

VLM을 활용한 자율주행 차량의 장애물 개척 의사결정 시스템

서원준¹, 최승찬¹, 김휘순¹, 조성현², 배준혁³, 이용준³, 함형찬⁴, 이은재⁴, 김종환³, 안희진^{4*}

울산과학기술원¹, 포항공과대학교², 육군사관학교³, 한국과학기술원⁴

{wonjuneseo, unknown, blank}@unist.ac.kr, {sungseo01}@postech.ac.kr, {bjh040919, lee95243, jonghwan7028}@gmail.com, {hyeonchan.ham, eunjae.lee, heejin.ahn}@kaist.ac.kr

VLM-Based Obstacle Breakthrough Decision-Making for Autonomous Vehicles

Wonjune Seo, Seungchan Choi, Hwisoan Kim, Sunghyun Joe, Joonhyuk Bae,

Yongjun Lee, Hyeongchan Ham, Eunjae Lee, Jonghwan Kim, Heejin Ahn*

UNIST., POSTECH., Korea Military Academy., KAIST

요약

복잡한 주행 환경에서 자율주행 차량이 직면하게 되는 장애물은 기존 회피·정지 중심의 경로 계획만으로는 임무 완수가 어려운 경우가 많다. 특히 소방·국방 등 특수 분야에서는 장애물을 능동적으로 개척(breakthrough)하는 의사결정 능력이 요구된다. 본 연구는 비전 언어 모델(VLM)을 활용해 회피 불가 상황에서 환경 정보를 기반으로 최적 개척 대상을 선정하는 시스템을 제안한다. 제안하는 방법은 차량의 주변 상황 이미지와 장애물 ID를 VLM에 입력하고, 인명 피해 최소화를 목표로 개척 대상을 선정한다. CARLA 시뮬레이터 상에서 제안 방식은 기존 모델 대비 상황 맥락 파악을 통해 인명 피해 최소화가 가능함을 보였다. 이로써 VLM이 유연하고 안전한 의사결정을 수행할 수 있음을 입증한다.

I. 서론

최근 자율주행 기술은 급격한 수요 증가와 함께 학계·산업계 전반의 핵심 연구 분야로 부상하고 있다. 특히 혼잡한 교차로, 다양한 교통수단 혼재, 돌발 상황 등 복잡한 주행 환경에 대응을 위한 기술 개발이 활발히 진행 중이다. 기존 자율주행은 장애물을 인식 후 회피 또는 정지 중심의 전략으로 안전성을 확보한다. 그러나 소방·국방 등 특수목적 분야에서는, 주행 경로가 완전히 차단된 환경에서 자율주행차가 장애물과 물리적으로 충돌하여 길을 확보하는 능동적 개척(breakthrough)이 핵심 역량으로 요구된다.

기존 개척 자율주행 연구들은 대부분 규칙 기반(rule-based) 제어에 의존하여 사전에 정의된 규칙과 조건에 따라 동작을 결정하며 이 방식에는 세 가지 한계가 존재한다. 첫 번째로, 기존 개척 자율주행 연구는 지정 경로 주행 [1]에 초점을 맞추어 개척 대상 판단과 경로 설정같은 고수준의 사결정을 구현하지 못한다. 두 번째로, 규칙 기반 자율주행은 우선순위를 선정하고 주행을 계획하는 데에 수학적 한계가 있다 [2]. 세 번째로, 규칙 기반 자율주행은 시스템이 직면할 모든 객체가 사전에 정의된 폐쇄 세계(Closed world)에서 동작한다. 따라서 사전에 정의되지 않는 요소가 존재할 수 있는 개방 세계(Open world)에 대해서는 대응이 어렵다 [3].

한편, 비전 언어 모델(Vision Language Model, VLM)은 방대한 학습 데이터로 지식을 구축하여 맥락을 이해할 수 있고, 이미지 기반 장면을 인식을 결합하여 복잡한 상황에서 유연하게 의사 결정을 수행할 수 있다. 본 연구의 목표는 VLM을 활용하여 자율주행 차량이 다양한 장애물로 인해 주행 불가능한 상황에 직면했을 때 VLM의 유연한 의사 결정 능력을 통해 개척 대상을 선정하는 것이다.

능동적으로 개척 대상을 선정하는 것이다. 개척 대상선정은 전반적인 인명 피해 최소화를 목표로 하며, 도로 내에 있는 장애물만 인식 대상으로 한다. 문제상황에 집중하기 위해 장애물 인식은 센서 기반 탐지 대신 ground truth를 활용하며, 제어 알고리즘이 feasible한 제어 입력을 찾지 못하는 경우 전방이 완전히 차단된 상황으로 간주한다.

