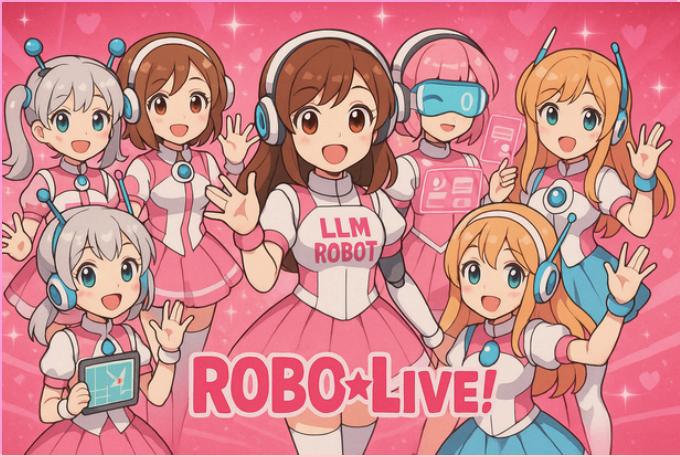


2025 한국통신학회 하계학술대회

LLM 기반 적응형 의사결정 로봇 시스템

참가 팀 소속 : 국민대학교 창의공과대학 전자공학부 (담당교수 : 주민철)

참가 팀 이름 : 러브라이브 (이지석, 박효준, 이우석, 박창균, 김예찬, 임현우, 정서현)



1. System Overview

Components of a system

사용자 중심 제어 진입 장벽

직관적 인터페이스
자연어 → ROS2 명령 자동 변환 GUI

험난 지형 맵핑 한계

자율 탐사-맵핑
Frontier-Based Exploration + SLAM

정밀 객체 파지 오차

정밀 객체 인식
RealSense D435i + YOLOv8

원격 제어 불안정

실시간 통신
5G 특화망으로 10ms 미만 지연 실현

시스템 확장성 요구

경량화-모듈화-데이터셋 구축
QLoRA 4bit 양자화로 GPU 부담 경감 및 빠른 파인튜닝
3,000개 자연어-ROS 명령어 쌍 데이터셋 구축으로 확장성 확보

Real-time Depth Sensing and Object Grasping

Depth Camera-Based Object Picking

로봇이 Depth 카메라를 활용해 물체의 3차원 위치를 인식하고, 해당 좌표를 바탕으로 로봇 팔을 제어해 물체를 파지하는 자동화 기법.

기술적 특징

- Depth 카메라를 통해 물체의 3D 위치를 추정하고, 로봇 기준 좌표계로 변환.
- 로봇 팔의 역기구학 계산을 통해 정확한 위치로 그리퍼 이동 및 파지 수행.
- 다양한 물체 형태 및 조명 조건에서도 동작 가능한 유연한 인식-제어 구조.

Structure

5G Private Network

- 로봇 간 통신 채널 제공
- 실시간 원격 제어 가능

Burger(탐지 로봇)

객체 탐지 및 좌표 송신

Computer

- LLM 기반 의사결정 및 명령 처리
- Turtlebot 제어 인터페이스 및 GUI 제공
- 지도 시각화 및 데이터 분석

Waffle(작업 로봇)

객체 파지 및 좌표 수신

2. Design

LLM-Based User-Centered Control

LLM 연동 GUI 기반 로봇 제어 시스템

직관적인 사용자 GUI 설계로 ROS 비전문가도 손쉽게 로봇 제어 가능
로봇팔 제어 및 맵핑 기능을 자연어 기반 챗봇 인터페이스로 구현
클릭 또는 명령어만으로 구성된 사용자 중심 UI → 시스템 접근성 향상

Qlora 기반 경량 Fine-Tuning

기존 16/32bits 부동소수점 학습 대비 4bit 양자화 적용
학습 속도 향상 및 GPU 메모리 사용량, 파인튜닝 비용 절감
명령어 추가 시 빠른 재학습 가능 → 유연한 시스템 확장성 확보

로봇 제어 특화 Dataset 구축

자연어 텍스트 명령어 ↔ ROS 실행 명령어로 구성
다양한 제어 시나리오 기반의 명령어-동작쌍 데이터셋을 수집 및 설계
LLM 학습에 최적화된 포맷으로 구성 → 높은 명령 인식을 달성

GUI

명령어
로봇 탐색 실행하고 맵 구축해줘

콘솔
ROS2 + Hugging Face

실행 성공!
***.o(=▽=)o..*☆(づ｡◕｡◕)づ(｡◕｡◕)/♡(づ｡◕｡◕)づ(｡◕｡◕)/♡(≥▽≤)♡

남- 실행!

