

산업용 로봇 조작 데이터 구축 도구의 성능 보장을 위한 설계

손영성, 한효녕, 박병훈, 이안영, 박준희*
한국전자통신연구원

{ysson, hyonyoung.han, juni}@etri.re.kr, {bhpark, aylee}@s-n.co.kr

Design of a Performance-Assured Framework for Industrial Robot Manipulation Data Acquisition

Youngsung Son, Byeounghun Park, An-Young Lee, Hyonyoung Han, Junhee Park*

Electronics and Telecommunications Research Institute

요 약

본 논문은 고정밀 조작(high-precision manipulation)을 수행하는 산업용 로봇의 학습을 지원하기 위해 고품질 조작 데이터셋을 효율적으로 구축할 수 있는 데이터 수집·저장 프레임워크의 설계 및 구현을 제안합니다. 본 연구에서 제안하는 시스템은 착용형 광학 측각 센서 기반 데이터 수집 시스템을 확장하여, 시각 이미지, 로봇 팔 자세, 다중 센서 스트림 등 다중 모달(multi-modal) 데이터를 통합적으로 획득하고 이를 표준화된 시계열 데이터 포맷으로 저장합니다. 특히, 데이터 수집 과정 전반에서 실시간 성능(throughput)과 데이터 일관성(consistency)을 동시에 확보하기 위한 주석 도구(annotation tool) 및 저장 관리 모듈의 설계를 진행했습니다. 또한, 제안 시스템의 아키텍처 구성, 데이터 저장 방식, 성능 검증 절차를 구체적으로 제시합니다. 성능 평가 결과, 본 시스템은 고속 데이터 스트림 환경에서도 안정적인 동작을 보였으며, 고충실도(high-fidelity) 조작 데이터를 안정적으로 기록할 수 있습니다. 이를 통해 본 연구는 고급 로봇 제어 모델의 학습에 필수적인 산업용 조작 데이터셋의 체계적 구축 기반을 제공합니다.

I. 서 론

산업 현장에서는 정밀 조작과 섬세한 상호작용이 가능한 고성능 로봇 시스템에 대한 수요가 빠르게 증가하고 있다. 이러한 로봇의 성능을 향상시키기 위해서는 인간의 조작 특성과 환경 변화를 정밀하게 반영한 고품질 조작 데이터셋이 필수적이다. 특히 안정적 파지와 적응적 힘 제어를 위해서는 시각, 촉각, 고유감각 등 다중 모달(multimodal) 센서 데이터의 통합 학습이 필요하다. 그러나 실제 환경에서 발생하는 비동기적 데이터를 정확히 수집·동기화·저장하는 것은 여전히 어려운 과제이다.

이에 본 연구는 산업용 로봇 조작 데이터셋 구축과 관리 효율을 극대화하기 위한 프레임워크를 제안한다. 착용형 광학 측각 센서 기반의 고정밀 데이터 수집 기술을 확장하고, 비동기 다중 모달 데이터 스트림의 시간·공간 동기화를 통해 실시간 성능과 데이터 일관성을 보장하는 도구 세트를 설계하였다[1]. 본 시스템은 고충실도 데이터셋 기반 로봇 파운데이션 모델 학습에 활용될 수 있으며, 향후 정밀 조작 중심의 산업용 로봇 응용 기술 발전에 기여할 것으로 기대된다.

II. 시스템 설계

제안된 산업용 로봇 조작 데이터셋 구축 시스템 아키텍처는 신뢰성·실시간성·확장성을 갖춘 데이터 수집을 위해 여러 소프트웨어 구성 요소를 통합하였다. 핵심 미들웨어인 ROS2는 로봇 그리퍼의 모터와 카메라

토픽(자세, 관절 각도, 토크, 영상, 촉각, 핑거 센서 등)을 비동기적으로 구독하여 다중 모달(multimodal) 데이터 스트림을 효율적으로 수집하고 시간·공간 동기화를 유지한다.

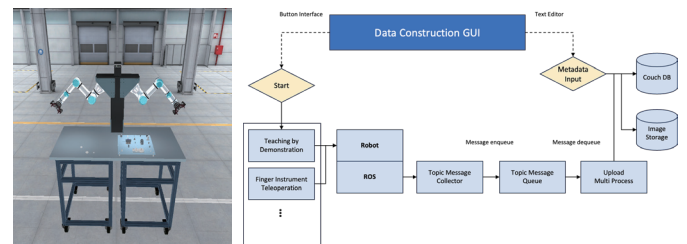


그림 1. 로봇 조작 데이터 수집 도구 구성

저장은 CouchDB 기반의 NoSQL 구조를 사용하여 메타데이터를 관리하고, 대용량 이미지는 별도 분리해 성능을 확보한다. PySide6 GUI는 작업자의 데이터 수집·생성을 지원하며, Pandas는 데이터를 Parquet 포맷으로 변환하여 효율적인 관리와 빠른 입출력을 가능하게 한다.

수집된 데이터는 Task-Episode-Frame의 계층 구조로 관리된다. Task는 조작 동작, Episode는 단일 수행, Frame은 시점별 샘플을 의미한다. 각 프레임에는 로봇 관절 데이터, RGB/깊이 영상 경로, 손가락·촉각 센서 입력이 포함된다. 학습 시 데이터는 Observation-

Action 구조로 변환되며, Observation 은 시점 t 의 이미지와 관절 각도, Action 은 이후 100 스텝의 관절 변화와 패딩 플래그로 구성된다. 이 구조는 시간적 연속성을 반영한 정밀 조작 학습에 적합한 고충실도(high-fidelity) 데이터셋을 제공한다.