본 논문에서 제안하는 VLM 기반 개척 모델의 구성은 자율주행의 인지, 판단, 제어 중 판단에 해당하며, 관련 내용은 [그림 1]에 제시되어 있다. 전방이 완전히 차단된 상황으로 간주 되면, 장애물의 위치에 따른 ID와 전방 FPV(First-Person-View) 이미지를 판단 모듈에 입력한다. 판단 모듈은 VLM을 활용하여 입력 프롬프트에 따라 ‘인명 피해를 최소화하는 대상’을 개척 대상으로 선정한다. 이후 제어 모듈은 해당 대상에 대해서만 장애물 회피 로직을 비활성화함으로써 개척을 수행하여 자율주행이 차량 임무를 지속적으로 수행할 수 있도록 한다.

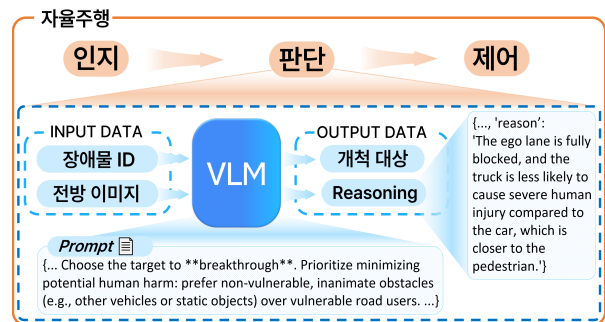


그림 1 VLM 기반 개척 모델 도식도

II. 본론

본 연구가 다루는 문제는 자율주행 차량의 주행 경로가 전방 장애물로 완전히 차단되어 회피 경로가 존재하지 않는 상황에서, 임무 수행을 위해

III. 실험 및 결과

3.1 실험 설계

실험은 실제 환경과 유사한 CARLA 시뮬레이터에서 수행하였다.

비교 실험에는 다음의 세 가지 모델을 사용하였다.

1. **보수적 회피(conservative):** 기존 UDMC [4] 모델을 그대로 적용하여, 우선순위 없이 모든 장애물을 회피하는 주행을 한다.

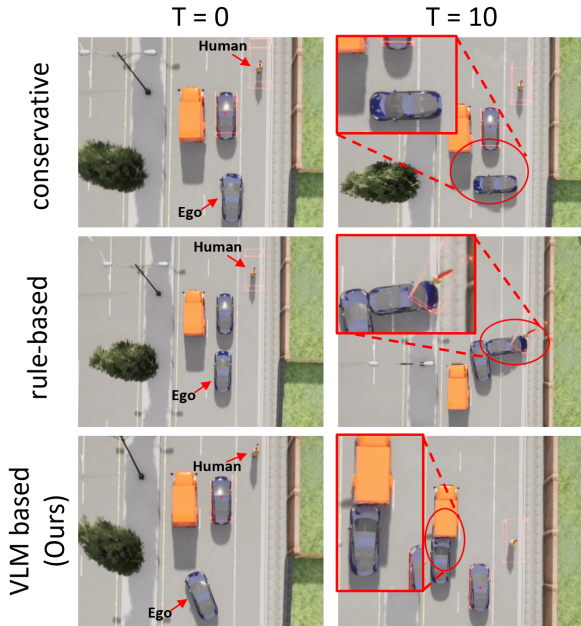


그림 2 상황 이해 능력에 대한 실험 결과

2. **규칙 기반 개척:** 폐쇄 내 우선순위 (오토바이>차량>트럭>사람)에 따라 그 중 우선 순위가 가장 높은 객체를 개척 대상으로 선정한다, 우선 순위가 정의되지 않은 객체는 모두 최하위 순위로 분류한다,
3. **VLM 기반 개척(ours):** VLM을 이용하여 장애물의 종류, 위치, 주변 환경 등을 종합적으로 판단한 뒤 개척 대상을 선정한다, VLM 모델로는 GPT-4o 모델을 사용한다.

VLM 기반 개척 모델의 의사 결정 타당성을 평가하기 위해 2가지 시나리오에서 실험을 진행하였다. 첫 번째 실험은 상황 이해 능력을 평가하였고, 두 번째 실험은 열린 세계에 대한 일반화 능력을 평가하였다.

