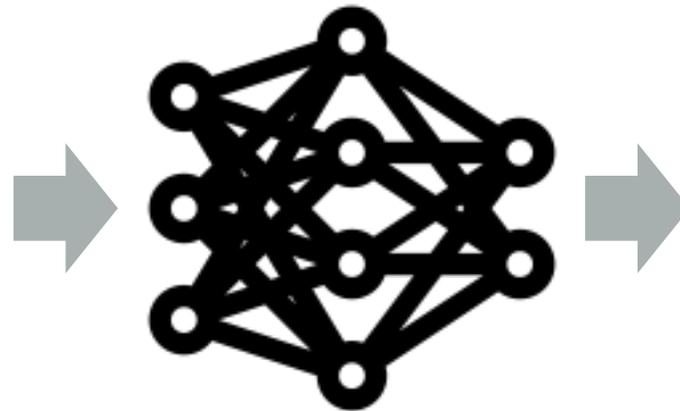


AUTOMATISCHE ERKENNUNG UND LOKALISIERUNG ZU REINIGENDER OBERFLÄCHEN

Florian Jordan, Fraunhofer IPA, florian.jordan@ipa.fraunhofer.de

Online-Seminar „Schlüsseltechnologien für die mobile Reinigung und Desinfektion“ am 18.11.2021



Motivation

- Der Roboter muss in der Lage sein die zu reinigenden Oberflächen in öffentlichen, sich verändernden Bereichen zu erkennen
 - Objekte sind nicht immer an exakt der selben Position (z.B. Türe ist offen statt zu)
 - Variierende Lichtverhältnisse erschweren die Erkennung, vor allem bei spiegelnden Oberflächen
 - Relevante Objekte können je nach Einsatzumgebung sehr unterschiedlich aussehen
- Genaue Lokalisierung relativ zum Roboter wichtig, damit Armbewegungen geplant werden können
- Roboter sollte einfach in neuer Einsatzumgebung installiert werden können und dabei alle Oberflächen vor Ort wahrnehmen können

→ Ziel: Entwicklung von zuverlässigen, möglichst allgemein funktionierenden Erkennungsalgorithmen

Grundsätzliche Vorgehensweise

Verwendung von RGB Bildern und Tiefendaten, aufgenommen von Sensoren am Handgelenk des Roboters.

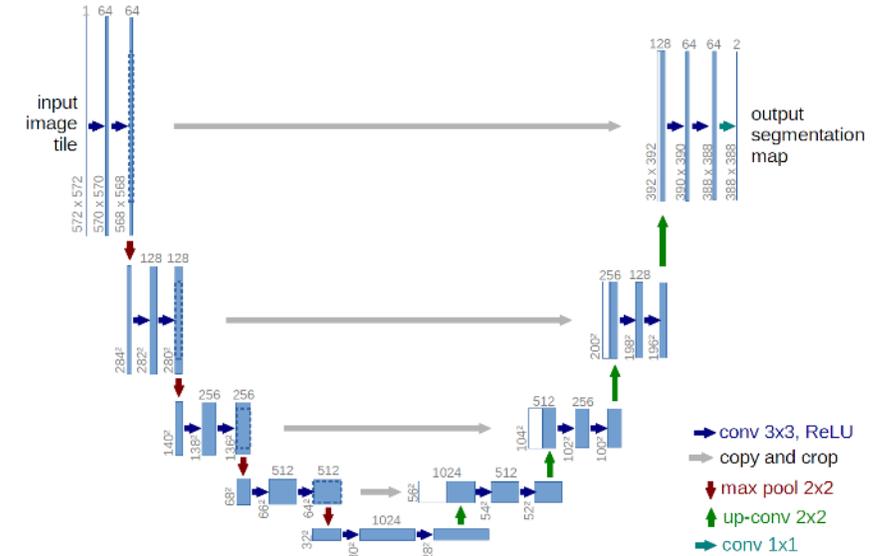
1. Erkennung der relevanten Oberflächen in RGB Bild aus der Ferne → Abdeckung eines größeren Ausschnitts der Szene um z.B. auch bei offenen Türen die Klinken/Knäufe erkennen zu können
2. Bestimmung einer Roboterposition vor der Oberfläche, sodass Roboter parallel dazu steht
3. Erkennung der relevanten Oberfläche in RGB Bild aus der Nähe
4. Abtastung der Oberfläche mithilfe eines 3D Sensors und Lokalisierung der relevanten Oberfläche unter Einbeziehung der Erkennung in RGB Bild



Verwendete 2D/3D Sensoren am Arm des Roboters.

