

자율 로봇 제조를 위한 계층적 작업 인식 프레임워크

A Hierarchical Task Recognition Framework for Autonomous Robotic Manufacturing

강민석¹·박상현²·김무림[†]

Minseok Kang¹, Sanghyun Park², Murim Kim[†]

Abstract: This study proposes a hierarchical task recognition framework that enables robots to precisely perceive and interpret their actions in complex and unstructured manufacturing environments. The system decomposes entire processes into modular primitive tasks and primitive actions, leveraging multimodal sensor data to evaluate task states in real time. By structuring each task stage as an independent, reusable module, the framework supports flexible composition and adaptation across diverse process scenarios. It further incorporates both success and failure conditions to ensure robust control under unexpected circumstances. Crucially, the system allows robots to dynamically adjust their execution based on real-time feedback, enhancing autonomy and responsiveness. The framework has been validated in real-world assembly tasks, demonstrating its ability to adapt to various task types and conditions without structural modification. This approach offers a practical foundation for intelligent robotic systems capable of real-time understanding and adaptive control in dynamic manufacturing settings.

Keywords: Autonomous Manufacturing Process, Robot Task Recognition, Multimodal Sensor Integration, Primitive Task Recognition, Hierarchical Task Recognition

1. 서론

최근 로봇이 인간과 유사한 방식으로 학습을 통해 지능적으로 자율 작업을 수행할 수 있도록 하는 연구가 활발히 진행되고 있다. 특히 Stanford 대학의 Mobile ALOHA^[1]와 같은 시스템은 사람의 시연과 조작 데이터를 기반으로 로봇이 다양한 작업을 학습하고 수행할 수 있도록 설계되어, 기존의 정형화된 로봇 제어 방식에서 벗어나 보다 유연하고 적응적인 작업 수행이 가능함을 보여주고 있다. 이러한 자율 작업 기술은 초기에는 주로

가정이나 서비스 환경에서의 단순하고 반복적인 작업에 적용되었다. 예를 들어, 간단한 물건을 집거나 정리하고, 음료를 서빙하는 등의 일상적인 작업은 비교적 구조화된 환경에서 안정적으로 수행될 수 있었으며, 이를 통해 자율 로봇 기술의 실용성과 가능성이 검증되었다.

그러나 기술의 발전과 함께 로봇에게 더 높은 수준의 정밀성, 상황 적응력, 복잡한 의사결정 능력이 요구되면서, 최근에는 제조업과 같은 복합적이고 비정형적인 산업 환경으로도 자율 작업 기술의 적용이 확대되고 있다. 제조업은 다양한 부품과 공정이 유기적으로 연결된 복잡한 환경으로, 작은 오차나 예기치 못한 변수도 전체 공정의 성공 여부에 중대한 영향을 미칠 수 있다. 특히 작업 실패는 제품 불량이나 공정 지연으로 직결되어 생산성 저하로 이어지기 때문에, 고정밀의 상황 인식 능력이 필수적이다.

이러한 환경에서 로봇이 효과적으로 자율 작업을 수행하기 위해서는, 정의된 명령어나 반복적인 제어 방식에 의존하는 것은 한계가 있으며, 작업 상태를 실시간으로 인식하고 변화하는 상황에 능동적으로 대응하며, 작업 전략을 유연하게 조절할 수

Received : May. 29. 2025; Revised : Jul. 15. 2025; Accepted : Jul. 17. 2025

※ This work was supported by the Technology Innovation Program (No. 00420520, Technology convergence of digital twin, artificial intelligence and robotics for AI-Robot based autonomous manufacturing in unstructured) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea).

1. Researcher, Human-Centered Robotics R&D Division, KIRO, Pohang, Korea (cwyh1118@kiro.re.kr)

2. Principal Researcher, Human-Centered Robotics R&D Division, KIRO, Pohang, Korea (shpark@kiro.re.kr)

† Chief Researcher, Corresponding author: Human-Centered Robotics R&D Division, KIRO, Pohang, Korea (mulimkim@kiro.re.kr)

있는 고차원의 자율성이 요구된다. 이를 위해 로봇은 시각, 촉각, 힘 등 다양한 센서를 통해 수집된 데이터를 통합 분석하고, 작업 상황을 정밀하게 인식하고 판단할 수 있어야 한다.

이러한 배경 속에서, 멀티모달 센서 데이터를 활용한 작업 상태 모니터링 및 이상 감지 기술은 복잡하고 비정형적인 제조 환경에서 로봇이 안정적이고 유연하게 자율 작업을 수행할 수 있도록 지원하는 핵심 요소로 주목받고 있으며, 이를 실현하기 위한 연구가 진행 중이다.

예를 들어, Park et al.^[2]은 PR2 로봇을 활용하여 전자레인지 문 닫기, 버튼 누르기, 사람에게 손가락으로 음식을 전달하는 등의 작업에서 발생할 수 있는 이상 상황을 감지하기 위해, 힘/토크, 오디오, 시각 정보를 포함한 멀티모달 센서 데이터를 수집하고 이를 기반으로 정상 실행의 패턴을 학습하였다. 이들은 Hidden Markov Model (HMM)을 활용하여 정상 실행 시의 로그 가능도를 시간에 따라 계산하고, 이를 기반으로 이상 여부를 판단하는 방법을 제안하였다. 학습에는 오직 정상 실행 데이터만을 사용하였으며, 실행 도중 로그 가능도가 시점별 임계값 이하로 떨어지는 경우를 이상으로 간주하였다. Thoduka et al.^[3]는 Toyota HSR 로봇이 책을 선반 위에 놓는 작업을 대상으로 시각적 이상 탐지 기법을 제안하였다. 이들은 optical flow를 통해 관찰된 움직임 분석하고, probabilistic U-Net 모델을 활용하여 정상 실행 중의 시각적 모션을 예측하도록 학습시킨 후, 실제 관찰된 움직임과의 오차를 바탕으로 이상을 감지하였다. 또한 로봇의 관절 상태로부터 계산한 카메라 및 본체의 예상 움직임과의 비교를 결합하여 이상 점수를 정량화 했다. 이 연구 역시 실패 데이터를 사용하지 않고, 오직 정상 실행 데이터를 기반으로 학습을 수행하였다. Inceoglu et al.^[4]는 Baxter 로봇을 이용하여 push, pick-and-place, pour 등 다양한 조작 작업에서 발생할 수 있는 작업 실패를 감지하기 위해, RGB, 깊이, 오디오 등 다양한 센서 데이터를 통합한 FINO-Net 프레임워크를 제안하였다. 이들은 성공과 실패 실행을 모두 포함하는 FAILURE 데이터셋을 구축하고, CNN 및 convLSTM 기반의 딥러닝 모델을 통해 작업 실행을 success/failure로 분류하도록 학습하였다. 즉, FINO-Net은 사전 정의된 특징 기반의 이상 탐지와 달리, 실패 사례를 명시적으로 학습한 감독 학습 기반 분류 방식으로, 다양한 실패 유형에 대해 높은 분류 성능을 보여주었다.

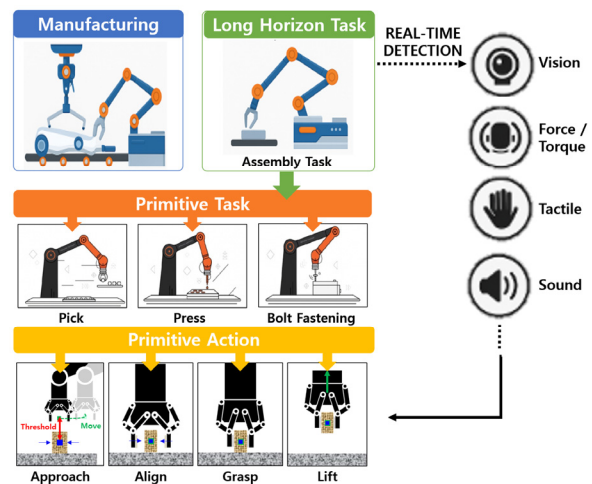
종합하면 기존 연구들은 멀티모달 센서 기반의 작업 상태 인식 및 이상 탐지를 중심으로 이루어졌고 대부분 단일 작업 시나리오나 정형화된 환경에 국한되어 있다. 그 결과, 비정형이거나 복잡한 작업 시퀀스에서는 작업 흐름 전반에 걸친 맥락 변화나 비정형 이벤트 발생 시의 상황 인식 및 대응에 한계가 있다.

최근에는 작업 수행 중의 상태를 실시간으로 인식하고, 이를 바탕으로 작업 전략을 유연하게 조정하려는 연구가 활발히 진행되고 있다. Duarte et al.^[5]는 이벤트 기반 시각 센서를 활용하

여 인간 작업자의 조립 동작을 Pick Up, Place, Screw 등의 primitive 단위로 실시간 분류하는 프레임워크를 제안하였다. 이 연구는 협동로봇이 인간의 현재 상태를 정확히 파악하고, 그에 맞는 협력 행동을 계획할 수 있는 기반 기술로 활용될 수 있음을 시사한다.

