

계층적 트랜스포머 기반 자율 이동체 경로 계획 및 제어

조예령, 김중헌

고려대학교

joyena0909@korea.ac.kr, joongheon@korea.ac.kr

A Transformer-Based Algorithms on Path Planning and Control for Autonomous Mobility

Yeryeong Cho, Joongheon Kim

Korea University

요약

자율주행 의사결정 및 제어는 환경 다양성, 실시간성, 안전성을 동시에 만족해야 하는 고난도 문제이다. 최근 연구는 트랜스포머 기반 시퀀스 모델링을 활용하여 멀티모달 융합과 전이학습을 결합함으로써 일반화 성능과 개발 효율을 동시에 향상시키는 방향으로 활발히 연구되고 있다. 특히, 사전학습된 트랜스포머의 미세조정을 통해 이중 차량 동역학과 다양한 도로 상황에 적응하는 적응적 알고리즘에 대한 연구가 제안되었다. 본 논문은 경로 계획과 제어 관점에서 이러한 트랜스포머 기반 방법들을 분석하고, 주요 핵심 연구 방향을 정리한다. 마지막으로, 연구 분야 분석을 통해 향후 연구를 위한 기술적 쟁점을 시사한다.

I. 서론

자율주행 의사결정 및 제어는 복잡한 환경 다양성, 엄격한 실시간 제약, 안전성 요구를 동시에 만족해야 하는 통합적 연구 과제이다. 전통적 파이프라인은 지각, 예측, 계획에 각각 기반했으나, 상호의존성이 높은 시나리오에서 전역 맥락을 충분히 반영하기 어렵다는 한계가 지적되어 왔다 [1]. 이러한 한계에 의해, 최근 트랜스포머 기반 시퀀스 모델링이 장기 의존성 처리와 전역 추론 능력을 바탕으로 멀티모달 융합을 자연스럽게 수행하며, 파이프라인 간 경계를 축소하는 방향으로 발전하고 있다. 특히 사전학습 후 미세조정 패러다임이 확산되면서, 대규모 멀티모달 데이터로 학습된 트랜스포머를 다양한 차량 동역학과 도로 상황에 적응시키는 연구가 활발하다. 이 접근은 학습 비용을 절감하고 이중 플랫폼으로의 전이를 가속할 수 있다는 장점을 가지며, 경로 계획과 저수준 제어를 아우르는 단일 또는 계층적 정책 설계로 이어지고 있다 [2].

이 중 제안된 계층적 트랜스포머 알고리즘은 고수준 목표 생성과 저수준 궤적 보간을 분리하여 추론 지연과 복잡도를 완화하는 경향을 보이며, 새로운 패러다임으로 제안되었다. 본 논문은 이러한 흐름을 경로 계획과 제어 관점에서 관련 연구를 분류 및 분석하고, 적용상의 장단점과 한계를 비교 분석한다. 마지막으로, 실시간성, 안전성 및 일반화의 균형 등 향후 연구를 위한 기술적 쟁점을 제시한다.

II. 계층적 트랜스포머 기반 경로 계획 및 제어 연구 개념

트랜스포머를 경로 계획과 제어에 적용하는 최근 연구는 지각, 계획, 제어를 개별 모듈로 고정 분할하기보다 멀티모달 정보를 시계열로 통합하여 하나의 정책으로 다루는 방향으로 발전한다. 이는 카메라, 라이다 등의 센서를 통해 인지된 환경에서 추출된 특징을 공통 표현 공간으로 정렬하고, self-attention으로 전역 맥락과 상호작용을 동시에 고려함으로써 장기 의존성과 즉시 반응을 함께 만족시키려는 접근이다 [3]. 이를 위해 모델과

