

# Dead Reckoning과 OCR 기반의 딥러닝 실내 위치 측위 시스템

전두선, 왕수현, 김우현, 엄두섭

고려대학교

jeonds1127@korea.ac.kr, 08shwang@korea.ac.kr, koniwin06@korea.ac.kr, eomds@korea.ac.kr

## Deep Learning Indoor Localization System based on Dead Reckoning and OCR

Du-Seon Jeon, Soo-Hyun Wang, Woo-Hyon Kim, Doo-Seop Eom

Korea Univ

### 요약

GPS를 활용할 수 없는 실내 환경에서 로봇의 Wheel Encoder와 IMU(Imperial Measurement Unit) 센서를 통해서 DR(Dead-Reckoning)을 수행할 때 이동 거리 및 방향각의 부정확성으로 인해 위치 추정 오차가 발생한다. 본 논문에서는 딥러닝을 활용한 OCR(Optical Character Recognition)을 통해서 실내의 호실 정보를 추출하고 주어진 지도와 비교하여 위치를 보정하는 실내 위치 측위 기법을 제안한다. 실험은 고려대학교 신공학관 5층 복도 환경에서 진행하였고, DR과 제안한 DR+OCR 기법을 적용하여 나온 결과를 비교했다. 제안한 DR+OCR 기법이 DR 기법에 비해 평균 위치 오차가 0.78m, 평균 거리 정확도가 2.7%만큼 개선된 성능을 보여준다.

### I. 서론

최근 식당, 카페, 공장 등 실내 공간에서 자율 주행이 가능한 지능형 로봇의 수요가 증가하고 있지만 GPS를 활용할 수 없는 실내 환경에서 로봇의 정밀한 위치 측위가 불가능하기 때문에 실내 자율 주행에 어려움이 있다. 로봇의 실내 위치 측위에 관한 다양한 연구가 진행되어 왔으며 대표적인 방법으로는 Wheel Encoder를 통한 DR 기법, LiDAR 또는 카메라를 통한 SLAM 기법, Wi-Fi AP, Bluetooth Beacon 등의 무선 인프라 또는 3축 자기장 센서를 활용한 Fingerprinting 기법 등을 들 수 있다. 최근에는 딥러닝의 발전으로 카메라를 이용한 영역이 점차 확대되고 있는 추세이다. 본 논문에서는 딥러닝을 적용한 DR과 OCR 기법을 사용하여 실내 위치 측위 시스템을 구성하였다. 자체 제작한 실내 측위 로봇 플랫폼을 실내 환경에서 실험을 통하여 제안한 실내 측위 시스템의 우수성을 보였다.

### II. 본론

#### 2.1. 실내 측위를 위한 로봇 플랫폼

로봇 플랫폼으로 가감속 제어를 위한 ESC 및 brushed DC 모터, 조향 제어를 위한 서보모터가 구성되어있는 Donkey Car를 사용하였다. MCU(Micro Controller Unit)로는 라즈베리파이 4(Raspberry Pi4)와 Arduino Uno를 사용하였다. 로봇 플랫폼의 이동 거리 추정을 위하여 Wheel Encoder를 사용하였으며, 방향각 추정을 위하여 6축 IMU 센서인 MPU6050을 사용하였다. OCR을 하기 위한 용도로 로봇 플랫폼의 상단에 카메라를 위치시키고 해당 카메라의 방향 제어를 위한 서보모터를 구성하였다. 위치 보정에 필요한 딥러닝 OCR 알고리즘은 노트북을 사용하여 처리하였다.

#### 2.2. 알고리즘

그림 1은 제안하는 DR+OCR 기반의 실내 위치 측위 시스템 구조이다. 본 논문에서는 3축 자이로스코프와 3축 가속도계를 포함한 6축 IMU 센서를 사용하여 확장 칼만 필터(EKF: Extended Kalman filter)를 기반으로

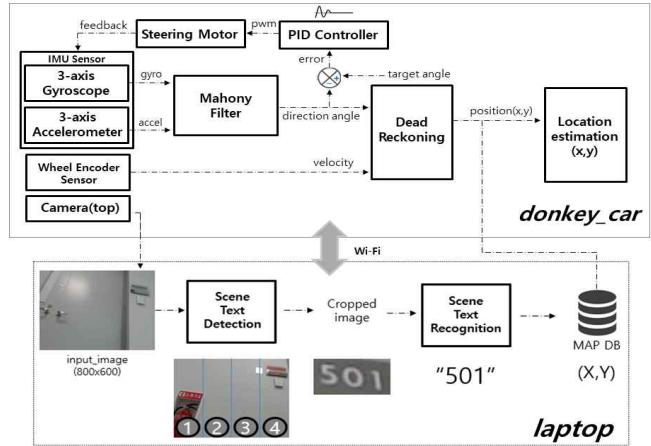


그림 1. 제안하는 DR+OCR 기반의 실내 위치 측위 시스템

하는 마호니 필터(Mahony Filter) [3]를 적용하여 두 센서를 융합함으로써 방향각을 추정하고, Wheel Encoder로 이동 거리를 추정하여 DR을 수행한다. 추가적으로 추정된 현재 방향각을 기준으로 목표 방향각을 추종하는 PID 제어기를 구성하여 정밀한 조향 제어를 가능하게 한다.