본 연구는 이와 유사한 작업 인식 기술을 로봇의 자율 작업 수행에 적용하고자 하며, 적용 대상을 인간이 아닌 로봇 자체의 작업 수행으로 확장한다는 점에서 차별성을 지닌다. 즉, 로봇이 수행 중인 작업을 의미 있는 단계 단위로 실시간 인식하고, 해당 상태 정보를 바탕으로 공정의 진행 상황을 추적하거나 실패 징후를 조기에 감지함으로써, 비정형적이고 동적인 작업 환경에서도 안정적이고 유연한 작업 수행이 가능하도록 한다. 이를 위해 복잡한 제조 공정에서 로봇이 수행하는 복합 작업 (Long Horizon Task)을 보다 정밀하게 인식하고 판단할 수 있도록, 전체 작업을 계층적으로 분해하여 처리하는 모듈 기반 작업 인식 전략을 제안한다. 복합 작업은 여러 개의 단위 작업 (Primitive Task)으로 구성되며, 각 단위 작업은 다시 논리적으로 구분 가능한 최소 단위 동작(Primitive Action)의 조합으로 이루어진다. 각 단위 동작은 로봇의 자세 변화, 조작기 구동 등 물리적으로 명확히 구분 가능한 행위를 의미하며, 작업의 맥락에 따라 로봇 및 조작기의 상태, 시각, 힘/토크, 촉각, 사운드 등 다양한 멀티모달 센서 데이터를 활용해 실시간으로 분석되고 인식된다.

제안하는 인식 프레임워크는 [Fig. 1]에 나타난 계층적구조를 기반으로, 각 단위 동작을 실시간으로 감지하고 이들의 연속성과 전이 관계를 분석함으로써 상위 수준의 단위 작업을 식별하며, 나아가 복합 작업 전체의 흐름과 상태를 추론하는 구조로 설계되었다. 이를 통해 정형화된 명령 기반 제어 방식의 한계를 넘어, 다양한 작업 시나리오와 환경 변화 속에서도 작업 상태를



[Fig. 1] Hierarchical Robot Task Recognition Framework for Complex Manufacturing Processes

정확히 인식하고, 공정 흐름 내에서 발생할 수 있는 이상 상황이나 실패 징후를 조기에 감지할 수 있는 유연성과 확장성을 확보하고자 한다. 본 연구에서는 복잡한 제조 공정을 구성하는 작업을 구조적으로 이해하고, 로봇이 공정 흐름을 실시간으로 해석하며 변화하는 상황에 능동적으로 대응할 수 있도록 지원하는 인식 기반 자율 시스템을 구현하는 방법을 제안한다.

2. 본 론

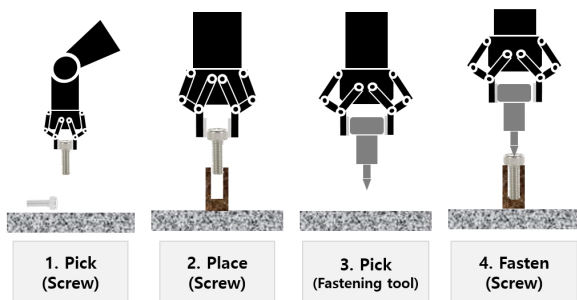
2.1 공정 인식 전략

로봇은 작업 시퀀스를 기반으로 각 단계에 적합한 동작을 수행하지만, 실제 제조 환경에서는 다양한 변수로 인해 의도한 결과가 항상 보장되지 않는다. 특히 복잡한 제조 공정에서는 예기치 못한 오차나 실패가 발생할 가능성이 높으며, 이러한 상황에 적절히 대응하기 위해서는 작업 상태를 실시간으로 인식하고 판단하는 능력이 요구된다. 본 연구에서는 작업 수행 과정에서 발생할 수 있는 이상 징후나 비정상 상황을 인식하고, 실패 또는 비정상적인 수행 결과에 대해 피드백을 제공함으로써, 공정의 안정성과 신뢰성을 높일 수 있는 모듈형 자율 작업 인식 프레임워크를 제안한다.

2.1.1 단위 작업 기반 모듈형 공정 인식 전략

본 연구에서는 단위 작업(Primitive Task)을 제조 공정을 구성하는 가장 기본적인 인식 단위로 정의한다. 단위 작업은 로봇이 실제로 수행하는 연속적인 동작 흐름 속에서 의미적으로 구분되는 최소 수준의 작업 단계이며, 각 단계는 인식의 기준으로 활용된다.

제조 공정은 일반적으로 여러 작업 단계를 순차적으로 포함하고 있어, 각 단위를 명확히 구분하고 인식하는 것이 공정의 신뢰성과 안정성을 확보하는 데 핵심적인 요소로 작용한다. 예를 들어, 스크류 체결 작업은 [Fig. 2]와 같이 '스크류 집기 (Pick)', '스크류 배치하기(Place)', '체결 도구 집기 (Pick Tool)', '체결 도구 집기 (Pick Tool)', '스크류 배치하기(Place)', '체결 도구 집기 (Pick Tool)', '스크류 체결 (Fasten)' 등으로 구성된다.



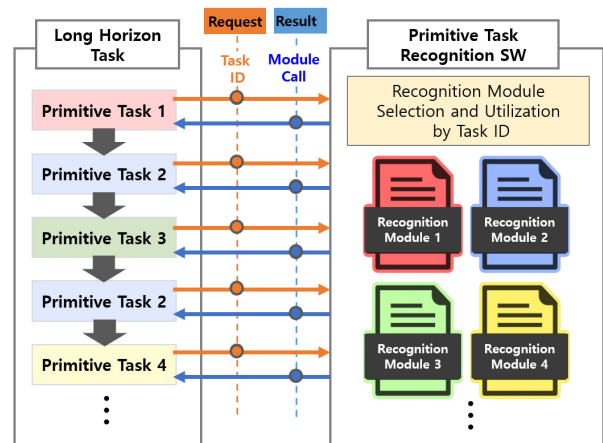
[Fig. 2] Primitive Task Composition of a Screw Fastening Process

'스크류 체결(Fasten)' 과 같은 세부 단계로 구분될 수 있다. 이러한 단위 작업 기반의 공정 분해 및 모듈화는 복잡한 작업 시퀀스를 구조화함으로써, 전체 공정에 대한 인식 난이도를 낮추고 다양한 상황에서도 유연하게 대응할 수 있는 기반을 제공한다. 나아가, 단위 작업의 세분화는 단순히 복잡한 공정을 구조화하는 데 그치지 않고, 다양한 제조 공정에서의 재구성과 확장 가능성이라는 중요한 이점을 제공한다. 즉, 각 단위 작업은 의미적으로 독립된 모듈로 분리될 수 있으며, 이를 조합함으로써 새로운 작업 시나리오나 공정 흐름을 유연하게 구성할 수 있다. 이러한 모듈형 접근은 로봇 시스템이 특정 작업에만 고정되지 않고, 다양한 생산 상황에 대응할 수 있는 범용성과 유연성을 갖추게 한다.

예를 들어, [Fig. 3]와 같이 각 단위 작업에 대응하는 인식 모듈은 독립적으로 설계되어, 공정의 맥락에 따라 동적으로 호출될 수 있으며, 동일한 모듈을 여러 작업 시나리오에서 재사용할 수 있다. 예를 들어, 스크류 체결 과정에서 활용된 'Pick' 인식 모듈은 다른 부품을 집는 작업에서도 동일하게 활용될 수 있다.

보다 정밀한 인식을 위해, 단위 작업은 다시 논리적으로 구성 가능한 최소 동작 단위인 단위 동작 (Primitive Action)으로 세분화된다. 이는 로봇 시스템은 인간과 달리 직관적으로 작업을 수행할 수 없기 때문에, 명확한 판단 기준에 따라 작업 단계를 정의되고, 센서 기반으로 인식 가능한 정량화된 동작 단위로 구성되어야 한다.

본 연구에서는 이러한 목적에 따라 실제 작업자의 조립 절차를 기반으로, 단위 작업을 구성하는 단위 동작 요소들을 분석하였다. 분석 항목에는 작업의 목적, 동작 방식 (예: 접근, 파지, 이송, 놓기), 물체 상태의 변화, 그리고 작업자가 직관적으로 활용하는 감각에 상응하는 센서 데이터가 포함된다. 작업 분석 과정에서는 단순한 작업 순서의 분해를 넘어, 실제 작업 수행 중인



[Fig. 3] Dynamically Invocable and Reusable Primitive Task Recognition Modules

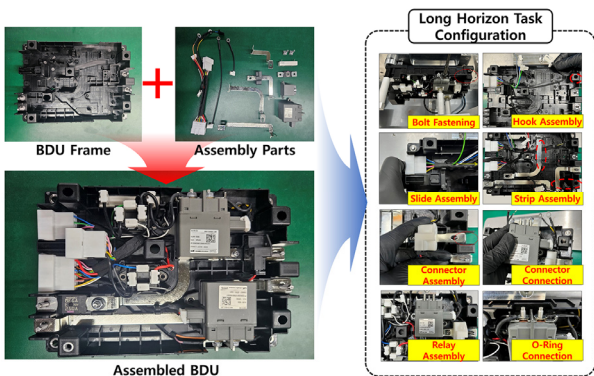
간이 어떻게 감각 정보를 활용하여 판단을 내리는지를 고려하였다. 이를 기반으로 해당 감각 정보를 로봇이 수용 가능한 형태로 정리하고, 센서 데이터 수준에서 통합적으로 해석할 수 있도록 구성하였다.