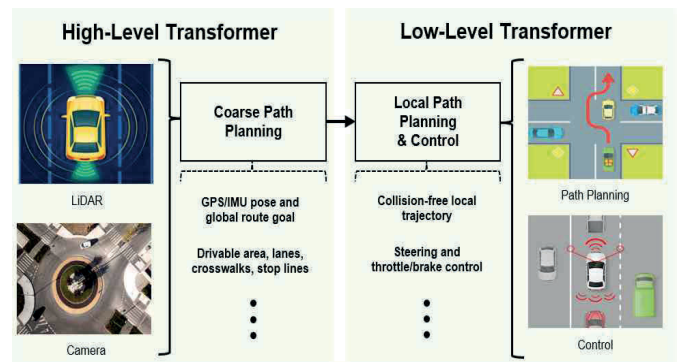


그림 1. 계층적 트랜스포머 알고리즘 구조도

플래너의 동시 설계를 전제로 하며, 과거 맥락을 재활용하고 재계산을 최소화한다. 이러한 설계는 입력 길이가 증가해도 지연이 선형적으로만 증가하도록 완화할 수 있으며, 동일 모델의 파라미터를 유지한 채 다양한 센서 구성과 차량 동역학에 비교적 쉽게 전이 가능한 이점을 제공한다 [4].

계층적 트랜스포머 구조는 이 흐름을 한 단계 더 구체화한다. 그림 1과 같이, 상위 모듈은 수 초 단위의 전역 웨이포인트와 중간 목표를 생성하여 차선 구조, 교차로 위상, 주변 행위자와의 상호작용을 요약된 형태로 제공한다. 이에 반해 하위 모듈은 이 요약 정보를 받아 짧은 지평에서의 조향과 가속을 산출하며, 제동 한계, 조향 각속도 등 동역학 제약을 직접 반영한다. 이렇게 분리된 구조는 유효 시퀀스 길이를 줄여 메모리 사용량과 추론 지연을 동시에 낮추고, 상위 모듈의 주기를 분리하여 재계획 빈도를 높여 정확도를 높일 수 있다. 결과적으로 급격한 장면 전환, 가시성 저하, 돌발 합류와 같은 난상황에서도 상위은 전역 일관성을 유지하고 하위는 미세한 안정화에 집중하여 전체 시스템의 강건성을 높인다 [5].

이렇게 계층적 알고리즘은 크게 3개의 장점을 가진다. 첫째, 전역 목표

참 고 문 헌

- [1] Felipe Codevilla, Matthias Müller, Antonio López, Vladlen Koltun, and Alexey Dosovitskiy, "End-to-End Driving Via Conditional Imitation Learning", in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, May 2018, pp.1-9.
- [2] Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch, "Decision Transformer: Reinforcement Learning via Sequence Modeling", In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, Sydney, Australia, December 2021, pp.1-14.
- [3] Lili Chen, et al., "Decision Transformer: Reinforcement Learning via Sequence Modeling", in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, Sydney, Australia, December 2021, pp.15084-15097.
- [4] K. Chitta, A. Prakash, B. Jaeger, Z. Yu, K. Renz and A. Geiger, "TransFuser: Imitation With Transformer-Based Sensor Fusion for Autonomous Driving," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 11, pp. 12878-12895, November 2023.
- [5] Zikang Zhou, Luyao Ye, Jianping Wang, Kui Wu, and Kejie Lu, "HiVT: Hierarchical Vector Transformer for Multi-Agent Motion Prediction", in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, USA, June 2022, pp.8823 - 8833.
- [6] Shaoshuai Shi, Li Jiang, Dengxin Dai, and Bernt Schiele, "Motion Transformer (MTR): Global Intention Localization and Local Movement Refinement", in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, November 2022, pp.6531-6543.
- [7] Michael Janner, Qiyang Li, and Sergey Levine, "Offline Reinforcement Learning as One Big Sequence Modeling Problem (Trajectory Transformer)", in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, Sydney, Australia, December 2021, pp.1-14.
- [8] Yihan Hu, et al., "Planning-Oriented Autonomous Driving (UniAD)", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, June 2023, pp.17853-17862.
- [9] Ashish Vaswani, et al., "Attention Is All You Need", in *Proc. Advances in Neural Information Processing Systems (NIPS)*, Long Beach, USA, December 2017, pp.6000-6010.

ACKNOWLEDGMENT

본 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2022-0-00907, (2세부) AI Bots 협업 플랫폼 및 자기조직 인공지능 기술개발). 본 논문의 교신 저자는 김중현임.

