나타낸다. (c) 3차원 상의 카메라 위치와 경로를 나타내고 사각형으로 표시된 것은 물체의 위치를 표시한다. (d) 영상 공간에서의 오차 경로를 표시한 것으로 점선은 u-방향의 영상 좌표 오차를 실선은 v-방향의 영상 좌표 오차를 나타낸다.

표 위치의 근처에서 수행되기 때문에 수렴성에 대한 조건을 만족하게 된다.

$$e_{3D} = \begin{bmatrix} \Delta t \\ \Delta r \end{bmatrix} \quad (12)$$

3.3.3 제어 벡터

일반적인 카메라 배치 문제를 앞서 언급한 바와 같이 목표치와 현재치의 오차를 0으로 만드는 문제로 정리할 수 있다. 따라서 수행해야 할 작업을 다음과 같이 정의할 수 있다.

$$e = C \begin{bmatrix} e_{2D} \\ e_{3D} \end{bmatrix} \quad (10)$$

$$e_{2D} = s - s^* \quad (11)$$

이때 C는 시변 가중치 행렬로써 PBVS에서 IBVS로 변환해주는 역할을 한다. s와 s*의 차이와 Δt , Δr 을 0으로 기하급수적으로 수렴하도록 하기 위해 식 $\dot{e} = -\lambda e$ 에 대입하고, 이때 λ 는 수렴 정도를 결정하는 변수이다. 오차의 시간에 대한 미분은 카메라의 속도 성분, T에 의해 식(13)과 같이 결정된다.

$$\dot{e} = LT \quad (13)$$

이때 L은 영상 자코비안으로 식(14)와 같이 얻어진다.