Objekterkennung in RGB Bildern

- Lokalisierung in 2D RGB Bildern: Verwendung eines U-Net-style CNN → Deep Learning basierte One-Stage Object Detection
- Erkennung der zu reinigenden Objekte in unterschiedlichen Bildkanälen (Knäufe, Türklinken, Lichtschalter/Knöpfe)
- Unterscheidung in verschiedene Reinigungstypen (Objektklassen) ermöglicht spezielle Anpassung der Reinigungsbewegung an Oberfläche der Objekte
- Erzielte Genauigkeit auf Testdatensatz: ca. 84%



Verwendete U-Net-style Architektur für Lokalisierung der Objekte in RGB Bildern.



Ergebnisse der Detektion von zu reinigenden Objekten in RGB Bildern.

Objekterkennung in RGB Bildern



Weitere Ergebnisse der Detektion von zu reinigenden Objekten in RGB Bildern.

Objekterkennung in RGB Bildern



Weitere Ergebnisse der Detektion von zu reinigenden Objekten in RGB Bildern.

Objekterkennung in RGB Bildern

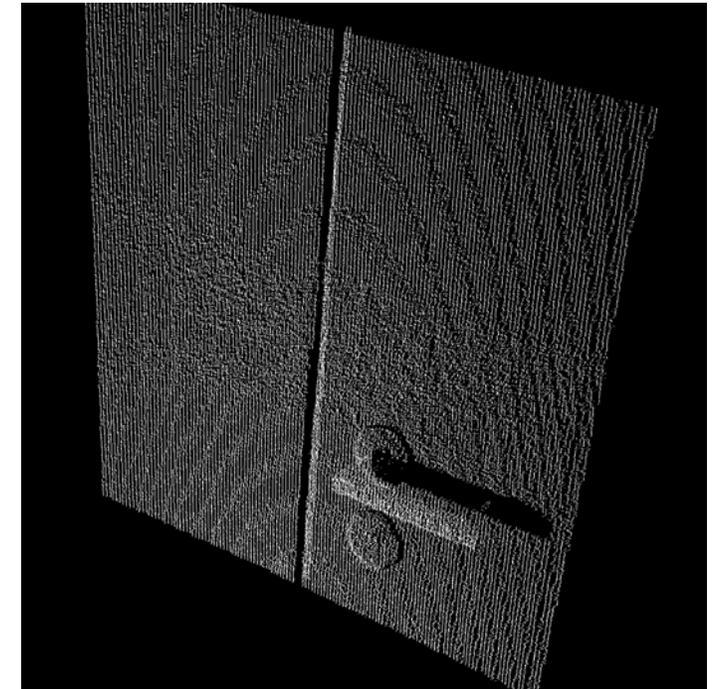
- Weitere Eckpunkte der Erkennung in RGB Bildern
 - Trainings-/Testdatensatzgröße: 2929 (Training) / 204 (Test)
 - Ca. 84% Erkennungsrate auf Testdatensatz
 - Ausführung auf Nvidia Jetson AGX Xavier (Embedded AI Processing Modul)
 - Berechnungsgeschwindigkeit auf 960x544 Bild: ca. 5.31Hz → Nicht optimiert, kann noch gesteigert werden
 - Objekte von Praxistests nicht im Trainingsdatensatz, konnten aber zuverlässig erkannt werden
 - Fehlerrate bei Praxistests: 1 Fehlschlag bei 45 Erkennungen
- Bestimmung der Roboterposition vor der Oberfläche auf Basis der Laserscanner-Daten (in Richtung der Detektion im RGB Bild) → RANSAC basierte Liniendetektion