구체적으로는, 부품의 위치 및 정렬 판단에 필요한 시각 정보, 조립 중 힘의 저항을 인지하는 촉각 및 토크 감각, 그리고 조립 상태를 판단하는 데 유용한 청각 정보 등이 이에 해당한다. 본 연구는 이러한 감각 기반 판단 메커니즘을 로봇의 인식 체계에 반영하였으며, 이에 따라 시각 센서, 힘/토크 센서 등 다양한 센서 정보를 조합하여 단위 작업 수준에서의 상태 변화를 정밀하게 인식할 수 있는 체계를 설계하였다. 이러한 분석을 바탕으로 전체 작업 흐름은 로봇이 수행 및 인식 가능한 단위 작업 수준으로 구조화 및 정량화 되었다.

본 연구에서는 이러한 구조를 기반으로, 복합 제조 공정 중 발생하는 다양한 작업 상황에 따라 적절한 단위 작업 인식 모듈을 선택적으로 호출·운용할 수 있는 공정 인식 기법을 제안한다. 이를 통해 공정 수행 과정에서 각 단위 작업의 성공 여부를 실시간으로 판단하고, 공정 흐름 상의 진행 상태를 정밀하게 파악할 수 있는 공정 인식 체계를 구축하고자 한다.

2.2 조립 공정 선정

제안한 공정 인식 전략의 적용 가능성을 실험적으로 검증하기 위해, 자동차 전장 부품인 BDU(Battery Disconnect Unit) 조립 공정을 테스트 시나리오로 선정하였다. 이 공정은 [Fig. 4]와 같이 릴레이, 커넥터, 기어, 하네스 와이어, 볼트 등 다양한 부품을 포함하며, 이에 대한 파지, 정렬, 삽입, 체결, 위치 보정, 검사 등의 작업이 다단계로 구성되어 순차적으로 수행되는 복합 조립 프로세스이다. 각 작업은 부품의 물성, 형상, 조립 방향, 오차 허용 범위, 센서 피드백 특성이 서로 달라, 정형화되지 않은 다양한 조립 조건과 로봇 동작 유형을 포함하고 있다. 이러한 특성 덕분에 BDU 조립 공정은 단일 제품 공정임에도 불구하고,



[Fig. 4] BDU (Battery Disconnect Unit) Manufacturing Process

제안된 계층형 인식 프레임워크의 유연성, 안정성, 적용 가능 범위를 통합적으로 검증할 수 있는 대표적이고 현실적인 테스트베드로 기능한다.

2.3 대표 단위 작업 도출 및 정의

제안된 공정 인식 전략을 구체화 하기 위해, [Fig. 4]에서 제시한 BDU 조립 공정을 분석하여 핵심 단위 작업을 도출하고, 각 단위 작업에 대한 인식 기준을 정의하였다.

특히, BDU 조립 공정 내에서도 릴레이(Relay) 삽입 작업은 구조가 비교적 단순하면서도 다른 조립 공정에서도 빈번하게 등장하는 핵심 작업 유형을 포함하고 있어, 대표 단위 작업 도출을 위한 분석 대상으로 선정하였다.

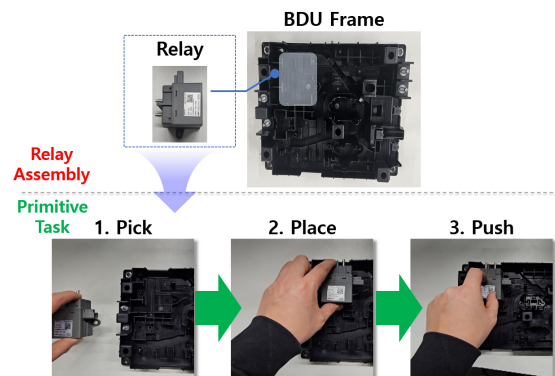
릴레이 조립 작업의 세부 절차를 분석한 결과, [Fig. 5]와 같이 Pick, Place, Push의 세 가지 단위 작업으로 구성됨을 확인할 수 있었다. 이 세 가지 작업은 비교적 기본적인 동작에 해당하지만, 제조 공정 전반에서 보편적으로 활용되는 범용 작업 유형으로, 복합 공정을 세분화하고 구조화하는 데 있어 핵심적인 역할을 수행한다.

공정 전체를 인식 가능하도록 구조화하기 위해 작업 흐름을 단위 작업 수준으로 분해하였으며, 이로 인해 각 단위 작업은 공정 인식 시스템 내에서 인식 모듈의 기본 단위로 기능하게 된다. 단위 작업을 명확하게 인식하고 모듈화하기 위해 각 단위 작업의 개념과 범위를 아래와 같이 정의하였다.

Pick : 위치와 방향이 결정된 상태의 물체를 정확히 잡고 들어 올리는 과정이다.

Place : 위치와 방향이 결정된 위치에 파지한 물체를 이송하여 해당 위치에 놓는 과정이며, 물체를 놓은 후 다음 작업을 위해 팔단을 들어 올리는 과정까지 포함된다.

Push : 물체를 손바닥으로 누르듯이 수직 방향으로 압력을 가해, 물체를 목표 위치에 정확히 끼우거나 고정하는 과정이다. Push 작업에서는 물체를 Grip 하는 동작이 포함되지 않는다.

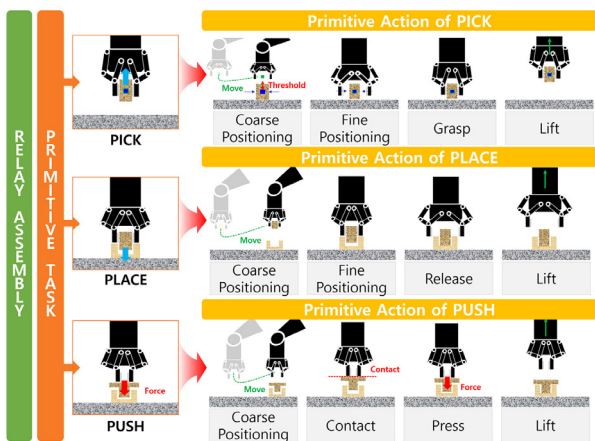


[Fig. 5] Primitive Tasks of the Relay Assembly Process in BDU Manufacturing

2.4 단위 작업 분석

제안된 공정 인식 전략을 기반으로, 릴레이(Relay) 조립 작업에 포함되는 Pick, Place, Push 세 가지 단위 작업에 대해 엄격한 정의를 바탕으로 각 작업의 세부 수행 특성과 요구 조건을 분석하였다. 이를 바탕으로, 각 단위 작업은 논리적으로 구성 가능한 최소 수준의 동작 단위인 단위 동작(Primitive Action)으로 세분화하였다. 세분화된 단위 동작의 조합을 통해 각 단위 작업의 수행 절차를 구조화하고 정형화하여 인식 기준 수립과 인식 모듈 설계의 기반을 마련하였다. 구체적으로, 세 가지 단위 작업은 [Fig. 6]과 같은 단위 동작들로 구성된다.