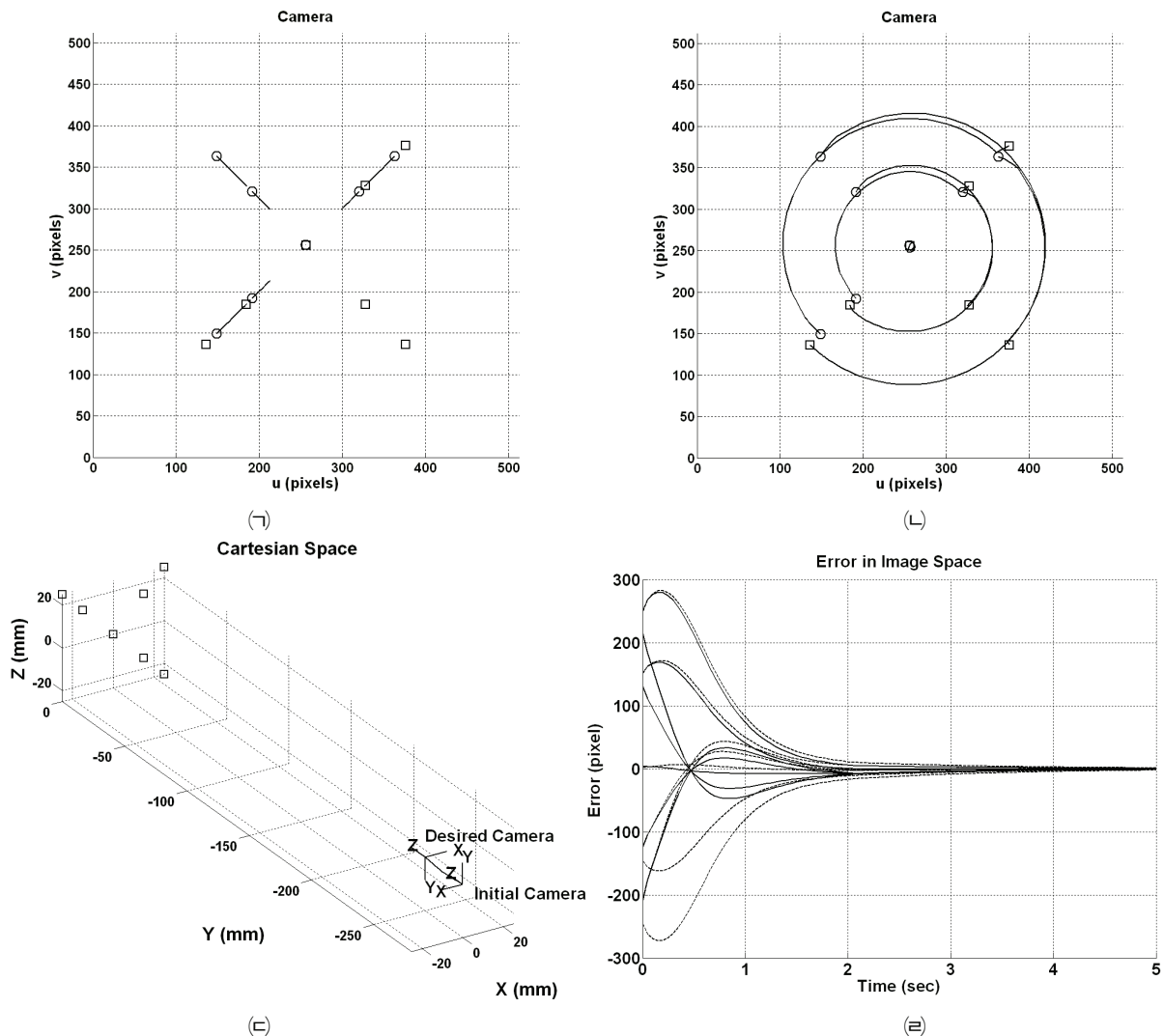


그림 5. 제안된 비주얼 서보 시스템의 작업 특이점에 의해 발생하는 문제를 해결하는 모의실험 결과. (㉠) 기존의 IBVS를 수행 하였을 경우의 결과로 영상 특징점들이 가운데로 모이는 방향으로 제어 입력이 발생된다. (㉡) 영상 특징점들의 경로로 원은 영상 특징점들의 처음 위치이고 사각형은 최종 위치를 나타낸다. (㉢) 3차원 상의 카메라 위치와 경로를 나타내고 사각형으로 표시된 것은 물체의 위치를 표시한다. (㉣) 영상 공간에서의 오차 경로를 표시한 것으로 점선은 u-방향의 영상 좌표 오차를 실선은 v-방향의 영상 좌표 오차를 나타낸다.

$$L(s, Z^*) = \begin{bmatrix} J_v & J_\omega \\ I_3 & 0_3 \\ 0_3 & I_3 \end{bmatrix} \quad (14)$$

식(14)에서 사용된 J_v 와 J_ω 는 식(15)와 같이 정의된다.

$$J_v = \begin{bmatrix} -1/Z^* & 0 & x/Z^* \\ 0 & -1/Z^* & y/Z^* \end{bmatrix}, \quad (15)$$

$$J_\omega = \begin{bmatrix} xy & -(1+x^2) & y \\ 1+y^2 & -xy & -x \end{bmatrix}$$

그러므로 목표 물체가 움직이지 않는 환경에서 제어 벡터는 식(16)과 같이 정의될 수 있다.

$$T = -\lambda \hat{L}^+ e \quad (16)$$

식(16)에서 구해진 카메라 속도 성분은 로봇 제어기에 입력으로 주어지게 되고 이때 e 는 측정된 오차로 카메라의 영상 정보가 들어올 때마다 매번 계산되며, \hat{L} 는 영상 자코비안 L 의 오차를 포함한 예측 값이다. 최종적으로 제어 벡터는 식(17)과 같이 주어진다.

$$T = -\lambda \begin{bmatrix} \hat{L}_v^+ & I_3 & 0_3 \\ \hat{L}_\omega^+ & 0_3 & I_3 \end{bmatrix} e \quad (17)$$

4. 모의실험

본 장에서는 제안한 비주얼 서보 시스템의 모의 실험 결과로써 기존의 비주얼 서보 시스템이 다루지 못한 작업에 대해 수행하고 그 결과를 얻었다. 목표 물체의 3차원 정보는 사용되지 않았고, 평면 물체로 7개의 특징점으로 구성되어 있다. Matlab을 이용하여 PC에서 테스트되었다. 카메라는 핀-홀 카메라 모델을 사용하였고 로봇의 엔드-이펙터(end-effector)에 장착된 아이-인-핸드(eye-in-hand) 구조로 모의실험 환경을 구성하였다. 영상의 크기는 512×512로 제안하였고 모든 영상 좌표 값들은 정수 값으로 설정하고 계산에 이용되었다. 또한 물체의 최종 목표 위치에서 촬영된 목표 영상은 미리 주어진다고 가정하였다.

