

Vom Greifen zum Begreifen

Interaktive Strategien zur Modelldatenakquisition und Algorithmus für die Modellsynthese in der Objekterkennung

M.Eng. **Jens Kubacki**, Fraunhofer IPA, Robotersysteme, Stuttgart

Kurzfassung

In der Vision sollen mobile Serviceroboter komplexe und anspruchsvolle Aufgaben im Einsatzort effizient und intelligent lösen können. Ein Problem bei der Programmierung dieser Roboter ist, dass die zur Aufgabe gehörenden Elemente (Werkzeuge, Geräte oder Räumlichkeiten) zur Zeit der Entwicklung noch nicht bekannt sind. Auch im Bereich der Objekterkennung sollen Roboter deshalb mit Mechanismen ausgestattet werden, die das Erlernen von relevanten Aspekten neuer Objekte langfristig gewährleisten. In diesem Beitrag werden aktuelle Arbeiten zur Modelldatenakquisition und Modellsynthese vorgestellt, die am Fraunhofer IPA im Rahmen des EU-Projektes Cogniron (www.cogniron.org) integriert werden.

Einleitung und Motivation

Während früher das Bildverarbeitungssystem als Komponente des Roboters betrachtet wurde, wird heute der Roboter auch als Ressource für das Bildverarbeitungssystem gesehen. Die „Verkörperung“ des Bildverarbeitungssystems ist für die Lösung eines fundamentalen Problems möglich: zur Modelldatenakquisition. Im Thema Modelldatenakquisition steht nicht die Art der Modelle im Vordergrund der Diskussion, sondern wie diese Modelle „in den Kopf“ des Roboters kommen. Die Modelltypen sind dabei üblicherweise durch die kontribuierenden Disziplinen der Robotik definiert.

Das Gebiet der Modelldatenakquisition umfasst zwei Teilbereiche:

1. Verkörperte Strategien zur Modelldatenakquisition
2. Algorithmen zur Modellsynthese

In diesem Beitrag werden aktuelle Arbeiten und Ergebnisse für diese beiden Bereiche mit Bezug auf das Problem der Objekterkennung und 6D-Lageschätzung für moderne kognitive Roboter vorgestellt.

Strategien für die Modellakquisition in der Objekterkennung und –lageschätzung

Wir unterscheiden zwei Arten von Strategien zur Modellakquisition: autonome Strategien und interaktive Strategien. Autonome Strategien sollen den Roboter befähigen, Modelle selber, ohne Beisein des Benutzers zu akquirieren. Im Bereich der Objekterkennung und 6-D-Lageschätzung gab es bisher weniger Arbeiten zur autonomen Modelldatenakquisition. Komplett autonome Strategien sind ohne massives Vorwissen nur schwer zu realisieren. Woher soll der Roboter z. B. wissen, dass eine Pflanze am Topf gegriffen werden muss?

Deshalb setzen wir am Fraunhofer IPA auf interaktive, benutzergestützte Strategien zur Modelldatenakquisition. Dabei sind verschiedene Strategien möglich, die jeweils andere Anforderungen an den Benutzer, an die technische Ausstattung des Roboters und an die anschließenden Algorithmen für die Modellsynthese stellen. Jede Strategie zur Modellakquisition setzt sich aus den folgenden Stufen zusammen.

1. Initialisierung
2. Variation der Ansichten und Segmentierung der Objektansichten
3. Speicherung der segmentierten Beispielbilder

In Bild 1 sind zwei mögliche Strategien dargestellt, die aktuell implementiert und getestet werden.

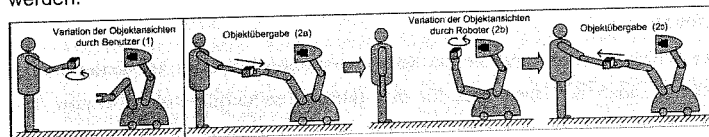


Bild 1: Zwei Strategien für die interaktive, benutzergestützte Modelldatenakquisition.

In Strategie 1: „Variation der Objektansichten durch den Benutzer“ (Bild 1, links) zeigt der Benutzer verschiedene Ansichten. In Strategie 2: „Variation der Objektansichten durch den Roboter“ (Bild 1, Mitte/rechts) muss zuerst das Objekt an den Roboter übergeben werden. Objektübergabe wird durch einen Greifer möglich, der mit Infrarot-Sensoren in den Greiferfingern ausgestattet ist, mit einem handelsüblichen Kraft-Momenten-Sensor und Motorstrommessung.

Algorithmus für die Modellsynthese in der Objekterkennung und 6-D-Lageschätzung

Es wurde bisher darauf verzichtet, den Begriff der „Ansicht“ genauer zu definieren. In dieser Arbeit sind die Ansichten sind Ausgangsdaten eines 3-D-Farbsensorsystems, das im Folgenden vorgestellt wird und in Bild 2 mit einigen Illustrationen der Resultate gezeigt ist.

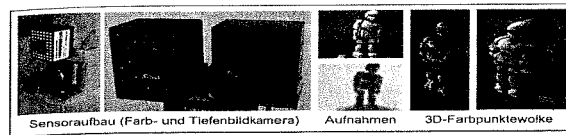


Bild 2: Kombiniertes Sensorsystem aus Farb- und Tiefenbildkamera. Links: verschiedene Anordnungen. Rechts: Farbbild, Tiefenbild (in Grauwerten kodiert), fusionierte Farbpunktewolke und interpolierte fusionierte Farbpunktewolke.