먼저, Pick 작업은 물체를 파지하여 들어 올리는 절차로, 네 단계의 단위 동작으로 구성된다. 로봇은 먼저 Coarse Positioning 단계를 통해 물체의 대략적인 위치와 방향을 맞추며 접근하고, 이후 Fine Positioning 단계에서 안정적인 파지를 위해 위치와 방향을 정밀하게 조정한다. 그리퍼가 최적의 위치에 도달하면 Grasp 동작을 통해 물체를 파지하며, 마지막으로 Lift 단계에서는 파지한 물체를 일정 임계값만큼 들어 올린다. Place 작업은 파지한 물체를 목표 위치에 정확히 배치하는 과정을 의미하며, 이 역시 네 단계의 단위 동작으로 구성된다. Coarse Positioning 단계에서는 로봇이 물체를 놓을 위치로 접근하면서 대략적인 위치를 맞추고, Fine Positioning 단계에서는 위치와 방향을 정밀하게 조정하여 정확한 배치를 준비한다. 이어서 Grasp 단계에서 그리퍼를 개방하여 물체를 내려놓으며, 마지막 Lift 단계에서는 물체를 놓은 후 다음 작업 단계로 전환하기 위해 로봇 말단을 상승시킨다. Push 작업은 물체를 수직 방향으로 눌러 목표 위치에 고정하는 작업이며, 역시 네 단계로 구성된다. Coarse Positioning 단계에서는 로봇이 접촉 지점 근처로 이동하여 위치를 조정하고, Contact 단계에서 로봇 말단부를 물체와 접촉



[Fig. 6] Primitive Action Composition of Relay Assembly Primitive Tasks

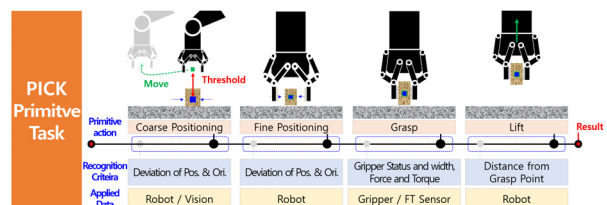
시킨다. 이후 Press 단계에서는 수직 방향으로 힘을 가하여 물체를 소켓 등에 정확히 삽입하거나 고정하며, 마지막 Lift 단계에서는 삽입이 완료된 후 로봇이 상승하여 작업을 마무리하고 다음 단계로의 전환을 준비한다.

2.5 단위 작업 인식 기준

앞 절에서는 전체 작업을 구성하는 기본 요소로서 단위 작업(Primitive Task)을 정의하고, 이를 다시 단위 동작(Primitive Action) 수준으로 세분화하여 작업 절차를 체계적으로 구조화하였다. 이러한 계층적 구조는 공정 흐름의 인식 가능성과 모듈화 가능성을 확보하게 하며, 이는 공정 인식 시스템의 설계를 위한 기반으로 작용한다.

본 연구에서 제안하는 단위 작업 인식 시스템은, 각 작업 단계의 수행 여부를 실시간으로 평가하기 위해 작업 특성에 기반한 센서 기반 정량 판단 기준을 설정하였다. 이러한 인식 기준은 단위 작업을 구성하는 단위 동작(Primitive Action) 수준에서 적용되며, 로봇의 작업 수행 과정에서 수집되는 시각(vision), 힘/토크(force/torque), 사운드(audio), 로봇의 상태 등의 다양한 센서 데이터를 종합적으로 분석하여 수행 결과를 판별한다. 이러한 인식 체계는 실제 제조 공정에서도 적용 가능하며, 릴레이 조립과 같은 다단계 작업에서 각 단위 작업의 성공 여부를 판단할 수 있는 기반을 제공한다. 본 절에서는 그 대표 사례로, 릴레이 조립 공정에서 도출된 Pick, Place, Push 세 가지 단위 작업을 중심으로 각 단위 동작 단계별 인식 기준을 설명한다.

먼저, [Pick 작업은 로봇이 목표 물체를 정확히 파지하여 들어 올리는 과정으로, [Fig. 7]과 같이 네 단계의 단위 동작(Primitive Action)으로 구성된다. 첫 번째 단계인 Coarse Positioning에서는 로봇 말단이 물체의 대략적인 위치와 방향에 접근하는 과정을 수행한다. 이때 로봇 데이터와 비전 센서를 활용하여 로봇 말단과 목표 물체 간의 위치 및 방향 오차가 사전에 정의된 허용 범위 내에 있는지를 확인함으로써, 초기 접근의 정확도를 판단한다. 두 번째 단계인 Fine Positioning에서는 파지점을 기준으로 더욱 정밀한 위치 조정이 이루어진다. 이 과정에서는 로봇 말단의 위치 변화량과 제어 정확도 지표를 기반으로, 로봇이 목표 파지 위치에 정확히 접근하고 있는지를 평가한다. 세 번째

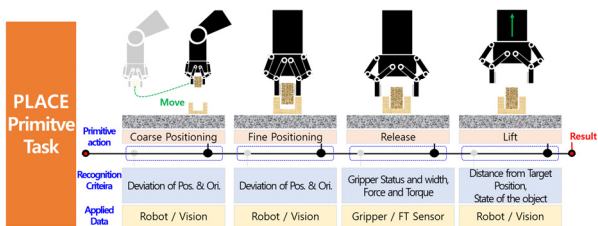


[Fig. 7] Recognition Criteria for Primitive Actions in Pick Tasks

단계인 Grasp에서는 물체 파지를 위한 그리퍼 동작이 수행된다. 이때 그리퍼 전류의 급격한 변화나 힘/토크 센서의 피드백 값, 그리고 그리퍼의 폐쇄 상태 및 파지 폭(state)을 종합적으로 분석하여, 물체가 정상적으로 파지되었는지를 판단한다. 마지막 단계인 Lift에서는 파지된 물체를 들어 올리는 동작의 성공 여부를 판단한다. 로봇 말단이 상승하는 궤적과 함께, 파지된 물체가 로봇과 동기화되어 안정적으로 이동하고 있는지를 비전 센서 및 말단의 위치 변화 데이터를 통해 확인한다. 동시에, 상승 중 발생하는 힘/토크 피드백의 안정성을 분석함으로써 파지 유지 상태에 이상이 없는지를 정량적으로 평가할 수 있다.

다음으로, Place 작업은 로봇이 파지한 물체를 목표 위치에 정확히 배치하는 과정으로, [Fig. 8]과 같이 네 단계의 단위 동작(Primitive Action)으로 구성된다. 첫 번째 단계인 Coarse Positioning에서는 로봇이 물체를 놓을 목표 위치에 대략적으로 접근하는 동작이 수행되며 로봇 말단의 현재 위치가 기준 좌표계에 대해 허용 가능한 오차 범위 내에 있는지를 판단한다. 두 번째 단계인 Fine Positioning 단계에서는 파지된 물체를 목표 위치에 정밀하게 정렬하기 위해 로봇 말단의 위치 편차가 충분히 작아졌는지를 판단한다. 이를 위해 비전 센서를 통한 위치 오차 측정이나, 로봇 말단의 미세 움직임 안정화 여부를 분석하여 목표 위치에의 정확한 정렬이 이루어졌는지를 평가한다. 세 번째 단계인 Release에서는 그리퍼를 개방하여 물체를 내려놓는 동작이 이루어진다. 이때 그리퍼의 개방 상태, 파지 폭(state)의 변화, 그리고 물체가 그리퍼로부터 물리적으로 분리되었는지 여부를 모니터링함으로써 물체 놓기 동작의 성공 여부를 판단한다. 비전 센서 또는 로봇 말단의 힘/토크 변화 또한 보조 판단 지표로 활용될 수 있다. 마지막 단계인 Lift는 로봇 말단이 물체를 놓은 후 상승하여 다음 작업을 준비하는 단계이다. 이 과정에서는 로봇 말단이 충분히 상승했는지를 확인함과 동시에, 놓인 물체가 목표 위치에 안정적으로 고정되어 있는지를 인식한다.

마지막으로, Push 작업은 물체를 수직 방향으로 눌러 소켓이나 지정된 위치에 삽입하거나 고정하는 작업으로, [Fig. 9]과 같이 네 단계의 단위 동작으로 구성된다. 첫 번째 단계인 Coarse Positioning에서는 로봇 말단이 삽입 대상에 근접한 위치로 도

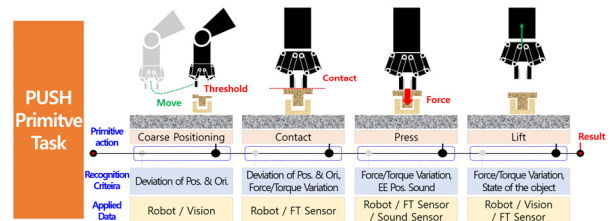


[Fig. 8] Recognition Criteria for Primitive Actions in Place Tasks

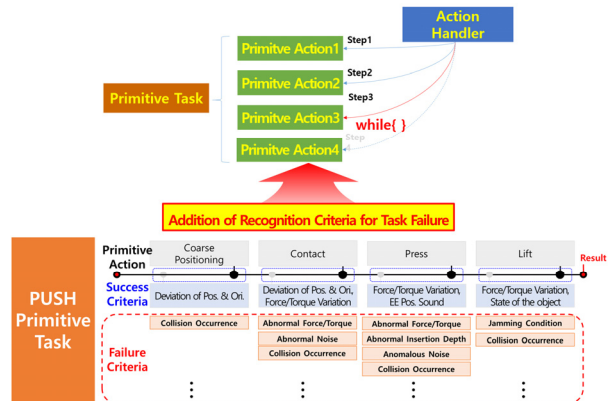
달했는지를 확인한다. 이때, 말단의 위치 및 자세가 사전 정의된 삽입 경로와의 편차 내에 존재하는지를 기준으로 판단하며, 주로 비전 및 로봇의 내부 위치 정보가 활용된다. 두 번째 단계인 Contact는 삽입 대상과의 실제 물리적 접촉을 인식하는 과정이다. Fine Positioning과 유사하게 정밀한 접근이 요구되지만, 이 단계는 힘/토크 센서를 통해 초기 접촉 시점의 반력 변화를 포착함으로써 수행된다. 특히, 비접촉 상태에서 접촉 상태로 전환되는 순간의 힘 피드백 곡선의 급변 또는 미세한 접촉력을 기반으로 접촉 여부를 정량적으로 평가한다. 세 번째 Press 작업은 부품을 삽입 방향으로 눌러 결합 또는 체결을 완료하는 과정이다. 이 과정에서는 삽입 중 저항력이 점진적으로 증가하고, 삽입이 완료될 때 힘이나 토크가 감소하는 패턴이 주요 판단 기준으로 사용된다. 또한, 삽입 완료 시 발생하는 체결음이나 로봇 말단의 위치 변화를 추가적인 판단 요소로 활용될 수 있다.

