

실내 문서 전달 자율주행 카트 개발을 위한 텍스트 추출 및 객체 인식 모듈, 회피 모듈 개발

이재빈, 김명현
경희대학교 컴퓨터공학과

Implementing Text and Object Detect module For The Indoor Automatic driving Document cart

요 약

최근 머신러닝 기술과 HW의 발전으로 ‘자율주행’에 관한 연구가 활발히 이루어지고 있다. 또한, 이에 관한 응용분야 역시 단순히 노상이 아닌, 실내에서도 충분히 적용가능 함에 따라, 본 문서는 실내문서전달용 자율주행 카트 개발을 위한 실시간 영상처리 및 머신러닝 모듈 개발을 목표로 한다. 나아가 최종적으로는 수집된 실내정보와 클라우드 컴퓨팅을 기반으로 정확한 실내정보를 인지하고 이에 맞추어 동적으로 동작할 수 있는 자율주행 카트 개발을 목표로한다.

1. 서 론

1.1 연구 배경

5G network 환경과 더불어 머신 러닝과 같은 SW기술 발전으로 인해 점차적으로 자율 주행 기술의 점진적 변화가 예상된다. 이에 미국 자동차 학회(SAE)에서는 자율 주행의 구현 정도에 따라 자율주행 기준 발전단계를 마련하였으며 2020년 현대자동차에서는 향후 2년 이내에 주행관련 정보를 판단하고 행동을 결정할 수 있는 레벨3단계를 대부분 구축할 수 있을 것으로 보고 있다. 앞서 보듯, 실외용 Autonomous Vehicle의 연구가 현재 활발히 진행되고 연구되고 있으나, 도로 위가 아닌 실내의 경우 또 다른 문제점에 직면하게 된다. 먼저 Guide Line인 차선이 없다는 점과 더불어, 비좁은 통로가 존재하고 사람들의 왕래가 실내에서는 더욱 잦다는 점을 감안하면 단순히 위치정보만을 활용한 자율주행카트 구현에는 무리가 있다. 따라서 본 연구는 기존의 실외 자율주행카트의 개념에 덧붙여 실시간 영상처리를 통한 Number Detection, Human Oriented Driving시스템을 구축하고 Edge Computing기술을 덧붙여보다 정확한 실내 주행과 원격제어가 가능한 카트를 제안한다.

1.2 연구 개발 목표

클라우드 컴퓨팅 기술을 기반으로 대략적인 위치정보를 수신하며, 이와 동시에 전면, 측면에 대한 실시간 영상처리를 통해 보다 정확한 위치정보를 얻을 수 있도록 자율 주행 카트 모듈을 구현한다. 이에 덧붙여, 실내에서의 다양한 특징 점을 가진 텍스트(실내의 각 태그)를 기반으로 위치정보를 얻을 수 있도록 기기를 특화한다. 또한 영상처리 모듈을 통해 실시간 객체 및 사람 인식 기능을 구현하여 1)Human Oriented Driving System을 구현한다.

2. 기존 연구

2.1 SLAM(Simultaneous Localization And Mapping)기반 자율주행 카트

초음파, 적외선, 관성 측정 장치인 IMU와 LIDAR, 모터 등을 사용하여 제어 가능한 로봇에 대해연구가 이루어진 바 있다. 이 원격제어 로봇은 LIDAR센서를 통해 물체와의 거리를 측정하여 장애물 회피기능을 구현하였다. 다만 2)LIDAR센서가 Object의 반사펄스를 이용하여 물체의 거리를 측정하는 특성을 고려하면, 여전히 반사율이 높은 환경에서 정확한 거리인식이 어렵고, 이에 따른 우발적 장애요소에 대한 회피문제점이 잔존한다.

2.2 클라우드 연계 자율주행 맵 시스템 기술 동향

클라우드와 연계하여 생성되는 자율주행 맵을 구성하면 보다 정밀성 있는 맵을 구성할 수 있다. 또한 도로 교통과 관련된 실시간 갱신주기 측면에서 더욱 세밀하고 정확한 위치정보를 제공할 수 있다. 그러나 실내에서의 경우 주행 경로 상의 장애물이 있을 가능성이 높고, 좁은 통로의 특성상 많은 우발적 요소가 예상되기 때문에 정적인 지도만을 구축하여 주행하기에는 한계점이 있다.