Neue low-cost 3D-Sensorik zur Abtastung der Oberflächen

- Erkennung in RGB Bildern liefert grobe Position der zu reinigenden Oberflächen in der Umgebung
 - Für Serviceroboter exakte Position relativ zum Roboter notwendig
- Daher: Entwicklung eines auf metallene Materialien angepassten 3D-Sensors, um Oberflächen abzutasten
- 3D-Erfassung von Objekten im Abstand bis zu 1m mit einer Genauigkeit im Millimeterbereich
- Optisches Messprinzip
 - Erweiterung des Lichtschnittverfahrens durch zusätzliche Hardwarekomponenten und einen darauf abgestimmten Auswertalgorithmus
 - Im Vergleich zu herkömmlichen optischen Sensoren deutlich zuverlässigere Erfassung von spiegelnden Objekten wie Türklinken, metallene Aufzugknöpfe und Lichtschalter
- Scan-Bewegung erfolgt durch Bewegung des Handgelenks



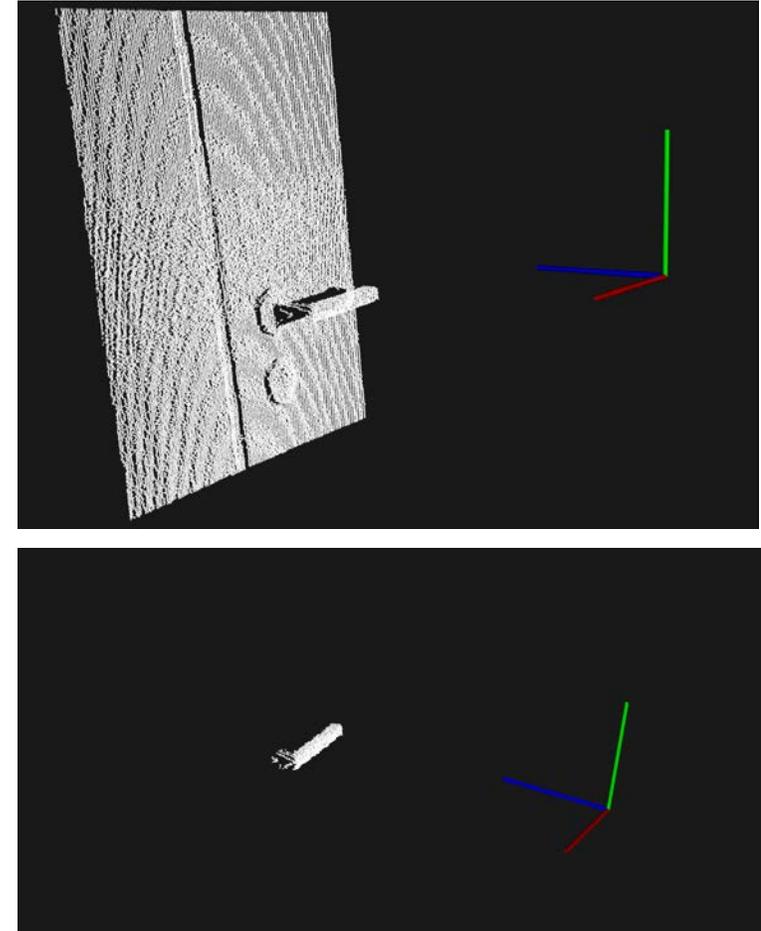
Entwickelter 3D-Sensor.