마지막 단계인 Lift는 삽입이 완료된 후 로봇 말단을 상승시켜 다음 작업을 준비하는 단계이다. 이때 말단의 상승 궤적, 힘/토크의 감소 추이, 삽입된 물체의 위치 안정성 등을 종합적으로 고려하여 삽입 작업의 완료 여부를 최종적으로 판단한다.

그러나 이러한 판단 기준은 정상적인 작업 흐름에서는 높은 신뢰도로 작동하였지만, 실험 과정에서 예기치 못한 상황이 발생하는 경우 시스템이 의도대로 동작하지 않는 한계가 나타났다. 예를 들어, 파지 실패, 삽입 시 오정렬, 비정상적인 저항 지



[Fig. 9] Recognition Criteria for Primitive Actions in Push Tasks



[Fig. 10] Failure Recognition Criteria for Primitive Actions (Example: Push Task)

속등 물리적으로 실패한 작업에서도, 단순히 성공 조건이 충족되지 않았다는 이유만으로 시스템이 무한 대기 상태에 빠지는 문제가 발생하였다. 이로 인해 단위 작업 루프에서 벗어나지 못하고 전체 공정이 정체되는 상황이 나타났다.

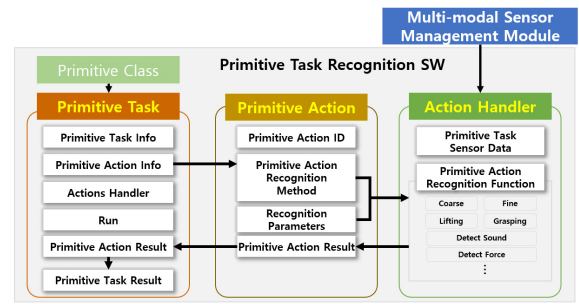
이를 해결하기 위해 본 연구에서는 [Fig. 10]과 같이 단위 작업의 성공 조건뿐 아니라, 실패 조건 또한 병행하여 정의하는 방식으로 판단 구조를 확장하였다. 즉, 일정 시간 내에 성공 조건이 만족되지 않거나, 센서 피드백이 비정상적인 패턴을 보일 경우 해당 작업을 명시적으로 '실패'로 판정하고, 이후 작업 전환 또는 오류 대응 절차가 실행되도록 설계하였다. 예를 들어, 그리퍼가 닫혔음에도 불구하고 힘/토크 변화가 없거나 부품이 움직이지 않는 경우는 파지 실패로 간주된다. 마찬가지로, Push 작업에서는 삽입 동작 중 목표 깊이에 대한 위치 변화 없이 과도한 힘/토크가 지속될 경우, 이는 삽입 저항 또는 삽입 깊이 이상으로 인한 비정상 상태로 인식된다. 이러한 실패 조건은 타입아웃, 센서 이상 패턴, 무응답 반복 등으로 정의되며, 공정 흐름의 정체를 방지하고 실시간 제어의 안정성을 높이는 데 기여한다.

결과적으로, 단위 작업 인식 기준을 성공 조건과 실패 조건의 이중 구조로 설계함으로써, 시스템이 단일 조건에만 의존하여 정체되는 문제를 해소하고, 공정 전체의 흐름을 유연하게 제어할 수 있는 기반을 마련하였다. 이러한 구조는 다양한 작업 조건과 예기치 못한 상황에서도 안정적으로 대응할 수 있는 판단 체계를 제공하며, 실시간 제어 성능과 오류 복구 능력을 동시에 강화하는 데 기여한다. 단위 작업을 구성하는 세부 단위 동작은 작업 간 공통적으로 반복되는 기본 행위 단위로 정의되기 때문에, 수행 대상이나 파라미터(위치, 방향, 힘 등)만 조정하면 다양한 작업 간에 재사용이 가능하다. 예를 들어, Coarse Positioning이나 Lift와 같은 동작은 Pick, Place, Push 작업 모두에 포함되며, 그 수행 맥락에 따라 유연하게 조합될 수 있다. 이는 단위 작업 수준에서의 모듈화뿐만 아니라, 단위 동작 수준에서도 높은 수준의 유연성과 확장성을 확보할 수 있게 한다.

결과적으로, 본 연구에서 정의된 단위 작업 기반 인식 구조는 복합 공정을 다루는 공정 인식 시스템의 핵심 구성 요소로 가능하며, 향후 다양한 제조 시나리오나 환경 변화에 적응 가능한 범용적이고 확장성 높은 인식 시스템 설계를 위한 중요한 기술적 기반으로 작용할 수 있다.

2.6 단위 작업 인식 소프트웨어 구조

앞서 정의한 단위 작업 및 단위 동작의 판단 기준을 기반으로, 본 절에서는 이를 실제 인식 시스템의 소프트웨어 아키텍처로 구현하기 위한 계층적 설계 방안을 제안한다. 제안된 구조는 [Fig. 11]에 나타난 바와 같이, 작업 인식과 판단 기능을 기능별로 분리하고 계층화함으로써, 유연성, 재사용성, 실시간성 등



[Fig. 11] Software Architecture of the Primitive Task Recognition Module

시스템 구현 관점에서의 요구사항을 효과적으로 충족한다.

제조 공정은 본질적으로 연속적이며 동적인 특성을 가지므로, 로봇이 이를 실시간으로 인식하고 판단하기 위해서는 작업 단계를 체계적으로 해석하고 구조화할 수 있는 메커니즘이 필요하다. 이에 본 연구는 인간 작업자의 인지-판단 구조를 모사하여, 전체 공정을 계층적으로 해석하고 인식 가능한 단위로 구성하는 소프트웨어 아키텍처를 설계하였다. 일반적으로 사람은 작업의 전체 목적(예: 체결)을 인지한 뒤, 이를 달성하기 위한 세부 행동들(예: 부품 집기, 정렬, 삽입)을 순차적으로 수행하며, 각 행동의 성공 여부를 시각, 촉각, 청각 등 다양한 감각 정보를 통해 실시간으로 판단한다.

이를 소프트웨어적으로 구현하기 위해, 전체 공정을 기능적으로 분리 가능한 단위 작업(Primitive Task) 수준으로 분해하고, 각 작업을 세부 단위 동작(Primitive Action)으로 계층화하였다. 각 계층은 작업 정의, 실행 판단, 센서 처리 기능을 명확히 분리함으로써, 복잡한 작업 환경에서도 유연하고 재사용 가능한 인식 체계를 실현할 수 있도록 설계되었다.

제안된 인식 구조는 크게 Primitive Task, Primitive Action, Action Handler의 세 계층으로 구성된다. 각 계층은 상위 작업 정의에서부터 하위 센서 기반 판단까지 기능적으로 분담되며, 상위 계층의 명령은 하위 계층으로 전달되고, 하위 계층의 판단 결과는 상위 계층으로 피드백 되는 방식으로 동작한다.

Primitive Task : 하나의 단위 작업을 의미적으로 정의하는 상위 모듈로 복수의 Primitive Action 시퀀스를 포함한다. 작업의 실행의 흐름과 조건을 정의하며, 하위 동작 결과를 통합적으로 평가하여 전체 작업의 성공 여부를 판단한다. 필요에 따라 동작 재시도, 작업 전환 등 판단 로직도 포함할 수 있다.

Primitive Action : Primitive Task를 구성하는 최소 단위의 조작 동작으로, 실행 메서드, 입력 파라미터, 판단 기준 등을 포함한다. 각 Action은 실행 시 판단이 필요한 시점에 Action Handler를 호출하며, 각 단위 동작의 수행 상태를 지속적으로 모니터링하고 판단 결과를 상위 Task에 전달한다. 하나의 Action은 다양한 Task에 공통적으로 재사용 가능하다.

