

심층 신경망 기반 추적기를 사용한 사용자 추종 로봇

손찬영, 이혜민, 이준구, 오지용

한국전자통신연구원 대경권연구센터

{cysohn, leehaemin90, leeje01679, jiyongoh}@etri.re.kr

A User Following Robot Using Deep Neural Network Based Tracker

Chanyoung Sohn, Hea-Min Lee, Joon-Goo Lee, and Jiyong Oh

Daegu-Gyeongbuk Research Center

Electronics and Telecommunications Research Institute

요약

본 논문은 이동 로봇이 사용자를 따라다니며 다양한 협업이 가능한 사용자 추종 기능을 다룬다. 본 연구를 통해 개발된 사용자 추종 기술은 RGBD 카메라 신호로부터 사용자의 위치를 실시간으로 계산하는 추적 모듈과 로봇이 사용자를 적절하게 추종하기 위해 사용자의 위치로부터 주행 명령을 계산하는 주행 모듈로 구성된다. 정적 환경에서 반복 수행된 사용자 추종 실험은 본 논문의 사용자 추종 기술이 효과적이라는 사실을 보여준다.

I. 서론

2016년 3월 알파고의 등장 이후 인공지능 기술은 4차 산업혁명의 핵심 기술로 자리 잡아 현재 다양한 분야에서 활용되고 있으며 로봇 분야에서도 인공지능 기술이 적극적으로 활용되고 있다. 로봇은 전통적으로 인간의 노동력을 대체하거나 노동환경을 개선하기 위해 개발되었다. 특히 농업, 수산업 등 비롯한 1차 산업에서부터 물류창고 및 배송, 제조공장, 병원 등 다양한 산업에서 요구되는 중량물 운반은 로봇이 담당하기에 적합한 작업 중 하나이다. 그러나 농작물 수확, 택배 배송 등은 중량물을 단순히 목적지까지 이송하는 것 이외에 적재물을 운반하는 과정에서 사용자의 작업이 빈번하게 요구된다. 사용자 추종[1]은 이러한 요구에 맞추어 개발되고 있는 대표적인 인간-로봇 협업 기술이다. 본 연구에서는 인공지능 기반의 추적 기술을 활용하여 사용자 추종 로봇을 개발한다. 특히 개발된 사용자 추종 로봇은 VOT-RGBD2019[2] Challenge에서 우승한 SiamDW-D 구조[3]를 활용하여 사용자가 카메라의 FOV 밖으로 사라지거나 다른 사람에 의해 가려진 뒤에 다시 나타나더라도 추적 중이던 사용자를 강인하게 추종할 수 있다.

II. 본론

1. 로봇 및 프로그램의 구조

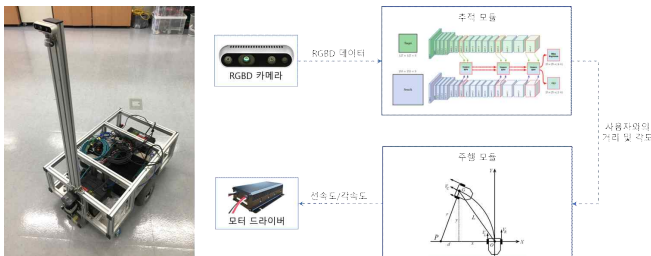


그림 1. 로봇 플랫폼 (좌), 사용자 추종 프로그램 흐름도 (우)

그림 1은 본 연구에서 사용된 2개의 구동 휠과 4개의 캐스터를 갖는 차동

휠 (differential wheel) 타입의 로봇 플랫폼과 사용자 추종을 위한 프로그램의 흐름을 보여주고 있다. ROS 기반의 사용자 추종을 위한 프로그램은 추적 모듈과 주행 모듈로 구성된다. 추적 모듈은 지면으로부터 120cm 높이에 장착된 Intel Realsense D435 카메라로부터 RGBD 데이터를 입력받아 20 fps 이상의 속도로 사용자의 상대 위치를 주행 모듈로 전달한다. 주행 모듈은 현재 사용자의 상대 위치로부터 사용자를 추종하기 위한 로봇의 선속도와 각속도를 계산하여 모터 드라이버로 전달한다.

2. 추적 모듈

2.1. 심층 신경망 기반 추적 모듈

추적기는 target(추적 대상)과 search(검색될 전제) 이미지를 각각의 신경망에 통과시키는 Siamese Tracker 구조를 갖는다. 특히 추적 속도가 빠른 RPN (region proposal network) 구조[4]를 지니며 추적 정확도 및 강인성을 향상시키기 위해 설계된 CResNet[3]을 backbone으로 사용한다. 신경망의 출력은 추적 대상의 영상 내 2차원 중심점과 경계 상자의 너비와 높이, 신뢰도 값이다.