2.3 위치 추정, 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율 주행 기술

실내 이동로봇의 자율주행 방법으로 지도생성, 위치 추정, 장애물 회피, 경로 계획에 대한연구가 이루어진 바 있다. 위치 추정을 위해서 지도 정보를 이용, 센서 데이터를 계산하였고 인공지능위계를 사용하여 장애물로부터의 척력, 목표위치로의 인력을 구하여 3)회피 알고리즘을 구현하였다. 그러나 4)LRF센서를 바탕으로 장애물을 인식하였고 때문에

2) 레이저 펄스의 반사 시점까지의 시간을 측정하여 반사체의 위치좌표를 측정하는 센서.

3) 해당 연구에서는 이동로봇의 위치추정방법으로 MCL알고리즘, 삼각측량, 칼만 필터 등을 사용하였음.

1) 이동객체(사람) 회피 주행

실내의 좁은 통로 특성상 센서 길이의 제약으로 인해 장애물 감지거리가 55cm에 불과하였다.

2.4 Multi-digit Number Recognition

CNN을 이용해 이미지에서 여러 자리의 숫자를 감지하는 연구는 있었지만, 사용된 데이터셋의 숫자 폰트가 너무 다양하고, 본 연구에서는 쓰이지 않는 숫자의 범위까지 포함되는 등의 이유로 모델 예측의 레이턴시가 커지는 위험이 있다. 따라서 실시간으로 영상을 크롭하여 숫자를 추출해야 하는 본 연구에 적합하지 않은 점이 있다.

2.5 기존 연구 분석

구분	연구내용	한계점
SLAM 기반 자율주행 카트	IMU, LIDAR, 모터를 사용하여 자율 주행 구현	반사펄스를 이용하는 LIDAR, 초음파 센서의 특성상, 반사율이 높은 환경에서 우발적 상황에 대한 한계점 잔존
클라우드 연계 자율주행 맵 시스템 기술동향	클라우드 기반 지도 구축 및 도로상의 엣지 클라우드-카트 간 상호작용을 통한 자율 주행 구현	실내의 경우 주행 경로 상에 장애물이 있을 가능성과 좁은 도로 폭을 고려해 볼 때 정적인 지도만을 바탕으로 한 자율주행에는 많은 어려움이 존재함
위치 추정 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율 주행 기술	LRF 센서를 바탕으로 장애물 인식, 인공 전위계를 사용하여 회피 알고리즘 구현	실내의 좁은 통로 특성상 LRF 센서의 길이가 짧아짐에 따라 장애물 감지거리가 약 55cm에 불과하여 실질적인 자율주행에는 어려움이 있음
Multi-digit Number Recognition	CNN을 이용해 이미지에서 여러 자리의 숫자를 감지하는 연구	실시간 영상을 크롭하여 숫자를 추출하기에는 한계점이 잔존

표 1 기존연구 분석

3. 연구 내용

3.1 시나리오 개념도

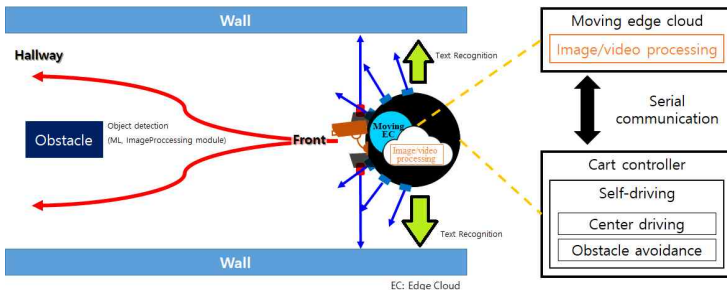


그림1 주행 개념도(위)

4) LIDAR 센서와 동일한 원리로, 레이저펄스를 이용하여 반사체 간의 거리를 측정하는 센서.

5) 참고문헌[5] 참조

6) SVHN(Street View House Numbers) Dataset이 사용되었다.