Action Handler : 각 Action이 요구하는 판단 기능을 수행하는 모듈이다. 시각, 힘/토크, 사운드 등 멀티모달 센서를 수집하고 분석하여 동작의 수행 상태를 실시간으로 평가한다. 예를 들어 파지 여부 판별, 삽입 시 힘 패턴 분석, 위치 정합 상태 확인, 체결 완료 감지, 이상 소음 감지 등 다양한 판단 기능을 제공한다. 각 기능은 독립적으로 정의되어 다양한 Action에 선택적으로 연동될 수 있다. 또한 작업 유형이나 센서 구성, 인식 조건이 달라지더라도 새로운 판단 로직을 모듈 단위로 유연하게 추가, 적용할 수 있도록 설계되어 다양한 공정 변화에 효과적으로 대응할 수 있다.

이러한 구조는 계층 간 명확한 역할 분담과 인터페이스를 통해 작업 정의와 센서 기반 판단을 효과적으로 분리하고, 각 모듈의 독립성과 재사용성을 극대화하는 방향으로 설계되었다. 제한한 인식 구조는 복잡하고 동적인 제조 공정 환경에서도 신뢰도 높은 실시간 판단과 유연한 공정 제어를 가능하게 하는 기반 기술로 작용한다.

2.7 멀티모달 센서 관리 모듈

단위 작업 인식 소프트웨어의 핵심 구성 요소인 Action Handler는 로봇의 상태 정보, 시각, 힘/토크, 촉각, 사운드 등 다양한 센서 데이터를 기반으로 작업 상태를 평가한다. 그러나 작업 유형마다 요구되는 센서 종류와 조합이 달라지고, 각 센서의 데이터 형식 및 처리 방식 또한 상이하기 때문에, 이를 통합적으로 관리하고 효율적으로 연계할 수 있는 센서 관리 체계가 필요하다.

이를 위해 본 연구에서는 멀티모달 센서 관리자(Multi-modal Sensor Manager)모듈을 설계하였다. 이 모듈은 다양한 센서로부터 수집된 데이터를 일괄적으로 관리하고, 각 Primitive Action 또는 Handler가 요구하는 시점에 필요한 센서 데이터를 선택적으로 제공한다. 이를 통해 인식 모듈은 복잡한 센서 제어 및 동기화 로직으로부터 분리되어, 본래의 인식 기능에 집중할 수 있다. 이와 같은 센서 데이터의 흐름 및 관리 방식은 [Fig. 12]에도

시된 바와 같이, Multi-modal Sensor Manager와 인식 모듈 간의 데이터 요청·제공 관계로 구현된다. 이 구조를 기반으로, 멀티모달 센서 관리자는 다음과 같은 기능을 수행한다.

센서 데이터 수집 및 통합: 시각, 힘/토크, 사운드, 촉각 등 다양한 센서로부터 실시간 데이터를 수집하고, 시간 동기화 및 정규화를 수행한다.

요청 기반 데이터 제공: Action Handler가 요청한 센서 유형과 포맷에 맞춰 필요한 데이터만 필터링하여 제공함으로써, 불필요한 연산 자원 소모를 줄인다.

센서 선택과 조합의 유연성: 작업 종류나 환경에 따라 필요한 센서 조합을 동적으로 구성할 수 있어, 특정 공정이나 제품에 특화된 인식 구성이 가능하다.

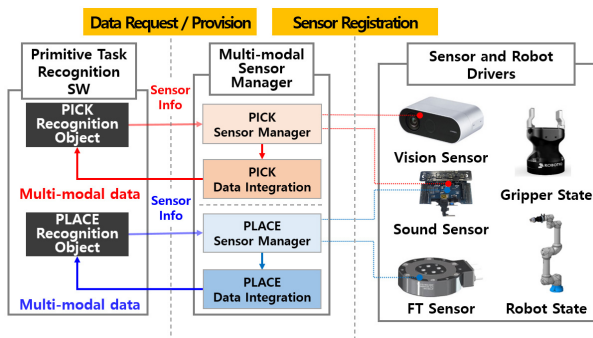
예를 들어, 동일한 ‘Insert’ 동작이라 하더라도 조립 대상의 재질, 허용 공차, 위치 정렬 조건에 따라 요구되는 센서 구성이 달라질 수 있다. 정밀 삽입 작업에서는 힘/토크 센서를 통해 삽입 저항을 실시간으로 모니터링해야 하고, 상대적으로 허용 오차를 큰 작업에서는 시각 정보만으로도 충분할 수 있다. 또한 일부 작업에서는 체결 완료 여부를 소리(예: 클릭음)로 판단해야 하는 경우도 있다. 이처럼 작업 조건에 따라 요구되는 센서 입력이 유동적이기 때문에, 센서 관리자는 각 Action에서 요청하는 센서 요구사항을 실시간으로 해석하고, 해당 데이터를 정확한 시점에 제공함으로써 전체 시스템의 안정성과 효율성을 높인다.

이 구조는 복잡한 제조 공정에서 요구되는 다양한 센서 사용 조건을 단일 통합 플랫폼 내에서 유연하게 처리할 수 있도록 설계되었으며, 공정 조건 변화나 시스템 구성 변경에도 센서 재설정을 최소화한다. 결과적으로, 멀티모달 센서 관리 모듈은 전체 인식 시스템의 모듈 간 독립성과 확장성을 지원하며, 실시간성·정확성·유연성을 동시에 만족시키며, 고신뢰성 작업 인식 프레임워크를 구성하는 데 핵심적인 역할을 수행한다.

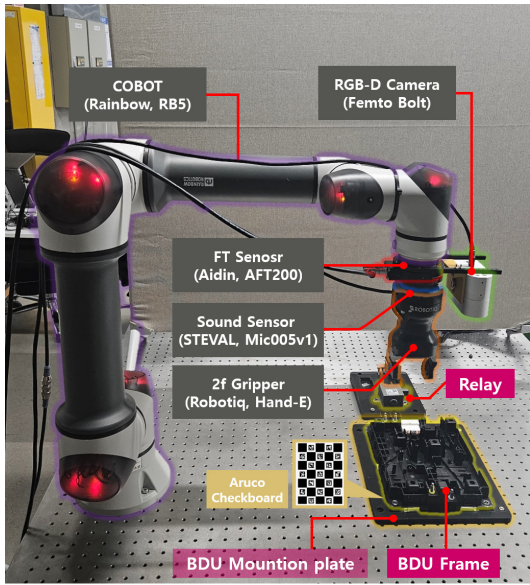
2.8 실험 및 적용 사례

제한한 계층형 단위 작업 인식 시스템의 실효성과 적용 가능성을 검증하기 위해 자동차 전장 부품인 BDU (Battery Disconnect Unit) 조립 공정의 일부 작업을 대상으로 실험을 수행하였다. 대표 작업으로는 릴레이 조립을 선정했으며 ‘Pick’, ‘Place’, ‘Push’의 세 가지 단위 작업(Primitive Task)으로 구성된다. 각 단위 작업은 로봇 상태 정보, 시각, 힘/토크, 사운드 등 다양한 센서 데이터를 통합적으로 분석하여, 조립 과정에서 발생하는 상태 변화를 신속히 파악하고 적절히 판단하는 것이 필요하다.

실험에 사용된 하드웨어는 [Fig. 13]과 같이 레인보우로보틱스의 6자유도 협동로봇 RB5와 Robotiq Hand-E 그리퍼로 구성되었다. 로봇 팔 플랜지 말단에는 에디인로보틱스의 6축 힘/토



[Fig. 12] Data Flow and Interaction between the Multi-modal Sensor Manager and Task Recognition Modules

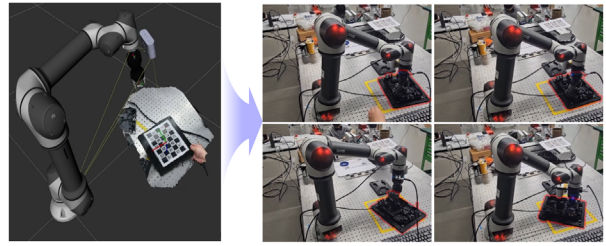


[Fig. 13] Hardware Con[figuration] of the Robotic System for BDU Manufacturing Process

크 센서(AFT200)가 부착되어, 작업 수행 중 발생하는 삽입 저항, 접촉력 변화, 과토크 등의 이상 상태를 감지하고 판단의 근거로 활용되었다. 또한 부품의 위치 및 자세 인식을 위해 Femto Bolt RGB-D 카메라는 핸드아이 구성으로 장착되었으며, 체결 완료 시 발생하는 클릭음을 감지하기 위해 그리퍼 인근에 마이크를 부착하여 사운드 기반 이벤트 인식을 구현하였다. 촉각 센서는 별도로 사용되지 않았으나, 그리퍼의 내장 전류 센서 및 상태 피드백 정보를 통해 물체 접촉 여부를 간접적으로 판단하였다. 전체 시스템은 ROS 기반의 제어 환경에서 구동되었으며, 각 인식 모듈은 Python 및 C++ 기반으로 개발되었다. 모든 센서 입력은 멀티모달 센서 관리자(Multi-modal Sensor Manager)를 통해 통합적으로 수집·관리되었고, Action Handler는 이를 기반으로 작업 상태를 실시간 평가하였다.