3D-Scan einer metallenen Türklinke

Objekterkennung in 3D Punktwolken

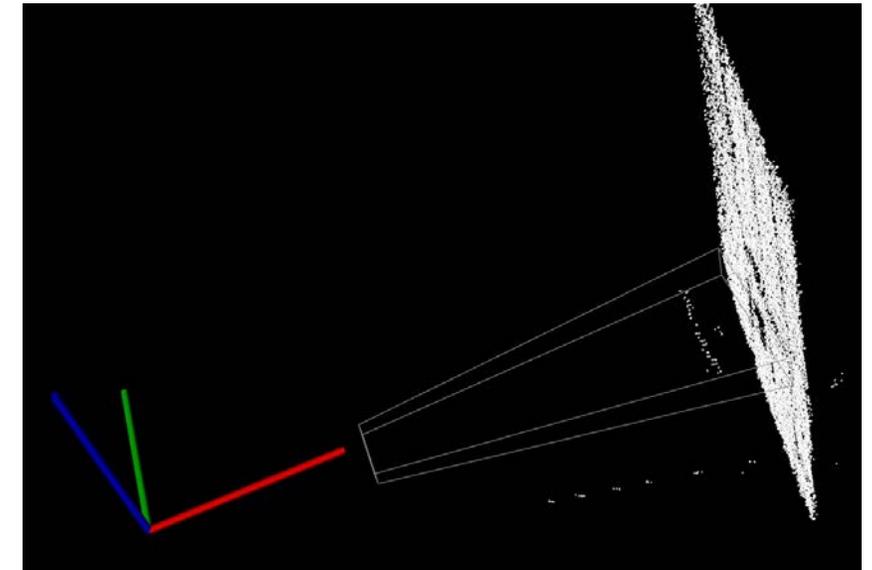
- In den durch den entwickelten 3D-Sensor aufgenommenen Tiefendaten kann mithilfe des RANSAC Verfahrens die Wandebene erkannt werden
 - Basierend auf der erkannten Ebene können die Tiefendaten der einzelnen Objekte segmentiert werden
 - Durch euklidisches Clustering werden die zu der relevanten Oberfläche gehörenden Tiefendaten extrahiert
 - Eine rotierte Bounding-Box liefert die Orientierung der zu reinigenden Oberfläche
- Liefert eine genaue räumliche Position und Ausdehnung der zu reinigenden Objekte



Aufgenommene Punktwolke einer Tür mit Türklinke und darin erkannte Türklinke.

Objekterkennung in 3D Punktwolken

- Oberflächen wie Lichtschalter stehen nur sehr schwach hervor, Aufzugsknöpfe sind meist wandbündig → nicht durch Clustering extrahierbar
- Erkennung der zu reinigenden Oberflächen in RGB Bildern liefert bereits 2 Koordinaten der Position
 - Kombination von 2D Erkennung und 3D Daten ermöglicht Lokalisierung von Lichtschaltern und Knöpfen
- Dazu: Transformation der 2D Erkennung in Pyramidenstumpf (da die genauen Tiefeninformationen fehlen) und Ausschneiden der 3D Punkte innerhalb dieses
- Zusätzliches Filtern und Clustern ermöglicht auch für solche Objekte eine genaue Lokalisierung