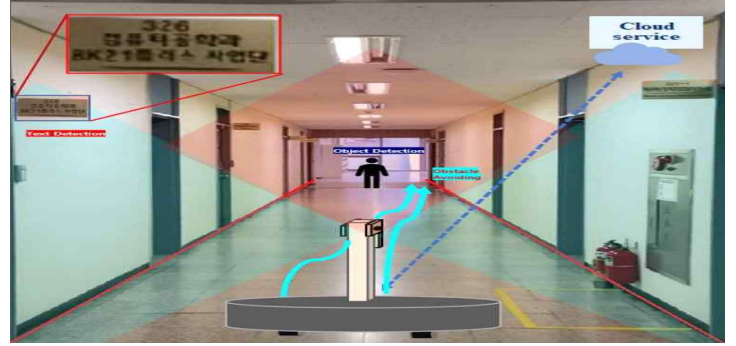


그림2 주행개념도 (후면)

3.2 세부 시나리오

각종 센서와 라즈베리파이 기반으로 구현된 자율주행 카트에 양 측면 그리고 전방 카메라를 장착하여 실시간으로 영상처리를 지원한다. 이는 로컬디바이스에서 처리되며 실시간 처리결과에 따라 카트가 자체적으로 행동결정을 하게 되고 주행 중 위치정보는 Cloud의 각Edge에 의해 보다 세밀화 된다. 각 카메라 렌즈에서 입력되는 영상은 실시간 영상처리 모듈에 의해 Detection과 Text추출이 이루어진다. 먼저, 전면 카메라는 물체 Detection에 활용되며 Object Detection 모듈에 의해 실시간 영상으로부터 객체를 회피한 주행경로를 출력하여 제어모듈로 전송한다. 다음으로 양 측면 카메라는 실내의 정보파악을 위한 Text추출 기능을 한다. Text추출 모듈의 정보인식은 영상의 각 프레임에서 Crop된 이미지에 대해서 이루어진다. 이때, 기존 OCR알고리즘은 Overhead가 크므로 라즈베리파이의 Local Device에 로드된 머신러닝 모델에 의해 예측된다. 이후 추출된 Text와 갖고 있는 DB정보를 비교하여 보다 세밀화 된 정보식별을 하게 된다. 최종적으로는 주행에 관한 정보가 도출되고 해당 정보를 카트의 제어모듈로 전송, 이를 기반으로 동적인 실내주행이 구현된다.

4. 요구사항

4.1 지연 시간 요구 사항

현재 pytesseract의 OCR라이브러리는 720*960 이미지 파일 기준, Text Extraction의 지연시간은 평균 0.38초로 측정되었다. 라즈베리파이의 경우 CPU, GPU환경이 좋지 않아 많은 Latency가 예상되므로 기존 OCR라이브러리 보다는 학습된 머신러닝 모델을 통하여 예측하는 것이 바람직할 것으로 보인다. 현재 라즈베리파이 환경에서의 실시간 Human Detection 실행시간은 약 0.05 - 0.06 초로 실시간 Object인식에는 무리가 없다. 다만 Text추출이나 장애물과의 거리측정의 경우 더 높은 해상도와 머신러닝 모델의 지원이 필요한 부분이므로 학습모델의 경량화와 프레임 조절 및 정확도 높은 크롭(Crop)이 요구된다.

4.2 객체 인식 요구사항

Object Detection의 경우 가장먼저 Human Oriented로 구현되어야 하며, 사람 중심으로 철저히 구동되어야 한다. 최소 5m에서 최대 25m거리에서 사람을 인식할 수 있어야 하며, 우발적인 동작에도 빠르게 반응할 수 있어야 한다. 이외에도 다른Object를 장애물로 인식하여 피해갈 수 있도록 카트의 HW제어 모듈에 해당 정보를 전달한다.

4.3 실내 Number Tag정보 인식 요구사항

Number Detection의 경우 라즈베리파이의 카메라모듈 프레임이

7) 320*240 해상도, Raspberry pi 4모델 기준으로 측정.

초당 30프레임을 감안하여 최소 4-5%이상의 이미지 크롭 정확도를 요구한다. 하지만 이는 최소범위이므로, 원활한 주행을 위해서는 약 40%이상의 이미지 크롭 정확도를 요구한다. 머신러닝 모델의 경우 약 80%이상의 정확도를 유지하여, 실내 환경에서의 Text를 예측하고 8)잡음을 걸러낼 수 있어야한다.

4.4 Number 인식용 데이터셋 요구사항

Number Detection의 용도로 사용될 데이터셋은 3~4자리의 숫자로 이루어진 이미지 집합을 요구한다. 다만 기존 연구에서 사용되었던 SVHN 데이터셋은 숫자의 범위가 너무 넓고 폰트가 다양하여 적합하지 않으므로, 본 연구에 특화된 데이터셋을 새로 구축할 필요가 있다. 1층에서 5층, 3~4자리의 숫자로 고른 분포를 가진 데이터셋이 요구된다.