그림 2. 로봇 동작 데이터 예시

III. 시스템 성능

본 연구에서 제안된 시스템의 성능 평가는 통신 속도(communication speed)와 저장 성능(storage performance)의 두 가지 측면에 중점을 두었다. 실시간 데이터 수집을 보장하기 위해 목표 통신 처리량(throughput)은 최소 80 MB/s 로 설정하였다. 실험 결과, 제안된 시스템은 500,000 개의 메시지 포인트(message points)를 6.04 초 내에 성공적으로 전송하였으며, 1 Gbps 네트워크 환경에서 36,000 개의 이미지 데이터 포인트를 82 초 내에 전송하였다. 이를 통해 제안된 아키텍처가 대용량 데이터 스트림 환경에서도 안정적으로 동작함을 확인하였다.

표 1. 실시간 에피소드 저장 성능

단위	성능 (속도 or 용량)
1 에피소드	5000(data) * 100 sec = 500K (data) / 167MB
Read/Write	100 에피소드 481 초/2043 초
데이터셋 크기	10 분 64GB
전체 용량	3 달 20 일 작업 데이터 로봇 조인트 2.1TB, 이미지 60TB

데이터셋의 저장 및 검색 효율을 향상시키기 위해, 수집된 데이터는 Parquet 포맷으로 변환되었으며 Zstandard(Zstd) 알고리즘을 이용해 압축하였다. 500 MB 규모의 원본 데이터(raw data)를 대상으로 한 압축 실험 결과, Zstd 압축 레벨 3 에서 데이터 크기가 295 MB 로 감소하였으며, 에피소드(episode)당 평균 처리 시간은 3.19 초로 측정되었다. 이 결과는 압축률과 처리 속도 간의 우수한 균형(trade-off)을 보여주며, 대규모 데이터셋의 효율적인 저장 관리에 적합한 것으로 나타났다.

데이터 수집 인터페이스는 Button Interface 기반으로 설계되었으며, 주요 기능으로는 Start, Teaching by Demonstration, Finger Instrument Teleoperation, Data Construction GUI 등이 포함된다. 시스템 구성 요소로는 ROS 기반 메시지 큐(Message Queue) 구조가 채택되어 있으며, Message enqueue/dequeue 모듈, Topic Message Collector, Metadata Input, Upload Multi-Process 모듈, CouchDB 기반 메타데이터 저장소, 그리고 이미지 전용 저장 영역(Image Storage)으로

구성된다. 이를 통해 다중 프로세스 환경에서도 안정적인 데이터 전송과 병렬 업로드가 가능하도록 구현하였다.

장기 운용을 위한 저장 용량(capacity) 계획 결과, 3 개월간 약 9,000 개의 에피소드 데이터를 수집할 경우 총 약 70 TB 의 저장 공간이 필요할 것으로 추정되었다. 이 중 카메라 영상 데이터가 약 60 TB, 로봇 메타데이터가 약 2.1 TB 를 차지하였다. 이와 같은 성능 분석 결과, 제안된 시스템은 고속·고용량 데이터 환경에서도 실시간성, 확장성, 안정성을 모두 충족함을 입증하였다.

IV. 결론

본 연구에서는 고정밀 로봇 제어를 위한 착용형 광학 촉각 그리퍼(EVO wearable optical tactile gripper) 를 기반으로, 산업용 로봇 조작 데이터셋을 효율적으로 구축하기 위한 통합 도구 세트(toolset) 의 설계와 검증 결과를 제시하였다. 제안된 프레임워크는 ROS2 를 활용하여 비동기적 다중 모달(multimodal) 데이터를 안정적으로 수집하고, CouchDB 기반의 NoSQL 저장 구조를 통해 대용량 데이터를 효율적으로 관리한다. 또한 PySide6 를 이용하여 직관적이고 사용자 친화적인 인터페이스를 구현하였으며, Pandas/Parquet 를 적용함으로써 정형화된 데이터셋의 구성과 분석을 용이하게 하였다.

본 시스템은 시각, 촉각, 고유감각 등 다양한 센서로부터 획득한 데이터를 시간적·공간적으로 동기화된 형태로 체계적으로 수집·주석화(annotate)할 수 있으며, 이를 통해 정밀 조작 학습에 최적화된 고충실도 (high-fidelity) 데이터셋을 생성할 수 있다. 검증 결과, 제안된 프레임워크는 실시간 성능과 데이터 일관성을 동시에 보장하면서, 고정밀 조작 작업을 위한 로봇 학습 데이터의 대규모 구축에 효과적으로 기여함을 확인하였다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부와 한국산업기술진흥원의 "국제공동기술개발 사업" 의 지원을 받아 수행된 연구 결과임(연구개발과제번호 P0028420)

참 고 문 헌

- [1] Chang-beom Kim, etc, "A Wearable Optical Tactile Data Collection System for Human-Robot Teaching of Precision," International Conference on Control, Automation, and Systems, 2025
- [2] Alexander Khazatsky, etc, "DROID: A Large-Scale In-the-Wild Robot Manipulation Dataset," Robotics: Science and Sytems, 2025.
- [3] Vaibhav Saxena, etc, "What Matters in Learning from Large-Scale Datasets for Robot Manipulation," International Conference on Learning Representations, 2025.
- [4] Yanlei Ye, etc, "ROS2 Real-time Performance Optimization and Evaluation," Chinese Journal of Mechanical Engineering. 36, Article number: 144, 2