Scenario

비타500이 어딴데?

모델에 명령을 전달할 Local AI 생성

명령 생성 및 전달

LLM을 통해 로봇에 명령 전달

5G 특화망을 통한 기기와 로봇 통신

MAP 생성 및 객체 탐지

Map 구축 완료

Frontier-Based Exploration 수행

서버에 Map 저장

LLM을 통해 로봇에 명령 전달

명령 키워드 추출 → 이동 시작

객체 파지 및 복귀

목표 위치 도착 → 목표물 탐지

IK 궤적 생성 → 목표물 집기

좋아, 복귀다!

고마워!!

Process Pipeline for Vision-Based Robotic Arm Operation

pick_and_place_input 노드

MoveGroupInterface(arm)

bottle_point 구독자

point_callback

tf2_ros:Buffer

카메라를 통한 객체 탐지

회전 탐색 시작

setValueTarget(joint1=angle)

move()

spin_some() + sleep(0.001s)

/bottle_point 메시지 수신

break

loop (오류로 종료 시 break)

setValueTarget(joint1=angle)

move()

spin_some() + sleep(0.001s)

/bottle_point 메시지 수신

break

break (오류로 종료 시 break)

객체 인식된 후 회전 중단

카메라에서 측정된 좌표를 로봇팔의 좌표로 변환

transform(target_point, "link5")

link5_frame_point or None

transform(link5_frame_point, "base_link")

base_point or None

로봇팔 경로 계획 & 이동

setPositionTarget(x,y,z)

setGoalPositionTolerance(0.01)

setPlanningTime(10.0)

setMaxVelocityScalingFactor(0.5)

setMaxAccelerationScalingFactor(0.5)

plan(plan)

SUCCESS

break (계획 실행)

execute(plan)

sleep_for(2s)

shutdown()

pick_and_place_input 노드

MoveGroupInterface(arm)

bottle_point 구독자

point_callback

tf2_ros:Buffer

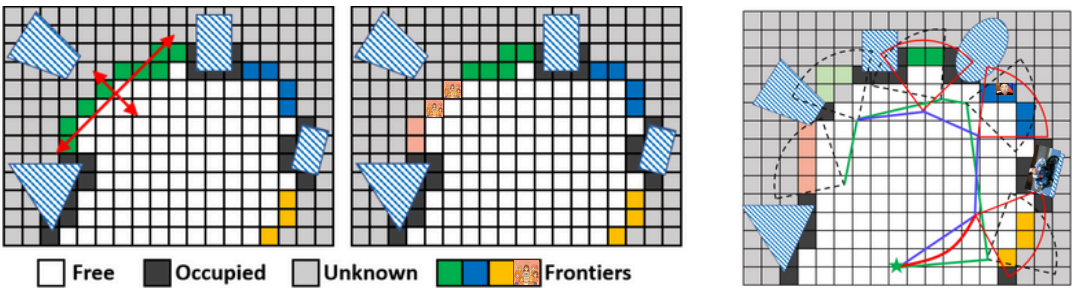
Autonomous Exploration

Frontier-Based Exploration

- 자율 로봇이 탐색된 영역과 미지 영역의 경계(frontier)를 따라가며 환경을 확장해가는 실시간 지도 구축 기법
- SLAM과 결합하여 실시간 지도 작성 및 미지 영역 탐색이 동시에 이루어짐

Frontier 탐지 및 클러스터링

- BFS 기반 경계점 식별: 점유 격자에서 자유 공간-미지 영역 경계를 탐지하고, 8방향 인접 셀 스캔으로 프론티어 후보 선정.
- DFS 기반 그룹화: 연결된 프론티어 포인트를 클러스터링 탐사 효율을 위해 5개 주요 그룹만 선택.