2.8.1 로봇 작업 구현 방식

로봇의 공정 작업 수행은 사용자가 사전에 지정한 작업 지점(Teaching Point)을 기반으로 계획되며, 각 동작은 개별 지점에서 수행되지만 전체 작업은 하나의 유기적인 모션 흐름으로 자율적으로 실행되도록 구성되었다. 제안된 인식 시스템은 전체 조립 과정을 실시간으로 모니터링하며, 단위 작업 단위로 해석하여 수행 상태를 판단한다. 본 연구에서는 BDU 보드를 고정할 수 있도록 설계된 전용 고정판에 Aruco 체크보드(Aruco Checkboard)를 부착하고, 로봇이 작업 시작 시 이를 인식하도록 시스템을 구성하였다. 로봇은 마커를 통해 작업판의 위치와 자세를 실시간으로 추정하고, 이를 기반으로 작업판 기준의 좌표계를 자동으로 설정한다.

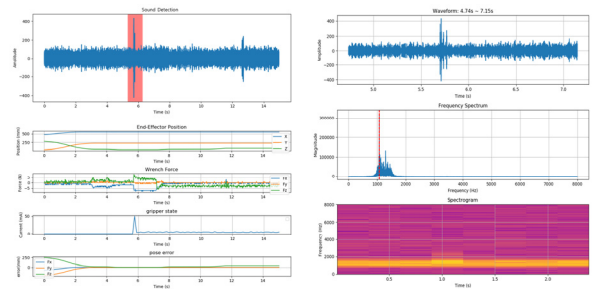


[Fig. 14] ArUco-Based Reference Recognition for Repetitive Assembly

[Fig. 14]는 마커 기반 기준 좌표계 재설정을 통한 조립 작업이 수행되는 과정을 나타낸다. 사용자는 이 기준 좌표계에 대해 작업 지점을 상대 좌표로 사전에 티칭하였으며, 작업판의 위치나 자세가 변경되더라도 마커 인식을 통해 기준 좌표계를 재설정함으로써, 기존에 정의된 상대 좌표를 그대로 활용하여 조립 작업을 반복 수행할 수 있다. 이와 같은 작업 구성 방식은 실험 수행 시 작업판의 위치를 유연하게 변경할 수 있도록 하며, 동일한 상대 좌표를 사용하더라도 로봇의 실제 조립 동작은 매 수행마다 미세하게 달라지게 된다. 이러한 차이는 단순한 위치 오차가 아니라, 다양한 조립 모션을 자연스럽게 유도하는 효과를 가지며, 제안된 인식 시스템의 유연성과 일반화 가능성을 평가하는 데 유리한 조건을 제공한다. 결과적으로, 본 실험 환경은 재티칭 없이 반복 수행이 가능하고, 다양한 작업 조건을 반영한 정량적 평가가 가능하다는 점에서, 유연하고 실효성 높은 평가 기반을 제공한다.

2.8.2 실험 결과 및 분석

총 50회의 반복 실험을 통해 제안한 계층형 단위 작업 인식 시스템의 정량적 성능을 평가하였다. 실험은 릴레이 삽입 작업을 중심으로 Pick, Place, Push 세 가지 단위 작업에 대해 수행되



[Fig. 15] Multi-modal sensor data acquisition and event detection during relay insertion: The left panel shows multi-modal data acquired during a relay insertion task: raw sound waveform, end-effector pose variation, gripper current, and positional deviation. The right panel visualizes audio analysis for click detection, including waveform, frequency spectrum, and time-frequency spectrogram

었으며, 인식 정확도, 실패 감지 성능, 실시간성, 작업 유연성 등을 중심으로 분석하였다.

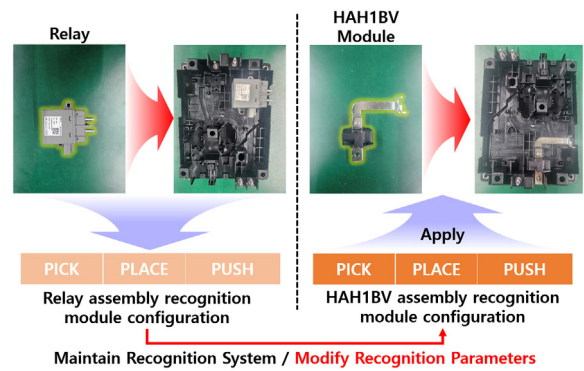
각 작업은 해당 단계에 대응되는 Action Handler를 통해 실시간으로 판단되었으며, 이를 위해 [Fig. 15]와 같이 로봇 말단의 위치, 속도, 관절 상태 등 동작 정보와 함께, 힘/토크 센서, 비전 카메라, 사운드 마이크, 그리퍼 전류 센서 등에서 수집된 멀티모달 센서 데이터가 통합적으로 활용되었다. 수집된 데이터는 정상 동작과 이상 상황을 구분하고, Action Handler가 작업 상태 및 이벤트 발생 여부를 판단하는 데 활용되었다. 전체 작업 흐름 내에서 단위 작업별 인식은 대체로 안정적인 정확도와 판단 일관성을 보였으며, 이는 제안한 시스템이 실시간 판단에 있어 충분한 신뢰성을 갖추고 있음을 보여준다.

Pick 및 Place 작업은 물체 위치 정합이 명확한 상황에서 높은 인식 일관성을 보였으며, Push 작업에서는 삽입 시 발생하는 힘/토크 패턴의 편차나 체결음 감지 실패 등으로 소수의 판단 오류가 관찰되었다. 추가적으로, 파지 실패, 삽입 저항 과다, 체결음 누락 등 일부 이상 시나리오를 의도적으로 삽입하여 시스템의 실패 감지 성능을 평가하였다. 그 결과 대부분의 이상 조건에서 센서 기반 판단 로직이 정상적으로 동작하며 상황을 탐지하였고, 전반적으로 90% 이상의 감지 성공률을 나타냈다. 일부 오판 사례는 센서 임계값 주변의 미세한 신호 차이로 인해 발생한 것으로 분석되었으며, 이는 향후 기준값 설정 정밀도 및 후속 처리 로직 강화 등을 통해 개선 가능할 것으로 판단된다. 이와 같은 결과는 제안한 인식 구조가 단순 반복이 아닌 작업 위치의 변동성과 다양한 작업 조건을 포함한 환경에서도 높은 신뢰도로 작업 상태를 인식하고, 이상 상황을 실시간으로 감지할 수 있음을 보여준다. 실험을 통해 본 시스템이 정밀성, 확장성, 유연성을 갖추고 있음을 확인하였다.

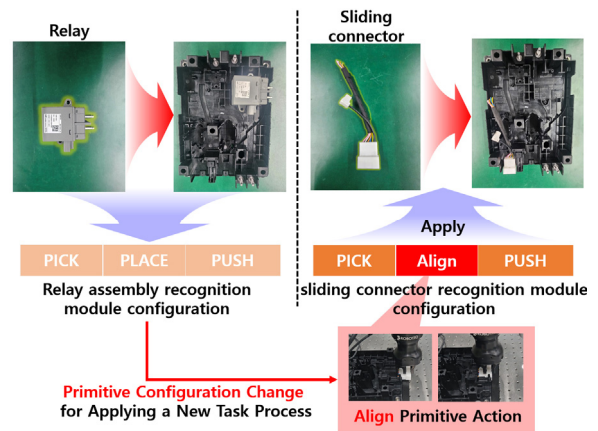
또한, 제안한 인식 구조의 범용성과 모듈 재사용 가능성을 확인하기 위해, HAH1BV 릴레이 모듈과 고정 브래킷이 결합된 일체형 조립체를 대상으로 추가 실험을 수행하였다. 해당 조립체는 형상 구조 및 체결 방식에서 기존 릴레이 삽입 작업과 유사하지만, 모듈 크기와 고정 지점, 삽입 방향 등에 있어 차이가 존재한다. 본 실험에서는 [Fig. 16]과 같은 방법으로 단위 작업 인식 소프트웨어의 인식 구조의 변경 없이, Action Handler 내의 판단 파라미터만 해당 조립 조건에 맞게 조정하여 인식 기능을 적용하였다.

그 결과, HAH1BV 조립 작업에서도 릴레이 삽입 작업과 유사한 수준의 인식 정확도와 판단 일관성을 확보할 수 있었으며, 제안 시스템이 형상 및 공정 조건이 다른 부품에도 유연하게 대응 가능함을 실증적으로 입증하였다. 이와 같은 결과는 향후 다양한 BDU 조립 시나리오에서의 범용적 인식 프레임워크로의 확장 가능성을 보여주는 기반으로 작용한다.