OCR의 입력 이미지는 로봇 상단 카메라에서 얻어오게 되는데, OCR은 이미지상에서 글자를 검출하는 STD(Scene Text Detection)와 검출한 글자 박스들을 자르고, 잘린 박스 안에 어떤 글자를 포함하는지를 인식하는 STR(Scene Text Recognition)의 과정을 가진다. 이를 통해서 호실 정보를 추출하여 주어진 지도와 비교하여 위치를 보정하게 된다. 사전에 호실 정보와 로봇의 방향각에 따른 촬영 이미지를 기반으로 DB를 구성했으며, 정밀한 위치 보정을 위해서 STD 단계에서 이미지의 구간을 나눠서 검출된 박스들의 위치 및 로봇의 진행 방향에 따라서 위치 보정을 수행한다.

STD 모델로는 글자 검출에서 높은 성능을 보이는 CRAFT 모델 [1]을 사용했다. CRAFT는 검출 단위를 단어 단위보다 작은 글자 단위를 사용해서 검출 능력을 향상시켰다. 또한 문자의 위치 정보를 사용하고, 인접한 문자 간의 관계성을 사용했다. 네트워크 구조는 backbone으로 Batch

Normalization을 추가한 VGG-16 기반 Fully Convolutional Network를 사용하고, 인코더의 각 레이어 특징들을 skip connection을 이용하여 디코딩 과정에서 문자 검출 성능을 올려준다. STR 모델은 Attention 기반 모델 [2]을 사용했다. NLP(Natural Language Processing) 분야 중 기계 번역에서 많이 사용하던 Sequence-to-Sequence의 Attention을 이미지에 처리하여 픽셀 간의 특징에 적용하여 STR 모델에 사용한다.

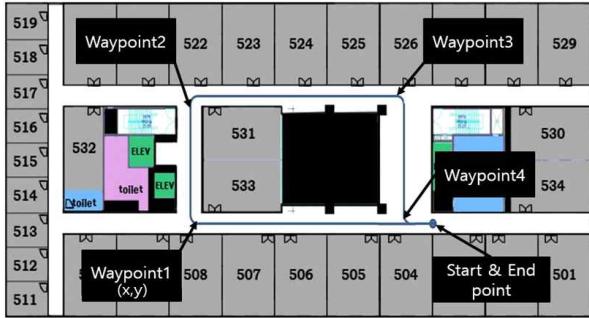


그림 2 고려대학교 신공학관 5층 복도 환경

### 2.3. 실험 환경

실험 환경은 고려대학교 신공학관 5층 복도에서 진행하였다. 그림 2는 실제 건물 도면을 바탕으로 만든 지도이다. 로봇은 출발지로부터 4개의 waypoint를 거쳐서 도착지에 오게 된다. 출발지에서 도착지까지의 총 이동 거리는 92.75m이고, 이동 경로는 직진과 90도 회전으로 이루어져 있다. 출발지는 주어진 상태에서 실험을 시작하며 각각의 waypoint와 최종 도착지에서 실제 위치( $x, y$ )와 DR, DR+OCR 기법의 추정된 위치( $\hat{x}, \hat{y}$ )간의 오차 및 추정된 거리 오차를 계산한다. 여기서 위치 오차는 좌표평면상에서 두 점 사이의 거리를 구하여 나타낸다.

호실 정보를 비교적 높이가 낮은 로봇에서 카메라를 통해 위로 올려다보기 때문에 전등에 의한 빛 번짐 현상이 발생하여 OCR에 어려움을 초래한다. DR+OCR 방식의 OCR을 통한 위치 보정은 호실 번호 바로 위에 전등이 위치해 있지 않은 505호, 507호, 522호, 525호에서만 수행된다. 또한 로봇이 움직이면서 이미지를 얻어오면 카메라 초점이 흔들려서 글자를 인식할 수 없는 문제점이 있기에 로봇이 멈춘 상태에서 이미지를 얻어온다. 따라서 지정된 호실 근처에 왔을 때 멈춤-촬영-움직임 단계를 반복하여 호실 번호를 검출한다. 실험은 총 6회 진행했다.

