

GPU 가속을 활용한 ROS2 기반 실시간 인지 시스템

김진완, 황정훈, 박민철*
한국전자기술연구원

jwk1939@keti.re.kr, hwangjh@keti.re.kr, *mincheol.p@keti.re.kr

ROS 2-based Real Time Cognitive System Using GPU-Acceleration

Kim Jin Wan, Jung Hoon Hwang, Park Min Cheol*
Korea Electronics Technology Institute.

요약

본 논문은 라이다-카메라 기반 인지 시스템을 실시간으로 적용할 때 발생하는 데이터 병목 현상을 분석하고, 이를 해결하기 위한 GPU 기반 데이터 처리 시스템을 제안한다. 기존 ROS2 환경은 CPU 기반으로 데이터를 처리하는 노드 구조로 구성되어 있어, 대용량 센서 데이터를 처리하고 인지 모델을 추론하는 과정에서 CPU-GPU 간 데이터 전송 병목 및 지연이 발생하였다. 이러한 문제를 해결하기 위하여 NVIDIA Isaac ROS 의 NITROS 를 활용하여 센서 데이터 수집부터 전처리, 추론에 이르는 전체 파이프라인을 GPU 가속 기반으로 구성하였다. 이를 통해 불필요한 메모리 복사와 데이터 전송을 최소화하고, 연속적인 센서 입력 환경에서도 안정적인 추론 수행이 가능하도록 하였다. 실험 결과, 제안한 시스템은 기존 CPU 기반 처리 방식 대비 연속 데이터 수집 상황에서도 frame drop 현상을 감소시키고, 안정적인 인지 성능을 유지함을 확인하였다.

I. 서론

최근 보행 로봇 및 자율 주행 로봇의 활용이 확대됨에 따라, 라이다(LiDAR)와 카메라와 같은 각종 센서를 활용한 실시간 인지 시스템의 중요성이 증가하고 있다. 이러한 인지 시스템은 주변 환경을 정확하게 인식하기 위해 대용량의 센서 데이터를 높은 주기로 처리해야 하며, 실시간성과 안정성이 동시에 요구된다. 특히 이동 로봇 환경에서는 인지 지연이나 frame 손실이 누적될 경우 시스템 전체의 안정성 저하로 이어질 수 있다.

ROS2[1]는 로봇 소프트웨어 개발을 위한 대표적인 미들웨어로, 분산 시스템 환경에서의 확장성과 실시간 통신을 지원한다. 그러나 기존 ROS2 기반 인지 시스템은 주로 CPU 중심의 노드(Node) 구조로 구성되어 있으며, 센서 데이터 처리와 딥러닝 기반 인지 모델 추론 과정에서 CPU 와 GPU 간 반복적인 데이터 복사가 발생한다. 이로 인해 카메라 이미지와 라이다 데이터를 동시에 처리하는 상황에서는 데이터 병목 현상과 추론 지연이 발생할 수 있다.

최근 이러한 문제를 해결하기 위해 GPU 가속을 활용한 인지 파이프라인에 대한 연구가 활발히 진행되고 있다. 특히 NVIDIA Isaac ROS[2]에서 제공하는 NITROS 는 ROS2 환경에서 GPU 메모리 기반의 데이터 전달을 지원함으로써, CPU-GPU 간 불필요한 데이터 전송을 최소화할 수 있는 구조를 제공한다. 이를 활용하면 센서 데이터 수집부터 전처리, 추론에 이르는 전 과정을 GPU 기반으로 구성할 수 있어 실시간 인지 성능 향상이 기대된다.

본 논문에서는 라이다-카메라 기반 인지 시스템을 대상으로, 기존 CPU 중심 ROS2 처리 구조에서 발생하는 데이터 병목 현상을 분석하고, NITROS 를 활용한 GPU 가속 데이터 처리 파이프라인을 제안한다. 제안한 시스템은 연속적인 센서 입력 상황에서의 추론 지연 및 frame drop 발생 양상을 비교 분석함으로써, GPU 가속 기반 접근 방식이 인지 시스템의 실시간성과 안정성 향상에 효과적임을 검증한다.