Es sind neuerdings 3-D-Kameras erhältlich (siehe www.mesa.ch, www.pmd-tec.com oder www.canesta.com), die per Pixelelement des Bildsensors die Laufzeit des Lichtes eines Projektors vom Sensor ausgesandt, an einem Objekt reflektiert und zurück messen können (TOF = time-of-flight). Am Fraunhofer IPA wurde die Kamera SwissRanger 3000 von MESA mit einer konventionellen Farbkamera, DFK41F02 (www.theimagingsource.com), kombiniert und durch eine spezielle Kalibrierung beide Bilder auf ein gemeinsames Bildformat abgebildet (Bild 2). Das führt zu einem neuen Sensorsystem, das 3-D-Farbpunktewolken in relativ hohen Wiederholraten (>10 Hz) liefern kann (siehe Bild 1, oben, für Beispiele der Farbpunktewolken). Dadurch dass alle Pixel dieses resultierenden Sensors nun den x, y, und z-Koordinaten im Koordinatensystem des Sensors und R, G, B-Farbdaten zugeordnet werden können, wird es möglich die Farbsichten eines Objektes vom Hintergrund zu trennen, indem im 3-D-Raum ein Bereich (z. B. Kugel oder Quader) als Maske verwendet wird. Nun wird ein neues Farbbild berechnet, das nur die Pixel einfärbt, die sich in dem maskierten 3-D-Bereich finden und alle anderen Pixel auf R=G=B=0 setzt. In Bild 3 sind Bilder von Objekten mit, die mit dieser Methode Segmentiert wurden dargestellt. Wir nennen dieses Verfahren Raumsegmentierung, da es Farbdaten mit Hilfe der 3-D-Informationen segmentiert (siehe Bild 3).

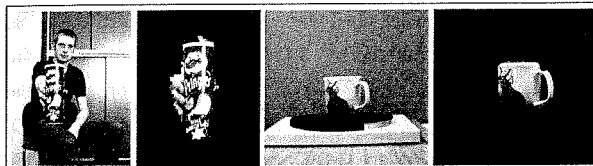


Bild 3: Das Konzept der Raumsegmentierung durch Kalibrierung von Farb- und Tiefenbildkamera. Links: ein Objekt wird gezeigt und segmentiert (ohne Interpolation). Rechts: ein Objekt steht auf einem „Lerntisch“ und wird segmentiert (mit Interpolation)

Der Algorithmus für die Modellsynthese ist in Bild 4 dargestellt. Er besteht prinzipiell aus zwei Stufen: (1) Nutzung der segmentierten Ansichten zum Lernen von objektspezifischen Ansichten auf Basis von Merkmalspunkten [2] und Dichteschätzungsverfahren für die Merkmalspunktidentifizierung und (2) Zuordnung der Merkmalspunkte zu 3-D-Koordinaten und

Registrierung der nun vorliegenden partiellen „Merkmalspunktewolken“ des Objektes in ein gemeinsames Modell.

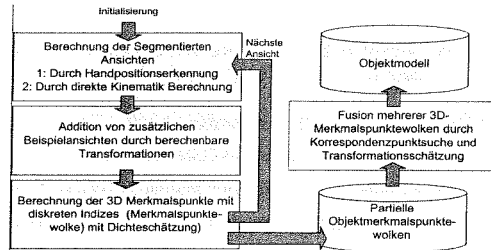


Bild 3: Der Algorithmus zur Modellsynthese (siehe Text).

Die Daten für die Segmentierung werden in den zwei vorgestellten Strategien durch einen zusätzlichen Algorithmus zur Handerkennung (Strategie 1) und durch die direkte Kinematik berechnet (Strategie 2). Dann werden künstlich erzeugte Beispiele durch Bildtransformationen für die Klassifikatoren berechnet, die hier durch Dichteschätzungsverfahren realisiert sind. Die Klassifikatoren werden auf die Beschreibungsvektoren der Merkmalspunkte angewendet. Nach Aufnahme aller Objektansichten und Generierung der zusätzlichen Ansichten werden alle Bilder mit dem Merkmalspunktdetektor bearbeitet und die Beschreibungsvektoren aller Merkmalspunkte in die Klassifikatoren gegeben, die daraufhin objektspezifisch trainiert werden. Dann wird zu jedem indizierten Merkmalspunkt die 3-D-Raumposition zugeordnet. Diese Daten bilden dann die Grundlage für die Registrierung des Objektmodells. Hier werden die Indizes der Merkmalspunkte und deren 3-D-Position dazu genutzt, die Teilansichten zusammenzufügen. Die Ergebnisse werden gerade auf speziell angefertigten 3-D-Datenbanken ermittelt. Auf konventionellen Datenbanken (z. B. COIL [3]) lieferte das Verfahren für die Merkmalspunkterkennung gute Raten (bis zu 86% Treffer, bei 100 Objekten). Das 3-D-Registrierungsverfahren wird aktuell auf speziellen 3-D-Datenbanken getestet und die Ergebnisse in dem finalen Beitrag veröffentlicht.

Themenzuordnung und Innovationsgrad

Dieser Beitrag kann am besten dem Thema „kognitive Robotik“ zugeordnet werden. Der Innovationsgrad ist nach Einschätzung des Autors hoch genug, da die Modelldatenakquisition und die Generierung von zur Erkennung geeigneter Modelle aus diesen Daten ein zentrales Problem der modernen Servicerobotik darstellt.

Literatur

- [1] Grabner, M.; Grabner, H.; Bischof, H.: Fast. Approximated SIFT, 2006
- [2] S. A. Nene, S. K. Nayar and H. Murase. Columbia Object Image Library: COIL-100, 1996.

ROBOTIK 2008

Leistungsstand - Anwendungen -
Visionen - Trends

Tagung München, 11. und 12. Juni 2008



VDI-Berichte 2012