### 2.4. 비교 및 분석

그림 3은 DR과 제안한 DR+OCR을 실험을 통해 나온 결과 그래프이다. 또한 실제 위치와 DR 및 DR+OCR의 추정된 위치 간에 평균 위치 오차를 표 1에 나타냈다. OCR 보정이 이루어지지 않는 구간들은 제외하고, 출발지에서 waypoint1, waypoint2에서 waypoint3, 출발지에서 도착지까지의 실제 거리와 DR 및 DR+OCR의 추정된 거리 간에 평균 거리 오차 및 정확도를 표 2에 나타냈다. 결과적으로 표 1을 보면 실제 위치와 비교하여 DR은 평균 1.4m, DR+OCR은 0.62m의 평균 위치 오차를 보인다. 따라서 제안한 DR+OCR 기법이 DR보다 평균 위치 오차가 0.78m만큼의 개선된 성능을 보이므로 더 안정적인 위치 추정이 가능한 것을 알 수 있다. 그리고 표 2를 보면 실제 거리와 비교하여 출발지에서 waypoint1은 DR 평균 1.41m, DR+OCR 평균 0.46m, waypoint2에서 waypoint3은 DR 평균 1.56m, DR+OCR 평균 0.06m, 출발지에서 도착지는 DR 평균 3.24m, DR+OCR 평균 0.78m의 평균 거리 오차를 보여준다. 따라서 실제 거리와 비교하여 제안한 DR+OCR 기법이 DR보다 거리 추정 정확도가 2.7%만큼 증가하여 더 안정적인 거리 추정이 가능한 것을 알 수 있다.

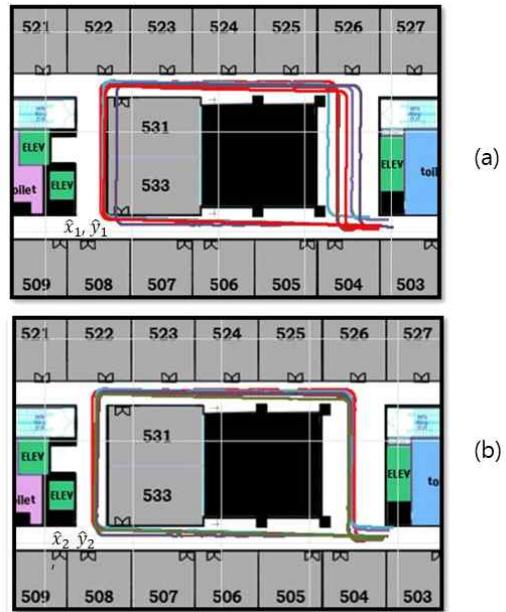


그림 3 실험 결과 (a) DR, (b) DR+OCR

표 1 실제 위치와 DR, DR+OCR간의 평균 위치 오차

Coordinates( $x, y$ )	DR( $\hat{x}_1, \hat{y}_1$ )	DR+OCR( $\hat{x}_2, \hat{y}_2$ )	
waypoint1	1.6 m	0.7 m	0.9 m
waypoint2	1.8 m	0.6 m	1.2 m
waypoint3	1.2 m	0.4 m	0.8 m
waypoint4	1.3 m	0.7 m	0.6 m
endpoint	1.1 m	0.7 m	0.4 m
total average	1.4 m	0.62 m	0.78 m

표 2 실제 이동 거리와 DR, DR+OCR간의 평균 거리 오차 및 정확도

Distance	DR		DR+OCR		
	err[m]	acc[%]	err[m]	acc[%]	
startpoint -> waypoint1 [32.4m]	1.41	95.7	0.45	98.6	2.9%
waypoint2 -> waypoint3 [28.625m]	1.56	94.6	0.06	99.8	5.2%
startpoint -> endpoint [92.75m]	3.24	96.5	0.78	99.2	2.7%

### III. 결론

본 논문에서는 제안한 DR+OCR 방식을 사용하여 차륜 형태의 로봇 플랫폼의 실내 자율 주행을 위한 정밀한 실내 측위가 가능함을 보였다. 제안한 DR+OCR 방식은 기존의 DR 방식에 비해 평균 위치 오차가 0.78m, 평균 거리 정확도가 2.7% 만큼 개선된 성능을 보인다. 실험 환경의 호실 번호는 흰색이어서 빛 번짐에 취약했는데, 글자가 진한 색상이라면 OCR 정확도가 향상될 것으로 예상되며 또한 카메라의 위치를 호실 번호와 비슷한 높이에서 촬영이 가능하다면 보다 좋은 성능을 기대할 수 있을 것으로 예상한다.

### 참고 문헌

- [1] Baek, Youngmin, et al. "Character Region Awareness for Text Detection." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [2] Baek, Jeonghun, et al. "What is wrong with scene text recognition model comparisons? dataset and model analysis." arXiv preprint arXiv:1904.01906 (2019).
- [3] R. Mahony, T. Hamel, and J.-M Pflimlin, "Nonlinear complementary filters on the special orthogonal group," IEEE Trans. Automatic Control, vol. 53, no. 5, pp. 1203–1218, Jun. 2008.