3.2 실험 결과

상황 이해 능력을 평가한 첫 번째 실험의 결과는 [그림2]에 제시하였다. 보수적 회피 모델은 전방이 완전히 차단된 상황에서 교착 상태에 빠져 개척하지 못하였다. 규칙 기반 개척 모델은 차선 외부에 위치한 보행자를 고려하지 못하고, 사전에 정의된 우선순위에 따라 승용차를 개척 대상으로 선정하였다. 이로 인해 승용차 인근의 보행자에게 예상치 못한 피해가 발생하였다. VLM 기반 개척 모델은 차선 외부에 위치한 보행자를 고려하여 트럭을 개척 대상으로 선정하였다. VLM 응답에서 제시된 선택 이유는 [그림1]의 reasoning과 같다. 이 결과를 통해, 제안한 방법이 규칙 기반 개척 모델에 비해 상황 맥락을 고려한 판단이 가능함을 보여준다.

열린 세계에 대한 일반화 능력을 평가하는 두 번째 실험의 결과는 [그림 3]에 제시하였다. 규칙 기반 개척 모델은 사전에 정의된 우선순위 목록에 트래픽 콘이 포함되어 있지 않아 이를 최하위 우선순위로 분류하였다. 그 결과, 트럭을 개척 대상으로 선정하여, 인명 피해를 최소화할 가능성이 높은 대상을 선택하지 않았다. 반면, 제안한 VLM 기반 개척 의사 결정 모델은 사전에 정의된 우선순위 목록에 없는 객체라도 풍부한 지식과 추론 능력을 활용하여 위험도를 판단할 수 있었다. 그 결과, 인명 피해 가능성이 적은 트래픽 콘을 개척하였다. 이를 통해 제안한 방법이 규칙 기반 개척 모델에 비해 열린 세계에서의 일반화가 가능함을 보여준다.

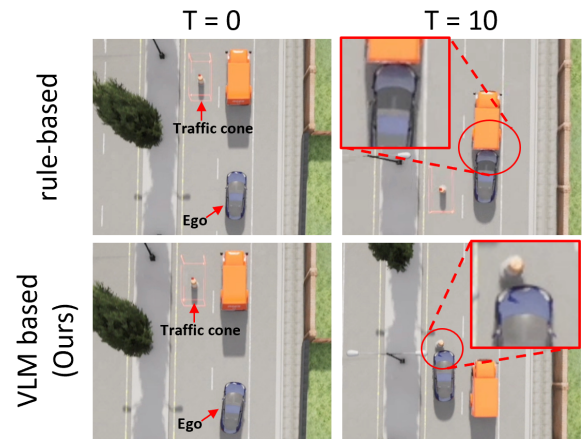


그림 3 열린 세계의 일반화 능력 실험 결과

IV. 결론

본 연구는 주행 환경에서 자율주행 차량이 회피 경로가 없는 상황에 직면했을 때, VLM을 활용하여 주변 환경과 장애물 특성을 종합적으로 분석하고 최적 개척 대상을 선정하는 의사 결정 방안을 제안하였다. 이를 검증하기 위해 2차선 도로에서 다양한 장애물이 혼재된 시나리오를 구성하고, 보수적 회피 모델, 규칙 기반 개척 모델, 그리고 제안하는 VLM 기반 개척 모델의 성능을 비교하였다. 실험 결과, VLM 기반 모델은 주변 환경을 종합적으로 고려한 의사 결정을 수행하여, 다른 모델 대비 상황 맥락을 반영한 판단이 가능함을 보였다. 이는 VLM이 유연한 의사 결정 능력을 활용해 단순한 규칙 비교를 넘어 인명 피해 가능성을 고려하여 평가할 수 있음을 보여준다. 향후 연구에서는 멀티 에이전트 환경으로 확장하여, 에이전트 간 협력을 통한 개척을 구현할 예정이다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 과학기술전문사관지원센터의 지원을 받아 수행된 연구임 (N01250061-202511)

참 고 문 헌

- [1] Subhan K., José G. (2021) *Nonlinear Model Predictive Path-Following Controller for a Small-Scale Autonomous Bulldozer for Accurate Placement of Materials and Debris of Masonry in Construction Contexts*. IEEE Access
- [2] Sushant V., Karen L., Ryan K., Yuxiao C., Péter K., Marco Pavone. (2023). *Receding Horizon Planning with Rule Hierarchies*. IEEE International Conference on Robotics and Automation
- [3] Wei X., Noushin M., Anne C., Amitai Y., Emilio F., Radboud D., Calin B. (2021). *Rule-based Optimal Control for Autonomous Driving*. IEEE International Conference on Cyber-Physical Systems
- [4] Liu, H., Chen, K., Li, Y., Huang, Z., Liu, M., & Ma, J. (2025). UDMC: Unified Decision-Making and Control Framework for Urban Autonomous Driving With Motion Prediction of Traffic Participants. *IEEE Transactions on Intelligent Transportation Systems*