

LERNFÄHIGE 3-D-OBJEKTERKENNUNG FÜR SERVICEROBOTER

Fraunhofer-Institut für Produktions- technik und Automatisierung IPA

Nobelstraße 12
70569 Stuttgart

Ansprechpartner
Dipl.-Inf. Jan Fischer
Telefon +49 711 970-1191
jan.fischer@ipa.fraunhofer.de

M.Eng. Jens Kubacki
Telefon +49 711 970-1043
jens.kubacki@ipa.fraunhofer.de

www.ipa.fraunhofer.de

Ausgangssituation

Für die autonome Ausführung von Handhabungsaufgaben in veränderlichen Alltagsumgebungen muss ein Serviceroboter in der Lage sein, Objekte in 3-D zu detektieren und deren genaue Lage (Position und Orientierung) zu berechnen. Aufgrund der Umgebungs- und Objektvielfalt sind dabei niemals die Modelle aller relevanten Objekte verfügbar. Im Sinne einer maximalen Flexibilität und Erweiterbarkeit sollte der Roboter neue Objekte selbstständig »erlernen« und somit seine Fähigkeiten kontinuierlich erweitern können.

Folgende Randbedingungen müssen durch die Lern- und Erkennungsalgorithmen erfüllt werden:

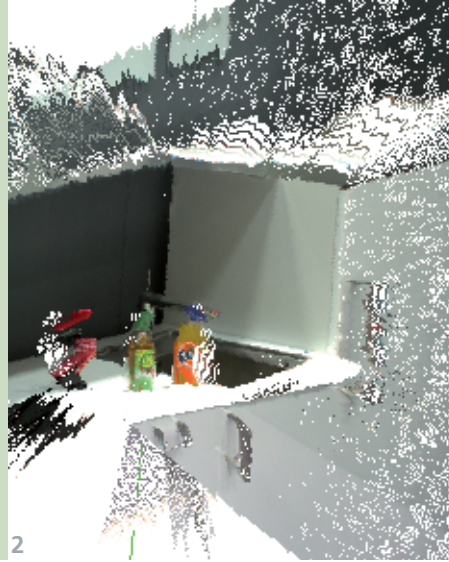
- Erkennung der Objektposition und Orientierung
- Einfaches Einlernen neuer Objekte durch den Benutzer

- Hohe Robustheit gegenüber Veränderungen in den Beleuchtungsverhältnissen am Einsatzort
- Hohe Robustheit gegenüber Verdeckungen (bis zu 80 Prozent möglich)
- Robustheit gegenüber Deformationen oder partiellen Oberflächenveränderungen
- Erkennung in weniger als einer Sekunde

Das nachfolgend dargestellte Verfahren wurde am Fraunhofer IPA zum Einlernen und Erkennen typischer Alltagsgegenstände entwickelt.

Sensoraufbau

Tiefenbildkameras oder Time-of-flight-Sensoren liefern ähnlich wie Farbkameras ein Bild, jedoch steht in jedem Bildelement ein Abstandswert anstelle der üblichen RGB-Werte. Durch die Kopplung einer Tiefenbildkamera mit einer Farbkamera können herkömmliche 2-D-Erkennungsverfahren



aus der klassischen Grau- und Farbbildverarbeitung zu 3-D-Erkennungsverfahren erweitert werden. Dafür wird mithilfe einer speziellen Kalibrierung zwischen den Sensoren ein approximatives Farbbild berechnet, welches mit seinem Pixelkoordinatensystem und dem der Tiefenbildkamera in Deckung gebracht wird. Damit ist für jedes Pixel im Bild nicht nur der entsprechende Farbwert, sondern auch der Abstand des zugehörigen Umgebungspunkts bekannt.