동적 목표 선택 전략

- 가중치 기반 최적화: 거리(지수 감소)와 그룹 크기 비율로 가중치를 계산함.

$$Score = \frac{frontier.group.size}{e^{distance}}$$

실시간 경로 계획

- A* 알고리즘: 8방향 이동 허용하고 유클리드 거리 휴리스틱 사용. 장애물 확장으로 안전 마진 확보.

Robotic Arm

Inverse Kinematics 기반 로봇팔 작동

-카메라로부터 수신한 목표좌표로 그리퍼의 손기구학 역함수를 이용해 대응하는 관절각을 구하고 다음식을 따름. $\theta^* = [\theta_1, \theta_2, \theta_3, \theta_4]^T$ 를 계산하여 목표위치에 도달.

뉴턴-랩스 반복법

- 로봇팔의 역기구학(IK) 문제는 비선형 방정식으로 구성되므로 해석적 접근이 어려워 수치 해석 기법을 이용. 목표 위치에 도달하기 위해 자코비안 행렬 기반의 뉴턴-랩스 반복법을 적용했고 반복식 $\theta_{k+1} = \theta_k + J^+(J_k)(x_d - x(\theta_k))$ 을 통해 관절 각도를 점진적으로 갱신.

Object Detection

- 로봇 팔이 원하는 물체를 파지할 수 있도록 카메라를 통한 Object Detection 구현.
- 실시간 객체 탐지에 최적화된 YOLOv8 모델을 적용하여 딥러닝 기반 물체 인식을 구현.

Coordinates Transformation

- Inverse Projection Transformation을 통해 2차원 이미지 좌표계에서 3차원 카메라 좌표계로 변환.
- 로봇이 물체를 파지할 수 있도록 카메라 좌표계에서 로봇 좌표계로 변환

Depth

- 물체의 3차원 좌표를 얻기 위해서 물체의 깊이 정보가 필요.
- RGB와 깊이 영상을 동시에 제공하는 Intel RealSense D435i 카메라를 활용해 체까지의 깊이를 추정.

Formula

-이미지의 픽셀 좌표에서 Inverse Projection Transformation을 통해 카메라 좌표계로 변환. 2차원 이미지로 투영되면서 차원을 잃게 되므로 Stereo Camera를 이용해 각 픽셀의 깊이 정보를 추정.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = Z \cdot K^{-1} \cdot \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix}$$

X_c, Y_c, Z_c

카메라 기준 좌표

Z

픽셀의 깊이 정보

 K^{-1} x_p, y_p

ROS2에서 제공하는 TF2 기능을 이용하여 로봇 좌표계와 카메라 좌표계를 연결. Transformation을 통해 로봇 좌표계를 기준으로 변환.

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

X_w, Y_w, Z_w

로봇 기준 좌표

R

Rotation Matrix

 T X_c, Y_c, Z_c

$\theta_1 = \arctan2(y, x)$

Joint1

회전각

 $\theta_2 = \arctan2(s, r) - \arctan2(a_3 \sin \theta_3, a_2 + a_3 \cos \theta_3)$

Joint2

 $\theta_3 = \arctan2(\pm\sqrt{1 - D^2}, D)$

Joint3

 $\theta_4 = \phi - (\theta_2 + \theta_3)$

Joint4

 $r = \sqrt{p_{w,x}^2 + p_{w,y}^2}$ $s = p_{w,z} - d_1$

평면거리, 높이 offset

 $D = \frac{r^2 + s^2 - a_2^2 - a_3^2}{2a_2a_3}$

엘보각도를 위한 중간변수

3. Performance

Mapping

탐색 시간

(1) MAP: 110.12 sec
(2) MAP: 83.64 sec

방문 위치

이동 경로

선택된 탐사 목표점

Occupied

Frontier Point 후보

Autonomous Exploration을 통해 생성한 Map과 로봇 이동 경로

Object Detection

카메라 좌표계 기준 3차원 좌표 추정

로봇 좌표계 기준 3차원 좌표 추정

카메라를 통한 객체 인식 및 Object Picking

4. Application

LLM-Based Decision-Making Robot