생성과 저수준 제어를 분리함으로써 시퀀스 길이와 계산 복잡도를 감소시키고, 동일 지연 예산에서 더 긴 맥락을 활용할 수 있게 한다. 둘째, 상하위 모듈의 입력 구성을 차별화하여 상위에는 멀티모달 전역 맥락을, 하위에는 동역학 및 차량 상태 중심의 근거리 정보를 집중시켜 안정성을 극대화한다. 마지막으로, 상·하위 모듈의 학습 목표를 구분함으로써 상위는 전역 일관성, 하위는 안정성과 승차감 같은 세부 목적을 각각 최적화하도록 하여, 전체 성능을 균형 있게 끌어올릴 수 있다 [6]. 결론적으로, 트랜스포머 기반 경로 계획 및 제어는 멀티모달 시퀀스 모델링과 계층적 분해를 통해 전역 추론, 실시간성, 안전성을 동시에 추구하는 방향으로 수렴한다. 계층적 트랜스포머는 계산 예산과 지연 제약을 고려한 실용적 구조를 제공하면서, 실제 도로 환경에서 요구되는 빈번한 재계획 및 센서의 상태 변화에 대응하기 위한 필수적인 선택으로 볼 수 있다.

III. 계층적 트랜스포머 기반 실증 연구 동향

최근 실증 연구는 전통적 환경 인지, 후보 생성, 비용 평가 파이프라인을 축소 및 통합하여, 트랜스포머가 멀티모달 맥락을 직접 통합하고 웨이포인트와 행동을 예측하도록 설계된다. 대표적으로 TransFuser는 카메라-라이다 표현을 다중 해상도 자기어텐션으로 융합하여 복잡 교차로·밀집 교통 시나리오에서 충돌을 크게 줄이는 결과를 보고하며, E2E 주행에서 트랜스포머 융합의 이점을 정량화하였다 [7]. 이후 제안된 계층적 도입은 다중 행위자 상호작용이 강한 도시 주행에서 특히 효과적이다. HiVT는 로컬 문맥 추출과 전역 상호작용 모델링을 분리한 계층적 벡터 트랜스포머로, 실시간 예측을 염두에 둔 계산 효율을 유지하면서 다중 행위자 제적 예측 성능을 향상시킬 수 있음을 증명하였다 [8]. 이와 달리, MTR은 소수의 학습 가능한 의도 쿼리로 전역 의도를 추정하고, 지역적 제적을 반복 정련하는 두 단계 구조를 도입하여 멀티모달 제적 예측의 안정성과 정확도를 높이는 방법을 제시한다 [9]. 이러한 분야에서의 활발한 연구는 계층적 트랜스포머 개념이 실제 벤치마크와 리더보드에서도 효율, 안정성, 그리고 일반화 사이의 균형을 확보하는 유효한 방향임을 시사한다.

IV. 결론 및 향후 연구 방향

본 논문은 트랜스포머 기반 시퀀스 모델링이 자율주행의 경로 계획과 제어에서 전통적 파이프라인의 경계를 축소하고, 멀티모달 융합과 계층적 분해를 통해 전역 추론, 실시간성, 안전성을 동시에 추구하는 방향으로 수렴하고 있음을 정리하였다. 특히 상위와 하위를 분리하는 계층적 구조는 유효 시퀀스 길이를 단축하고 재계획 빈도를 높여 급격한 장면 전환에서도 안정적 거동을 보임을 확인하였다.

향후 연구 방향 중 하나는 계층 간 인터페이스의 표준화이다. 상하위 모듈 사이의 주고받는 정보를 명확히 정의하고, 인터페이스의 차원, 주기, 제약 등을 체계화하면 모듈 교체와 도메인 전이가 용이해지며, 이는 곧 일반화 성능의 향상 및 범용성 증대로 이어질 수 있다. 또한 대규모 사전학습과 지시 기반(planning-oriented) 파인튜닝을 결합한 범용 정책으로의 확장, 모델 기반 플래너(MPC 등)와의 하이브리드 결합, 다중 에이전트 상호작용을 고려한 공동 계획으로의 일반화가 요구된다. 이러한 연구가 축적된다면 계층적 트랜스포머는 다양한 센서 구성과 차량 플랫폼에 이식 가능한 실증적인 방안으로 자리 잡을 것이며, 실도로 조건에서 요구되는 안전성과 실시간성을 동시에 만족하는 경로 계획 및 제어의 표준 기술로 발전할 것이다.