Objektrepräsentation

Objekte werden durch sog. »6-D-Merkmalpunktwolken« repräsentiert. Diese bezeichnen eine Menge von 3-D-Punkten, die im Farb- und Tiefenbild bestimmten wiedererkennbaren Mustern zugeordnet werden können. Durch die Kopplung von Farb- und Tiefenbild ist es möglich, jedem im Farbbild detektierten Merkmalspunkt eine eindeutige 3-D-Position und -Orientierung zuzuordnen. Um die Lage der einem Objekt zugeordneten Merkmalspunkte zueinander zu definieren, wird für jeden Merkmalspunkt ein lokales Koordinatensystem errechnet, das dessen Lage relativ zu einem zentral im Objekt liegenden Koordinatensystem beschreibt.

Die beschriebene Objektrepräsentation ist geeignet, um variantenreiche Objektarten zu modellieren und zu detektieren – teilweise sogar mechanisch bewegbare oder plastisch verformbare Objekte.

Detektion bekannter Objekte

Um ein bestimmtes Objekt in einer vorgegebenen Szene zu detektieren, werden zunächst alle in der Szene detektierten Merkmale zu einer Szenenmerkmalspunktwolke zusammengefügt. In dieser wird nachfolgend nach möglichen Korrespondenzen mit der Objektmerkmalspunktwolke gesucht. Für jede untersuchte Korrespondenz wird berechnet wie das zentrale Objektkoordinatensystem in der Szene läge, wenn der Merkmalspunkt tatsächlich zum Objekt gehörte. Von den errechneten Hypothesen wird diejenige mit der höchsten Wahrscheinlichkeit ausgewählt. Eine verfeinerte Suche kann je nach Genauigkeitsanforderungen nachgeschaltet werden. Die Erkennung dauert ca. 2 Sekunden. Die Genauigkeit bewegt sich entsprechend der Auflösung und Präzision aktuell verfügbarer Tiefenbildkameras im Bereich von 1 bis 2 cm.

Einlernen neuer Objekte

Für das Einlernen neuer Objekte stehen zwei verschiedene Lernmodi zur Verfügung: entweder wird das Objekt in den Roboter greifer gegeben und der Roboter dreht es beim Einscannen oder der Mensch zeigt das Objekt. Das dabei angewandte Lernverfahren besteht aus den folgenden, aufeinander aufbauenden Algorithmen:

1. Akquisition einer 3-D-Farbbildsequenz mit unterschiedlichen Objektansichten
2. Raumsegmentierung der Objektansichten auf Basis einer 3-D-Maske

3. Berechnung stabiler 6-D-Merkmalpunkte (3-D-Position und -Orientierung) für alle Ansichten
4. Fusion der partiellen Objektmerkmalspunktwolken in ein Objektmodell

Das errechnete Objektmodell wird in einer Datei abgespeichert und kann zur Erkennung genutzt werden.

Unser Leistungsangebot

Das vorgestellte Verfahren ist Teil einer am Fraunhofer IPA entwickelten Software-Bibliothek für 3-D-Bildverarbeitungsprobleme. Die darin enthaltenen Algorithmen sind dabei nicht auf die Erkennung und Lagebestimmung der in den Beispielen dargestellten Haushaltsgegenstände beschränkt. Auch herausfordernde industrielle Einsatzszenarien, wie die Detektion von Stellrädern oder von Kuheutern wurden mit den am Fraunhofer IPA entwickelten 3-D-Bildverungsverfahren bereits erfolgreich gelöst.

Das Fraunhofer IPA unterstützt Sie in allen Phasen der Entwicklung Ihrer spezifischen Erkennungssoftware: von der Systemkonzeption und Auswahl geeigneter Sensoren und Bildverarbeitungsverfahren bis zur Weiter- oder Neuentwicklung geeigneter Algorithmen für Ihren spezifischen Anwendungsfall.

- 1 *Eigenständiges Objektlernen (detektierte Merkmalspunkte farblich dargestellt, links)*
- 2 *3-D-Ansicht einer Szene*
- 3 *Haushaltsassistent »Care-O-bot® 3« beim Greifen einer Flasche nach erfolgreicher Detektion*