2.2. 추적 대상의 3차원 좌표 추정

로봇이 추종 대상을 따라다니도록 만들기 위해서는 실제 세계의 3차원 좌표가 필요하다. 이때 좌표계는 카메라 좌표계를 사용하며 로봇에 부착된 RGBD 센서의 중심을 원점으로 하는 좌표계이다. 카메라 좌표계에서의 추적 대상 좌표를 구하기 위해 먼저 추적 대상의 깊이 정보를 먼저 획득해야 한다. 일반적으로 깊이 정보는 depth 이미지의 화상 강도로 표현되며, 객체 영역의 평균 화소 값을 이용한다. 하지만 객체 영역은 수시로 변하는 배경을 포함하며, depth 센서의 특징으로 인한 결측이 발생한다. 이를 해결하기 위해 다음과 같이 경계 상자의 너비와 높이의 일정 비율만큼 포함하는 중심 상자의 좌표를 구한다.

$$l = m - w \times r/2, \quad r = m + w \times r/2 + 1,$$

$$t = n - h \times r/2, \quad b = n + h \times r/2 + 1$$

여기서 l, r, t, b 는 각각 중심 상자의 좌, 우, 상, 하의 좌표이며, (m, n) 은

중심 상자의 중심점 좌표이다. 또한 r 은 경계 상자와 중심 상자의 비율을 의미하는데 본 연구에서는 0.1로 설정하였다(그림 2 참조). 이 중심 상자 내에서도 깊이 정보 결측치가 발생하게 되므로 해당 값들을 최대한 제거해야 한다. 중심 상자를 기준으로 평균과 표준 편차값을 구하고, 중심 상자 내 임의의 값의 편차가 표준 편차값에 비례한 기준값보다 작은 값들에 대해서만 평균을 구한다. 이 값을 추적대상의 깊이 값으로 사용한다.

다음으로 영상 좌표계의 경계 상자의 중심 좌표 (m, n) 과 추적 대상의 깊이 값 d 를 이용하여 추적대상의 카메라 좌표계에서의 좌표 (x, y, z) 를 구한다. 이를 위해 센서 카메라의 내부 파라미터 행렬을 사용한다.

$$x = d \times (m - p_x) / f_x, \quad y = d \times (n - p_y) / f_y, \\ z = \sqrt{x^2 + y^2 + d^2},$$

여기서 f 는 초점 거리이며, p_x 는 주점의 x 좌표, p_y 는 주점의 y 좌표이다.

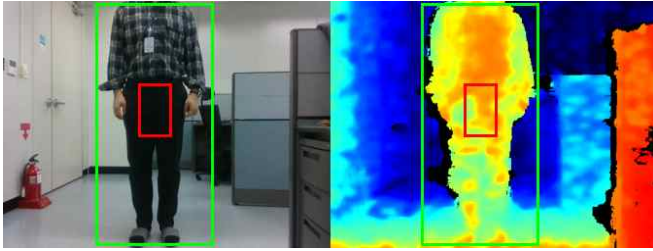


그림 2. 경계 상자와 중심 상자를 표시한 RGB 이미지(좌)와 색상화된 깊이 정보(우)

2.3. 제어기 입력 좌표계 변환

로봇의 모터를 구동하기 위한 제어기는 로봇과 추종 대상과의 상대 거리를 사용한다. 이때 θ 는 카메라 방향에 수직인 영상 중심점을 지나는 평면으로부터 카메라 방향과 추종 대상 간의 각도를 의미한다.

$$\phi = \sin^{-1}\left(\frac{x}{z}\right), \quad x_r = z \cos(\phi), \quad y_r = z \sin(-\phi)$$

앞서 구한 x_r, y_r, z, ϕ 를 제어기에 전달하게 되는데, 낮은 신뢰도를 갖는 추적 결과에 대한 로봇의 오작동을 방지하기 위해 신경망이 추론한 신뢰도 값이 일정 값 이하일 때 제어기에 해당 값을 전달하지 않는다.