5. 구현 방안

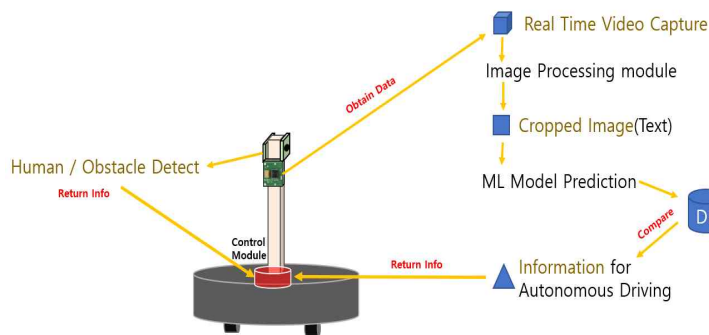


그림3 모듈간 데이터 처리과정



그림4,5 실내 태그 크롭이미지(좌) 및 사람 인식(우)

· Number Detection : 양 측면 카메라에서 전달받은 실시간 RGB 영상에서 Number Tag에 해당하는 이미지를 48*48 사이즈로 크롭한다. 크롭된 이미지는 Grayscale로 변환되어 CNN 모델에 입력된 후, 모델은 6개의 결과 값을 도출한다. 이 때 6개의 결과 값은 이미지가 숫자를 포함하는지 여부, 인식된 숫자의 길이, 최대 4자리의 숫자 값으로, 이 결과 값을 통해 실내 정보 데이터베이스와 비교하여 현재 카트의 정확한 실내 위치를 카트의 제어모듈로 전송한다.

· Human/Obstacle Detection : 전방 카메라에서 받은 실시간 영상으로부터 OpenCV의 Hog디스크립터를 이용하여 사람을 인식한다. 이외에 특정 크기 이상의 장애물을 인식하여 현재 카트의 주요 장애물 동선을 10개 요소를 가진 1차원배열로 작성하고 카트의 너비를 고려, 가장 이상적인 경로 정보를 1차원 배열 형태로 제어모듈에게 전송한다.

6. 결론 및 향후 연구

실내용 자율 주행 카트를 개발함에 있어서, 이미 저장된 지도정보가 아닌, 직접 학습을 하고 동적으로 상황에 따라 유연한 대응을 하는 자율주행카트를 개발함으로써, 실내 환경에 최적화된 자율주행 시스템을

연구할 수 있다. 또한 근거리의 경우 센서를 활용한 Object Detection과, 원거리의 경우 영상처리 및 머신러닝 모델에 의한 예측으로 우발적인 상황에는 하드웨어적인 명령으로 빠른 대처가 가능하며 실시간 영상처리를 기반으로 향후 이동경로를 미리 예상할 수 있음을 감안하면 정적 로드맵을 구축하는 것보다 훨씬 효율적인 주행을 기대할 수 있다. 향후 연구로는 실제 문서전달 기능을 감안하여 하드웨어 및 소프트웨어적으로 최적화된 기기를 구현할 계획이다.

참 고 문 헌

- [1] 김용년, and 서일홍. “바닥 특징점을 사용하는 실내용 정밀 고속 자율 주행 로봇을 위한 싱글보드 컴퓨터 솔루션” 로봇학회논문지 14.4 (2019): 293-300.
- [2] 박세준, 서용호, and 양태규. “위치센서와 레이저거리센서를 이용한 실내용 이동로봇의SLAM” 한국정보기술학회논문지 8.11 (2010): 61-69.
- [3] 노성우, 고낙용, 김태균, Noh, Sung-Woo, Ko, Nak-Yong, and Kim, Tae-Gyun. “위치 추정, 충돌 회피, 동작 계획이 융합된 이동 로봇의 자율주행 기술 구현.” 한국전자통신학회 논문지 = The Journal of the Korea Institute of Electronic Communication Sciences 6.1 (2011): 148-56. Web.
- [4] 최정단, 민경옥, 성경복, 한승준, 이동진, 박상현, 강정규, 조용우, Choi, J.D., Min, K.W., Sung,K.B., Han, S.J., Lee, D.J., Park, S.H., Kang, J.G., and Jo, Y.W. “클라우드 연계 자율주행 맵 시스템 기술 동향” 전자통신동향분석 = Electronics and Telecommunications Trends 32.4(2017): 40-47. Web.
- [5] Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks, Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibraz, Sacha Arnoud, Vinay Shet, 2011

8) 실제 Text정보가 아닌, 반사율과 같은 여러 외적요인에 의해 잘못 크롭된 이미지