4.1. 예제 1: 카메라의 시계 제한 문제

기존 비주얼 서보잉 기법의 전형적인 문제인 영상의 특징점들이 서보잉 도중에 카메라의 시계를 벗어나는 경우를 고려하고자 한다. 모의실험에서 카메라의 목적 위치는 위치 성분은 (117, -633, -43) [mm]로 주어졌고 회전 성분은 (-1.57, 0, 0) [rad]으로 각각 롤(roll), 피치(pitch), 요(yaw) 순서대로 기준 좌표계를 기준으로 주어졌다. 큰 위치 차이를 주기 위해 초기 카메라 위치를 위치는 (-545, -845, 160) [mm], 회전은 (-1.62, -0.12 -0.69) [rad]로 주어졌다. 앞서 언급했던 바와 같이 3차원 제어 벡터 생성시 회전 성분에 해당하는 부분이 영상의 특징점들을 카메라의 시계 안에 유지시키기 위한 제어 성분이 추가되었기 때문에 그림 4의 모의실험 결과에서 살펴볼 수 있듯이 모든 특징점들이 서보잉이 수행되는 동안 영상의 경계를 벗어나지 않고 유지되는 결과를 얻을 수 있다.

4.2. 예제 2: 작업 특이점 문제

기존 IBVS의 전형적인 문제점 중의 하나는 작업 특이점에 의해 발생하는 문제이다. 이는 영상 자코비안의 특이점 문제로 생각할 수 있고, 이러한 현상은 3차원 정보를 사용하지 않고 2차원 정보만을 사용하여 서보잉을 하기 때문에 발생하는 현상이다. 그림 5에서 보여주는 예제는 작업 특이점 문제를 나타내는 가장 대표적인 예로 초기 위치의 카메라와 목표 위치의 카메라가 상대적인 위치가 카메라의 광축으로 180°만큼 회전되어 있는 상황이다. 그림 5(㉠)에서 보여주는 것처럼 기존의 IBVS을 이용한 서보잉을 하게 되면 영상의 특징점들이 가운데로 모이는 카메라의 움직임을 형성하게 되고 이는 광축의 음의 방향으로 진행되는 것을 의미한다. 이러한 문

제는 PBVS기법으로 손쉽게 해결할 수 있는데 제안된 전환 기법은 초기에 PBVS에 의해 카메라의 위치를 제어하게 되므로 그림 5(㉡)과(㉢)에서 보는 바와 같이 카메라의 목표 위치로 수렴하는 것을 보여주게 된다.

5. 결론 및 추후과제

본 논문에서는 새로운 전환 비주얼 서보 시스템을 제안하였다. 현재 상태와 목표 상태 사이의 3차원 정보를 추출하기 위해 가상 영상을 이용한 새로운 위치 예측 방법을 제안하였고, 예측된 위치를 이용하여 PBVS를 통해 대략적인 목표 위치로 이동하는 동안 3차원 상에서의 제어를 통해 불필요한 3차원상의 로봇 움직임을 억제하고 대략적인 목표 위치에 도달하면 IBVS를 통해 좀 더 정밀하고 교정 오차에 강인하게 실제 목표 위치에 수렴하도록 하였다. 이러한 순차적인 전환 비주얼 서보 시스템을 통해 PBVS에서의 장점인 3차원 상에서의 직선 경로에 가까운 움직임 제어를 가능하게 하였고, 이어지는 IBVS를 통해서 IBVS의 장점인 카메라의 부정확한 교정에도 강인하게 수렴하는 것이 가능하게 되었다. 본 연구에서는 위치 예측을 하기 위해 싱글 카메라의 움직임을 통해 두 장 이상의 영상을 얻었지만, 만약 스테레오 카메라를 사용한다면 움직이는 물체에 대해서도 확장하여 이 방법을 적용할 수 있다. 왜냐하면 현재 위치에서 서로 다른 위치에 위치한 두 장의 영상을 동시에 얻을 수 있기 때문이다.