II. 본론

2.1 시스템 전체 구성

본 연구에서 제안하는 시스템의 환경은 Livox-MID360 라이다 및 E-con 의 RouteCAM CU25 카메라를 기반으로 하며, RTX-4080 과 ROS2 환경에서 실시간으로 동작하도록 구성하였다. 시스템은 크게 데이터 전처리 그리고 Depth Completion 모델인 DMD3C[3]의 추론 두 단계로 구성되며, 각 단계는 GPU 가속을 활용할 수 있도록 설계되었다. 기존 ROS2 인지 시스템에서는 센서 드라이버 노드와 전처리 노드가 CPU 상에서 동작하고, 추론 단계에서만 GPU 를 사용하는 구조가 일반적이다. 이러한 구조에서는 센서 데이터가 CPU 메모리에서 GPU 메모리로 반복적으로 복사되며, 고주파 데이터 입력 시 병목 현상이 발생한다. 이를 해결하기 위해 본 연구에서는 NVIDIA Isaac ROS 의 NITROS 를

활용하여, 센서 데이터가 GPU 메모리 상에서 직접 전달되는 파이프라인을 구성하였다.

2.2 GPU 가속 기반 데이터 전처리

라이다와 카메라 센서로부터 입력되는 원시 데이터는 추론 모델에 바로 입력되기 어려우므로 해상도 조정 및 필터링과 같은 전처리 과정이 필요하다. 본 시스템에서는 이러한 전처리 과정을 CPU 기반 노드에서 수행하는 대신 GPU 가속 노드로 이전하였다. 카메라 이미지의 경우, NITROS에서 제공하는 Resize 를 수행하고 라이다 데이터 또한 GPU 메모리 상에서 필터링 및 좌표 변환을 수행하도록 구성하여, CPU 개입을 최소화하였다. 이를 통해 전처리 과정에서 발생하던 데이터 복사 지연을 줄이고, 센서 입력 주기에 맞춰 안정적인 데이터 처리가 가능하도록 하였다.

2.3 딥러닝 기반 추론 파이프라인

전처리가 완료된 센서 데이터는 GPU 메모리 상에서 딥러닝 기반 인지 모델로 전달되어 추론이 수행된다. 본 연구에서는 객체 인식 및 환경 인지를 수행하는 DMD3C 를 적용하였으며, 모델 추론 또한 ROS2 노드 형태로 구성하여 시스템 통합성을 유지하였다. NITROS 를 활용함으로써 전처리 노드와 추론 노드 간의 데이터 전달이 GPU 메모리 상에서 직접 이루어졌으며, 이로 인해 CPU- GPU 간 데이터 이동으로 인한 지연을 효과적으로 줄일 수 있었다. 또한 연속적인 센서 입력 상황에서도 보다 안정적인 인지 결과를 제공할 수 있었다.

of the Computer Vision and Pattern Recognition Conference. 2025.

III. 결론

제안한 시스템은 라이다 및 카메라 데이터를 연속적으로 수집하는 환경에서, 기존 CPU 기반 처리 구조와 GPU 가속 기반 구조를 비교하였으며, 실험 결과, GPU 가속 기반 시스템은 연속적인 데이터 수집 환경에서도 frame drop 발생을 최소화하며 안정적인 추론 성능을 유지하였으며, 이는 실제 로봇 운용 환경에서 시스템 안정성 향상에 기여함을 기대할 수 있다.

ACKNOWLEDGMENT

본 연구는 산림청의 고성능 목재수확 기계장비 기술개발사업(No. 00403540, 고성능 목재수확 기계장비용 다목적툴 적용이 가능한 경량 유압 매니퓰레이터, 운반체 설계 및 통합 제어 시스템 기술 개발)의 지원을 받았습니다.

참 고 문 헌

- [1] Macenski, Steven, et al. "Robot operating system 2: Design, architecture, and uses in the wild." *Science robotics* 7.66 (2022): eabm6074.
- [2] <https://developer.nvidia.com/isaac/ros>.
- [3] Liang, Yingping, et al. "Distilling monocular foundation model for fine-grained depth completion." *Proceedings*