더불어, 기존의 수직 삽입 방식과는 다른 형태인 슬라이드



[Fig. 16] Adapting relay-based primitive task sequences via parameter tuning for HAH1BV assembly



[Fig. 17] Reconfiguration of primitive task sequences in the recognition module for slide-type assembly

방식(side-insert) 조립 작업에 대해서도 제안한 인식 구조의 적용 가능성을 평가하였다. 해당 작업은 커넥터를 측면에서 밀어 넣는 방식으로 수행되며, 기존 Push 작업과는 상이한 물리적 특성을 갖는다. [Fig. 17]의 절차에 따라 기존 Place 작업 대신 정렬 동작(Align)을 정의하고, 이후 측면 압입 또는 파지 상태에서의 삽입 동작으로 구성된 작업 흐름을 단위 작업 시퀀스로 재구성하였다.

이 실험에서는 기존 릴레이 조립 작업을 위한 단위 작업 구성과는 다른 새로운 작업 흐름이 요구되었으며, 이에 따라 기존에 정의된 Primitive Skill을 재조합하고, 작업 특성에 맞는 신규 단위 작업도 추가하여 슬라이드 방식 조립 작업에 대응하였다. 각 단위 작업은 해당 작업 방식에 맞춰 Action Handler 내 판단 기준이 새롭게 정의되거나 조정되었으며, 소프트웨어의 전체 구조나 모듈 구성은 그대로 유지된 채 인식 기능이 적용되었다. 그 결과, 슬라이드 조립 작업에서도 상태 인식 및 이상 감지 기능이 정상적으로 작동하였으며, 제안한 인식 구조가 조립 방향이나 절차가 상이한 새로운 작업 유형에도 유연하게 적용될 수 있음을 실증적으로 입증하였다.

3. 결론 및 향후 계획

본 연구에서는 복잡하고 비정형적인 제조 환경에서 로봇이 수행하는 작업을 정밀하게 인식하고 판단할 수 있도록, 계층적 구조에 기반한 모듈형 단위 작업 인식 시스템을 제안하였다. 작업 공정을 의미 있는 단위 작업(Primitive Task)으로 분해하고, 이를 다시 세부 동작(Primitive Action) 수준으로 계층화함으로써, 로봇의 실시간 작업 상태를 정밀하게 파악할 수 있도록 하였다. 또한, 멀티모달 센서 데이터를 활용하여 다양한 피드백을 통합 분석하고, 각 단계의 성공·실패 조건을 명확히 정의함으로써, 예상치 못한 상황에도 안정적으로 대응할 수 있는 실시간 판단 구조를 구현하였다.

제안된 프레임워크는 단순한 성공/실패의 이진 판단을 넘어, 각 동작의 정상 및 비정상 조건을 정량적으로 정의하고 이를 기반으로 작업 중 이상 징후를 조기에 감지할 수 있도록 설계되었다. 센서 피드백 분석을 통해 문제 발생 시점을 실시간으로 인식하고, 원인을 설명할 수 있도록 구성함으로써, 오류가 다음 공정으로 전파되기 전에 작업을 중단하거나 보정할 수 있는 실행 중심의 피드백 구조를 구현하였다.

실험을 통해 자동차 BDU 조립 공정 내 다양한 작업 시나리오에 본 시스템을 적용하였으며, 실제 릴레이 삽입, 커넥터 체결, 볼트 조임 등의 작업에서 높은 인식 정확도와 이상 감지 성능을 확인하였다. 특히 부품 형상이나 조립 방식이 바뀔 때에도 모듈 구조와 파라미터 조정을 통해 적용할 수 있음을 보였다. 제안된 인식 시스템은 계층적 구조와 모듈화된 설계를 기반으로 구성되어 있어, 신규 작업 공정이 추가되는 경우에도 기존에 정의된 단위 작업 및 단위 동작 모듈을 재사용하거나, 공정 특성에 따라 판단 전략을 재구성하거나 인식 파라미터를 조정하는 방식으로 쉽게 통합할 수 있다. 이러한 구조는 작업 정의와 판단 로직을 명확히 분리함으로써 시스템 구성 요소 간의 독립성을 보장하고, 다양한 제조 공정에 대한 높은 범용성과 유지보수 효율성, 그리고 구조적 확장성을 동시에 충족하는 공정 인식 프레임워크로 기능한다.

현재는 판단 파라미터가 실험적 설정에 따라 수동으로 구성되어 있으나, 향후에는 반복 수행 중 수집되는 정상 센서 시퀀스를 기반으로 조건값을 자동으로 보정하거나, 작업 초기에 수집된 제한된 데이터를 바탕으로 판단 임계 범위를 추정하는 기능을 통합할 예정이다. 이러한 기능은 anomaly-aware 판단 구조와 결합되어, 시스템의 자율성과 일반화 가능성을 한층 강화하는 기반으로 작용할 수 있다. 나아가 강화학습 기반의 전략 학습 모듈과 연계함으로써, 로봇이 작업을 수행하며 수집한 경험 데이터를 바탕으로 판단 기준을 점진적으로 조정하거나, 예외 상황에 대한 대응 전략을 스스로 학습할 수 있도록 시스템을

고도화할 계획이다. 이러한 구조는 판단 로직의 실시간성과 안정성을 유지하면서도, 예측 기반의 유연한 의사결정을 가능하게 하는 기술적 확장 경로로 기능할 것이다.

향후 연구에서는 다음과 같은 방향으로 본 시스템을 확장할 계획이다. 첫째, 현재는 단일 복합 작업 공정(BDU 조립 내 개별 공정 단위)에 대해 인식 체계를 구성하였으나, 이를 다수의 복합 작업이 연속적으로 구성된 상위 공정 수준(BDU 전체 조립 공정)으로 확장함으로써, 공정 간 연계 및 흐름 전체를 포함한 통합 인식 체계를 구축할 예정이다. 둘째, 본 연구에서는 실험 조건 하에서 사전 정의된 티칭 좌표 기반 경로를 통해 로봇의 작업이 이루어졌으나, 향후에는 강화학습 기반의 조립 전략과 연계하여, 학습된 모델이 조립 작업을 수행하는 과정에서 발생할 수 있는 비정상적 동작이나 실패 조짐을 조기에 감지하고 차단할 수 있도록, 제안된 작업 인식 시스템을 보완적 안전 판단 구조로 적용할 계획이다. 셋째, 본 연구는 조립 작업을 중심으로 시스템을 설계·검증하였으나, 제안된 인식 프레임워크는 그 구조적 일반성과 모듈화된 판단 방식 덕분에 이송, 검사, 인간-로봇 협업과 같은 다양한 제조 작업으로도 확장 가능성을 전제로 하고 있다. 향후 연구에서는 이러한 다양한 작업 유형에 대한 적용 실험을 통해 프레임워크의 범용성과 실용성을 더욱 폭넓게 검증하고자 한다.

References

- [1] Z. Fu, T. Z. Zhao, and C. Finn, "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," *arXiv:2401.02117*, 2024, DOI: 10.48550/arXiv.2401.02117.
- [2] D. Park, Z. Erickson, T. Bhattacharjee, and C. C. Kemp, "Multimodal execution monitoring for anomaly detection during robot manipulation," *2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, pp. 407-414, 2016, DOI: 10.1109/ICRA.2016.7487160.
- [3] A. Inceoglu, E. E. Aksoy, A. Cihan Ak, and S. Sariel, "Fino-net: A deep multimodal sensor fusion framework for manipulation failure detection," *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, pp. 6841-6847, 2021, DOI: 10.1109/IROS51168.2021.9636455.
- [4] S. Thoduka, J. Gall, and P. G. Plöger, "Using visual anomaly detection for task execution monitoring," *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, pp. 4604-4610, 2021, DOI: 10.1109/IROS51168.2021.9636133.
- [5] L. Duarte and P. Neto, "Classification of primitive manufacturing tasks from filtered event data," *Journal of Manufacturing Systems*, pp. 12-24, Jun., 2023, DOI: 10.1016/j.jmsy.2023.03.001.



강민석

2020 고려대학교 제어계측공학과(공학사)
2022 성균관대학교 기계공학과(공학석사)
2023~현재 한국로봇융합연구원 주임연구원

관심분야: 로봇 제어, 비전, 파지, 공정 자동화



박상현

2002 한국과학기술원 기계공학과(공학사)
2004 한국과학기술원 기계공학과(공학석사)
2024 울산과학기술원 기계공학과(공학박사)
2018~현재 한국로봇융합연구원 책임연구원

관심분야: 여유자유도 로봇, 비선형제어, 로봇 동역학, 로봇 설계



김무림

2004 한국과학기술원 기계공학과(공학석사)
2008 한국과학기술원 기계공학과(공학박사)
2008~2016 포항산업과학연구원 선임/책임
연구원
2016~현재 한국로봇융합연구원수석연구원/
본부장

관심분야: 재난대응로봇, 로봇 제어, 협동로봇, 공장자동화, 유압구동
로봇, 휴머노이드 로봇, 공중매니플레이션