

Humanoide Roboter: Hören, Sehen und Verstehen

Ausgangslage

Auf der diesjährigen Aktionärsversammlung (AGM) verkündete Elon Musk, der wie ein Rockstar gefeiert wurde, dass Teslas Fabrik in Fremont, Kalifornien, im nächsten Jahr mit der Produktion seiner humanoiden Optimus-Roboter mit einer Stückzahl von 100.000 pro Jahr beginnen werde – eine Produktion, die schnell auf eine Million Roboter pro Jahr gesteigert werden soll. «Die Produktion von Optimus», erklärte Musk, «wird schneller sein als alles andere in der Geschichte der Menschheit.» Tesla ist bei weitem nicht der einzige Autobauer, der in Zukunft auf humanoide Roboter setzen möchte. Weitere Beispiele sind die Hyundai Motor Groups aus Korea, welche sich stark bei Boston Dynamics für die Entwicklung von humanoiden Robotern einsetzt, oder der chinesische Autobauer XPENG, welcher mit dem Roboter IRON ein wahres Meisterwerk gelungen ist. Mittlerweile gibt es gut ein Duzend Hersteller von humanoiden Robotern, welche mit immer ausgeklügelteren Robotern die Entwicklung schnell vorantreiben.

Die Entwicklung von fantastischer Hardware – einer Meisterleistung aus Mechanik und Elektronik – ist aber nur die eine Hälfte der Medaille. Die andere ist die Entwicklung der notwendigen Software zur Steuerung der Hardware und darüber hinaus die Automatisierungs-Software, welche es erst ermöglicht, dass die Hardware autonom agieren kann. Während die Steuerungssoftware schon weitgehend vorhanden ist, fehlt es im Bereich der Automatisierungs-Software noch an vielen Enden.

Zielsetzung / Methodik / Vorgehen

Automatisiertes Handeln ist meistens kontextabhängig und der Kontext eines humanoiden Roboters ist sein Einsatzort. Autonomes Handeln ist also stark davon abhängig, ob der Roboter in einer Autofabrik oder in einem Altenheim eingesetzt werden soll, aber auch davon, ob das Altenheim in China oder in Europa steht. Daraus ergibt sich, dass die Erstellung der Automatisierungs-Software in den Zielländern entwickelt oder zumindest adaptiert werden soll.

In diesem Masterprojekt geht es nur um den Kopf eines humanoiden Roboters und diesbezüglich um das Sehen und Hören. Der untenstehende Roboterkopf soll mit Software versehen werden, so dass der Roboter Personen unterscheiden, erkennen und wiedererkennen kann und auch verstehen kann, welche Person sich gerade mit ihm unterhält.

Dabei sollen das Robot Operating System ROS und bestehende Deep Neuronal Networks zum Einsatz kommen, so dass der Roboter in Echtzeit seine unmittelbare Umgebung hören, sehen und verstehen kann.

Der Kopf enthält vier Kameras, mehrere Mikrofone, einen Lautsprecher und ein animiertes Augenpaar. Die einzelnen Geräte sind über Controller mit einem NVIDIA Jetson verbunden, auf welchem die Automatisierungs-Software betrieben wird. Die Steuerungssoftware der einzelnen Geräte ist bereits vorhanden. Um ein Hör- und Sehverständnis zu ermöglichen, müssen alle sensorischen Daten zusammen in Echtzeit verarbeitet werden. Um diese Echtzeitverarbeitung geht es in diesem Projekt.

Verlangte Skills: Analytische Ingenieurfähigkeiten (Konzept, Aufbau, Durchführung), Freude an Bild- und Video- und Audioverarbeitung.

Teilaufgaben

Das Projekt kann in 3 Stufen durchgeführt werden: Fortlaufende Projekte P7 (ROS und bestehende Software verstehen bzw. überarbeiten), P8 (Hören und Hörverständnis) und P9 (Sehen und Sehverständnis)

Studienart: ☒ Vollzeitstudium
☒ Teilzeitstudium 50%

Projektorganisation: Einzelarbeit (im Umfeld eines Projektteams)

Projektfinanzierung: Innosuisse

Arbeitsort: Windisch

Advisor: Prof. Dr. Christoph Stamm