3. 주행 모듈

구현한 모바일 로봇은 두 가지 센서 정보를 이용해 주행한다. 첫 번째, 영상처리 모듈에서 로봇이 따라가야 할 추종 대상의 상대 위치와 각도 차이, 거리 차이, 신뢰도 값이 계산되어 입력된다. 두 번째, 엔코더에서 글로벌 좌표 기준 실시간 좌표와 각도 정보가 입력된다. 이 두 데이터를 이용한 경로 계획을 통해 로봇이 추종 대상을 따라 주행하게 만든다. 그림 3의 왼쪽과 같은 이륜 구동 로봇을 제어하는 시스템을 구현하기 위해 다음과 같은 모델을 이용한다.

$$\rho = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2}, \quad \theta = \phi - \tan^{-1}\left(\frac{y_r - y_i}{x_r - x_i}\right) \\ v = K_1 \rho \cos \theta, \quad w = -K_1 \sin \theta \cos \theta - K_2 \theta,$$

여기에서 (x_i, y_i) 와 (x_r, y_r) 은 각각 로봇의 현재 위치와 추종 대상의 위치 정보이고, ρ 는 추종 대상과의 거리, θ 는 추종 대상과의 각도 차이이다. 또한 K_1 과 K_2 는 주행 제어를 위한 제어기의 이득에 해당한다. 주행이 시작되면 로봇은 이미지 센서로부터 도출된 정보로 추종 대상의 위치 정보를 만든다. 추적 모듈로부터 목표 지점의 좌표를 전달받으면 그림 3의 왼쪽과 같은 모바일 로봇의 기구학 모델을 기반으로 로봇의 경로를 계획한다. 위 식을 바탕으로 로봇에 선속도 값(v)과 각속도 값(w)을 제어 입력으

로 주어 로봇을 이동시킨다.

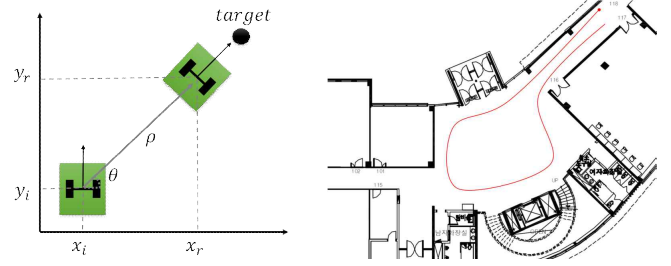


그림 3. 이륜 구동 모바일 로봇(좌), 사용자 추종 실험 환경 및 시험 경로(우)

4. 실험

개발된 사용자 추종 로봇의 성능을 검증하기 위해 그림 3의 오른쪽과 같이 비교적 넓고 지나가는 사람들이 많지 않은 실내 공간에서 미리 정의된 경로를 따라 움직이며 추종의 성공률을 계산하였다. 사용자 추종은 로봇이 출발지부터 도착지까지 지정된 사용자를 추종하며 이동하는 동안 관찰자의 개입이 없는 경우 성공이라고 간주하였다. 추종 실험에는 다른 사람이 사용자와 로봇 사이를 지나가는 의도적인 방해 활동을 포함하였다. 그럼에도 불구하고 개발된 사용자 추종 로봇은 20번의 반복 실험에서 20번 모두 사용자 추종에 성공하였다. 이러한 사실은 본 연구를 통해 개발된 사용자 추종 기능이 효과적이며 강인하다는 것을 의미한다.

III. 결론

본 연구에서는 로봇이 사용자를 따라다니며 사용자와 협업할 수 있게 하는 사용자 추종 기능을 개발하였다. 제안된 사용자 추종 프로그램은 SiamDW 구조의 심층 신경망을 활용하여 RGBD 카메라의 입력을 받아 실시간으로 사용자의 위치를 추적하는 추적 모듈과 현재 사용자의 위치를 입력받아 해당 사용자를 추종하기 위한 로봇의 이동 선속도와 각속도를 계산하는 주행 모듈로 구성된다. 그리고 정적 환경에서 수행한 반복 실험을 통해 개발된 프로그램이 사용자 추종에 효과적이며 다른 사람이 로봇과 사용자 사이를 이동하는 방해에도 강인하다는 것을 입증하였다.

ACKNOWLEDGMENT

본 논문의 연구는 한국전자통신연구원 연구운영지원사업의 일환으로 수행되었음. [20ZD1130, 지능제어기반 스마트 기계 및 로봇 기술 개발].

참 고 문 헌

- [1] Islam, M. J., Hong, J., and Sattar, J., "Person-Following by Autonomous Robots: A Categorical Overview," The International Journal of Robotics Research 38, no. 14, pp. 1581-1618, Dec. 2019.
- [2] Kristan, M. et al., "The Seventh Visual Object Tracking VOT2019 Challenge Results," In Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop, pp. 2206-2241, 2019.
- [3] Zhang, Z. and Peng, H., "Deeper and Wider Siamese Networks for Real-Time Visual Tracking," In Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4591-4600, 2019.
- [4] Li, B., Yan, J., Wu, W., Zhu, Z., and Hu, X., "High Performance Visual Tracking with Siamese Region Proposal Network," In Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8971-8980, 2018.