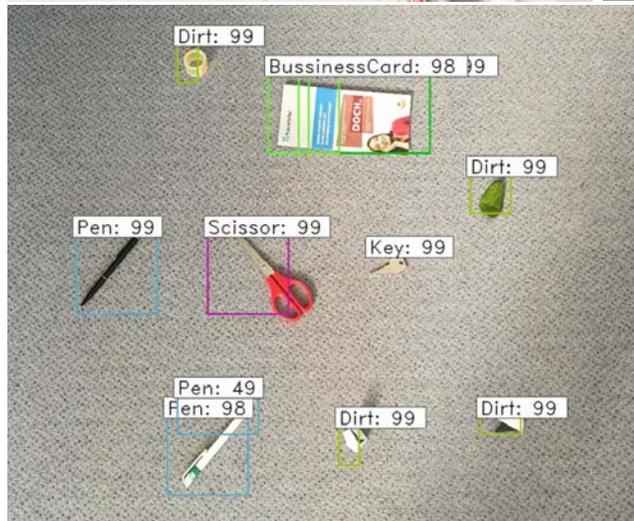
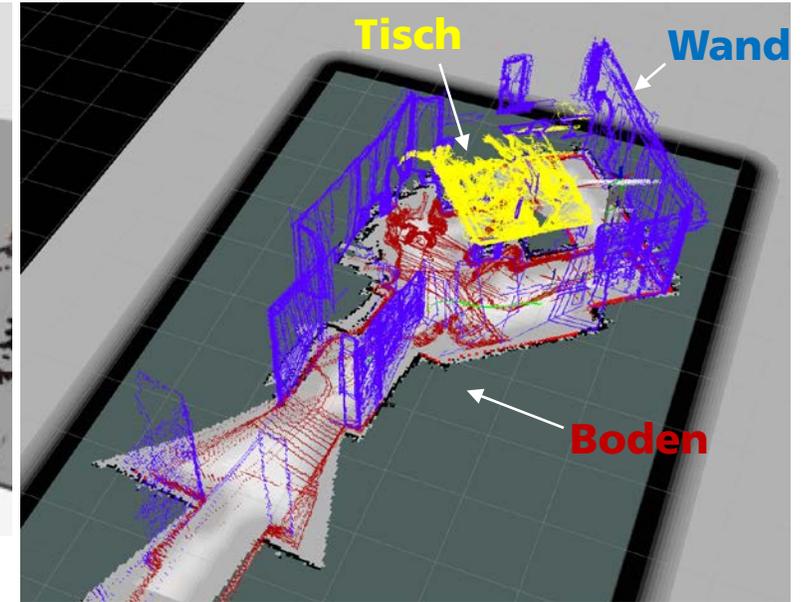
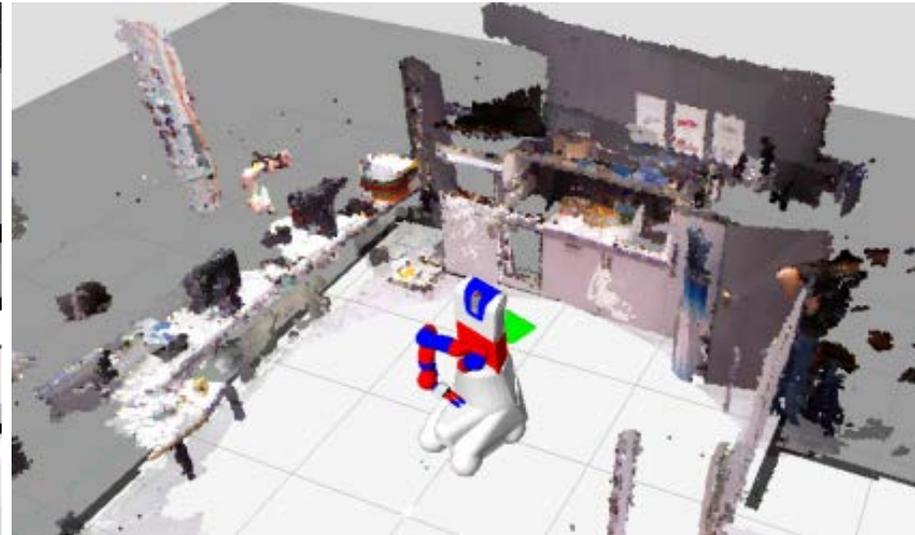
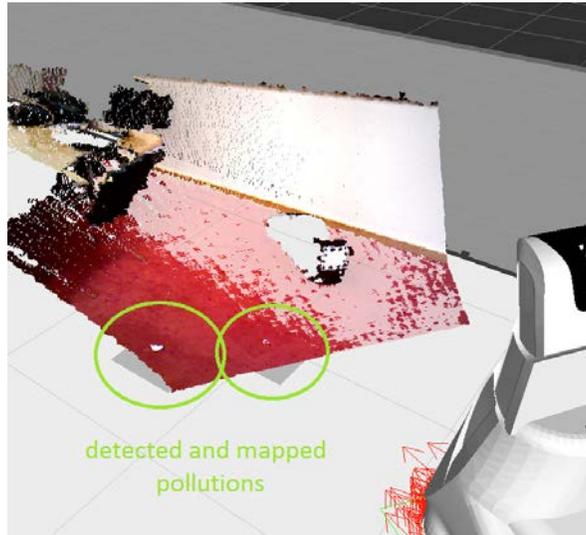


Transformation der 2D Erkennung in 3D Koordinaten.

Objekterkennung in 3D Punktwolken

- Weitere Eckpunkte der Erkennung in 3D Punktwolken:
 - Erfassungszeit der Punktwolke: ca. 30s → limitiert durch Möglichkeiten des Arms
 - Berechnungszeit <1s (keine genaue Messung, da insignifikant im Vergleich zu anderen Teilen)
 - Vorwissen (relevante Objekte an einer geraden Ebene) ermöglicht robuste Verarbeitung in allen Einsatzfeldern
 - Lokalisierung der Oberflächen in Punktwolken hat bei Praxistests nie fehlgeschlagen

Bildverarbeitung über das Projekt hinaus



Bildverarbeitung über das Projekt hinaus

