FRAMED

Semantic understanding for Large-scale Robotics: Real-time foundation models enabling dynamic task and motion planning

Programm / Ausschreibung	FORPA, Dissertaionen 2024, Industrienahe Dissertationen 2025	**Status**	laufend
Projektstart	01.09.2025	**Projektende**	31.08.2028
Zeitraum	2025 - 2028	**Projektlaufzeit**	36 Monate
Keywords	Automation of outdoor machines; Foundation models; AI on the edge; Semantic planning; Task and motion planning		

Projektbeschreibung

Um den wachsenden Bedarf an autonomen Maschinen, die in komplexen, unstrukturierten Außenumgebungen arbeiten können, zu decken, entwickelt dieses Projekt ein neuartiges Framework für modernste künstliche Intelligenz für große Arbeitsmaschinen. Ziel ist es, Industrie-Roboter – etwa Gabelstapler und Kräne – in die Lage zu versetzen, anspruchsvolle Aufgaben in dynamischen Umgebungen wie Baustellen, Logistikhubs oder Holzlagerplätzen sicher, effizient und autonom auszuführen.

Derzeitige Robotersysteme für den Außenbereich stoßen an Grenzen: Sie leiden unter fragilen Wahrnehmungsmodellen, hohem Rechenaufwand und Abhängigkeit von cloud-basierten KI-Diensten, was zu Latenzproblemen führt und die Datensouveränität einschränkt. Dieses Projekt begegnet diesen Herausforderungen mit der Entwicklung eines schlanken, edge-fähigen KI-Systems, das modernste Umgebungserkennung, semantisches Verstehen und Echtzeit-Planung vereint. Im Zentrum steht ein selektives Zustandsraummodell (SSM) der Architektur Mamba, ursprünglich für Sprachverarbeitung entwickelt, das hier für die semantische Bild- und Szenenanalyse in Echtzeit adaptiert wird. Dank der linearen Laufzeit von Mamba kann das System dynamische Szenen kontinuierlich und schnell interpretieren und langfristige Aufgabenpläne erstellen – essenziell für den Robotereinsatz im Freien.

Durch die Kombination komprimierter Foundation-Modelle (z. B. Vision-Language-Modelle) mit Mamba können Roboter nicht nur Objekte erkennen, sondern auch deren physikalische Eigenschaften (z. B. stapelbar, beweglich) und funktionale Einsatzmöglichkeiten ableiten. Dieses semantisch-physikalische Verständnis wird in ein dynamisches Task-and-Motion-Planning-Modul (TAMP) eingespeist, das komplexe Aufgaben ad hoc zerlegt und Pläne bei veränderten Bedingungen anpasst. Das System läuft vollständig auf Edge-Hardware, verzichtet auf Cloud-Abhängigkeiten, minimiert Latenz und senkt den Energieverbrauch erheblich.

Die Innovationen werden auf zwei großen Maschinen validiert: einem autonomen Geländegabelstapler und einem Lkw-montierten Forstkran. Unter realen Bedingungen – etwa schlechter Sicht, unebenem Gelände oder unbekannten Hindernissen – wird das System auf Sicherheit, Effizienz und Anpassungsfähigkeit geprüft.

Das Projekt bietet erhebliche Vorteile:

Stärkung der industriellen Wettbewerbsfähigkeit durch technologische Souveränität, da auf nicht-europäische Cloud-Dienste verzichtet wird.

Nachhaltigkeit durch bis zu 70 % weniger Energieverbrauch dank effizienter Edge-KI und Reduzierung von Emissionen.

Förderung der Inklusion durch Senkung der Qualifikationshürden beim Bedienen schwerer Maschinen, was die Diversität in traditionell männlich dominierten Sektoren stärkt.

Durch die Verbindung von Foundation-Modellen mit echtzeitfähiger Robotik-Autonomie treibt dieses Projekt die Führungsrolle Österreichs in Arbeitsmaschinen voran und bereitet die Industrie auf die nächste Generation klimafreundlicher, robuster und flexibler Robotiklösungen vor.

Abstract

In response to the growing demand for autonomous machines capable of operating in complex, unstructured outdoor environments, this project proposes a novel framework that integrates advanced artificial intelligence with large-scale robotics. The focus is on enabling industrial robots—such as forklifts and cranes—to perform challenging tasks autonomously, safely, and efficiently in dynamic environments like construction sites, logistic hubs, and timber yards.

Current robotic systems for outdoor use are limited by brittle perception models, high computational demands, and reliance on cloud-based AI services, which introduce latency and compromise data sovereignty. Our project addresses these challenges by developing a lightweight, edge-deployable AI system that unites cutting-edge perception, semantic reasoning, and real-time planning. At its core lies a selective state-space model (SSM) architecture, Mamba, originally designed for language processing, adapted here for real-time, vision-based semantic scene understanding. Mamba’s linear-time efficiency allows continuous, high-speed interpretation of dynamic scenes and long-horizon task planning—capabilities crucial for outdoor robotics.

We fuse compressed foundation models (e.g., vision-language models) with Mamba to enable robots to not only identify objects but infer their physical properties (e.g., stackable, movable) and functional affordances. This semantic-physical understanding feeds into a dynamic task and motion planning (TAMP) module that can decompose complex tasks on the fly and adapt plans as conditions change. The entire system operates directly on edge hardware, eliminating cloud dependencies, minimizing latency, and drastically reducing energy consumption.

The system’s innovations will be validated on two large-scale platforms: an autonomous off-road forklift and a truck-mounted forestry crane. These robots will be tested under real-world conditions—including poor visibility, uneven terrain, and unfamiliar obstacles—providing rigorous benchmarks for safety, efficiency, and adaptability.

Our approach offers significant benefits:

Industrial competitiveness through technology sovereignty, as the solution minimizes reliance on non-European cloud

services.

Sustainability by cutting energy use up to 70% through efficient edge AI and reducing emissions.

Workforce inclusivity by lowering the skill barrier for operating heavy machinery, helping diversify traditionally male-dominated sectors.

By bridging foundation model capabilities with real-time robotic autonomy, this project advances Austria’s leadership in smart automation and positions its industrial sector for the next generation of climate-friendly, resilient, and inclusive robotics solutions.

Projektpartner

- AIT Austrian Institute of Technology GmbH