추후과제로는 제안된 비주얼 서보 시스템의 유용성을 실제 실험을 통해 살펴볼 필요가 있다는 점이다. 그리고 제안된 하이브리드 제어 벡터는 여러 개의 직관적으로 결정된 변수들을 포함하고 있어 시스템을 조금 더 일반화하기 위해서는 변수들을 자동적으로 선택할 수 있는 기법이 요구된다.

참고문헌

- [1] J. S. Park and M. J. Chung, "Path Planning with Uncalibrated Stereo Rig for Image-Based Visual Servoing under Large Pose Discrepancy," *IEEE Trans. Robotics and Automation*, vol. 19, no. 2, pp. 250-258, 2003.
- [2] K. Hashimoto, *Visual Servoing: Real Time Control of Robot Manipulators Based on Visual Sensory Feedback*, Singapore: World Scientific Press, 1993.
- [3] S. A. Hutchinson, G. D. Hager and P. I. Corke, "A Tutorial on Visual Servo Control," *IEEE Trans.*

Robotics and Automation, vol. 12, no. 5, pp. 651-670, 1996.

[4] E. Malis, F. Chaumette and S. Boudet, "2-1/2-D Visual Servoing," *IEEE Trans. Robotics and Automation*, vol. 15, no. 2, April 1999.

[5] K. Deguchi, "Optimal Motion Control for Image-Based Visual Servoing by Decoupling Translation and Rotation," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Victoria, B.C., Canada, 1998, pp. 705-711.

[6] P. I. Corke and S. A. Hutchinson, "A new partitioned approach to image based visual servo control," *IEEE Trans. Robotics and Automation*, vol. 17, no. 4, pp. 507-515, 2001.

[7] Y. Mezouar and F. Chaumette, "Path Planning in Image Space for Robust Visual Servoing," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2000, pp. 2759-2764.

[8] G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino, "Keeping Features in the Field of View in Eye-In-Hand Visual Servoing: A Switching Approach," *IEEE Tran. Robotics*, vol. 20, no. 5, pp. 908-913, 2004.

[9] L. Deng, F. Janabi-Sharifi, and W. J. Wilson, "Hybrid Motion Control and Planning Strategies for Visual Servoing," *IEEE Tran. Industrial Electronics*, vol. 52, no. 4, pp. 1024-1040, 2005.

[10] F. Chaumette, "Potential Problems of Stability and Convergence in Image-Based and Position-Based Visual Servoing," in *The Confluence of Vision and Control*, D. Kriegman, G. Hager, and S. Morse, Eds. Berlin, Germany:Springer-Verlag, 1998, vol. 237, pp. 66-78.

[11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge, UK: Cambridge University Press, 2000.

[12] J.-I. Park and S. Inoue, "Arbitrary View Generation from Multiple Cameras," in *Proc. IEEE Int. Conf. Image Processing*, Washington, DC, USA, 1997, pp. 149-152.

[13] M. A. Akhloufi, V. Polotski, and P. Cohen, "Virtual View Synthesis from Uncalibrated Stereo Cameras," in *Proc. IEEE Int. Conf. Multimedia Computing and Systems*, Florence, Italy, 1999, pp. 672-677.

[14] B. S. Boufama, "The Use of Homographies for View Synthesis," in *Proc. of IEEE Int. Conf. Pattern Recognition*, Barcelona, Spain, 2000, pp. 563-566.

[15] N. R. Gans and S. A. Hutchinson, "An Experimental Study of Hybrid Switched System Approaches to Visual Servoing," in *Proc. of IEEE Int. Conf. Robotics and Automation*, 2003, pp. 3061-3068.



김도형

1998 전북대학교 제어계측 공학과 (학사)

2000 한국과학기술원 전자전산학과 (석사)

2000~2001 Ecole Polytechnique Federal de Lausanne 교환학생

2000~현재 한국과학기술원 전자전산학과 박사과정
 관심분야: Visual Servoing, Computer Vision, Human Computer Interaction



정명진

1973 서울대학교 전기공학과 (학사)

1977 미시간대학교 전기공학과 (석사)

1983 미시간대학교 제어공학과 (박사)

1976 국방과학연구소 연구원
 1981~1983 미시간대학교 CRIM 연구 조교
 1983~현재 한국과학기술원 전자전산학과 교수
 관심분야: Service Robot, Human Robot Interaction, Intelligent Robot