

PERSES - ein interaktiver Einkaufsassistent *

PERSES - an interactive Shopping Assistant

Hans-Joachim Böhme, Torsten Wilhelm, Christof Schröter & Horst-Michael Groß
Technische Universität Ilmenau
Fachgebiet Neuroinformatik
PF 10 05 65, 98684 Ilmenau
{hans,wilhelm,schroeter,homi}@informatik.tu-ilmenau.de

1 Einleitung

Service-roboter sind in aller Munde, die Visionen über ihren zukünftigen Einsatz nahezu grenzenlos. Das Ziel unserer Arbeiten besteht in der Entwicklung eines interaktiven mobilen Service-roboters, der im Bau- oder Supermarkt als Informationskiosk (Shopping-Assistent) fungiert. Der Beitrag beschreibt den aktuellen Stand der Arbeiten in diesem langfristig angelegten Projekt, das gemeinsam mit der toom BauMarkt GmbH als potentiellem Endanwender solcher Systeme bearbeitet wird. Dabei liegt der Schwerpunkt zum einen auf einer robusten und sicheren Navigation im Einsatzfeld, zum anderen werden die aus aktueller Sicht notwendigen bzw. sinnvoll umsetzbaren Fähigkeiten zur intuitiven Mensch-Roboter-Interaktion vorgestellt und diskutiert.

Service robots have been becoming more and more popular over the past few years. The goal of our research is to develop an interactive mobile service robot serving as an information kiosk in home or department stores. The paper describes the current state of our project, which is joint work together with the toom BauMarkt GmbH, one of the largest german home store operators. The emphasis lies on two major aspects: first, the robust and save navigation within the operation area, and second, the ability of the service robot to interact with its users in a natural and intuitive manner.

2 Navigation

Abb. 1 fasst die wesentlichen Komponenten des entwickelten Systems zur Roboternavigation in der Baumarkt-Umgebung zusammen.

Basisrepräsentation Die Modellierung der Umgebung des Roboters durch Belegtheitsgitterkarten [9] hat sich als sinnvoll und weitgehend praktikabel erwiesen. Der ursprüngliche Nachteil dieser Art der Umgebungsrepräsentation, der im hohen Speicherplatzbedarf bei der Kartierung weiträumiger Umgebungen bei geringer Zellengröße bestand, tritt mit dem heutigen Stand der Rechentechnik mehr und mehr in den Hintergrund. Gleiches gilt für die effiziente Pfadplanung auf einer solchen Repräsentation, die für zwei diametral gelegene Positionen des Baumarktes (Dimension 100 × 60m)

*gefördert durch die Projekte PERSES und SERROKON des Thüringer Ministeriums für Wissenschaft, Forschung und Kunst

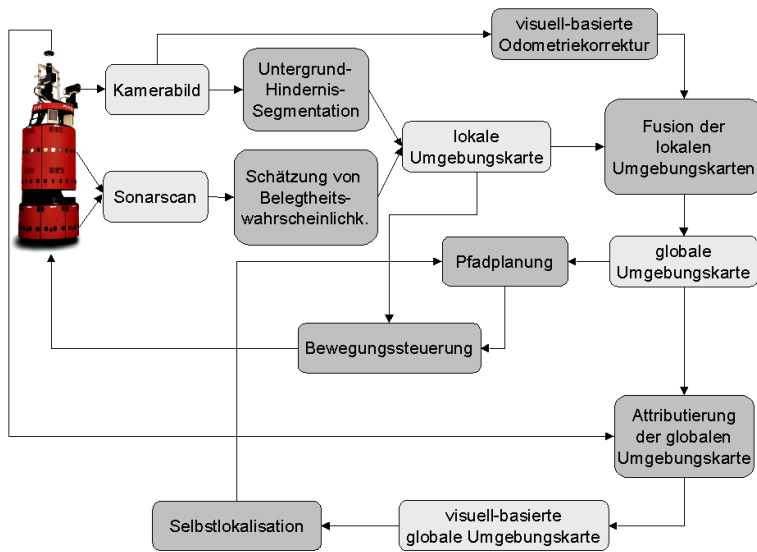


Abbildung 1: Überblick über die wesentlichen Komponenten des vorgestellten Systems zur Roboternavigation. Repräsentationen sind hellgrau, informationsverarbeitende Module dunkelgrau dargestellt.

und einer Zellengröße von 20×20 cm nur wenige Sekunden in Anspruch nimmt. Vorteilhaft ist weiterhin die Bewahrung der metrischen Informationen der Umgebung, was die Attributierung der globalen Umgebungskarte mit Artikelstandorten und Marktbereichen deutlich vereinfacht.

Lokale Umgebungskarte und Hinderniswahrnehmung Zunächst wurde ein aus der Literatur bekanntes Verfahren [10] implementiert, dass die Messwerte der Sonarsensoren mit Hilfe eines neuronalen Netzwerks auf Belegtheitswahrscheinlichkeiten der Zellen in der egozentrisch repräsentierten lokalen Umgebung des Roboters abbildet. Allerdings erscheinen in der so aufgebauten lokalen Karte nur die Hindernisse, die mit der vorliegenden Anordnung der Sonarsensoren auch tatsächlich wahrgenommen werden können. Bei alleiniger Verwendung der Sonarsensoren würden daraus starke Einschränkungen für die Umgebung resultieren, die gerade im Baumarkt mit der hohen Vielfalt an möglichen Hindernisgeometrien unmöglich eingehalten werden können. Dies führte zur Entwicklung eines visuell-basierten Verfahrens zur Hinderniswahrnehmung. Dieses beruht auf dem Prinzip der invers-perspektivischen Transformation [7, 6]. Es nutzt die Tatsache aus, dass sich der Roboter ausschließlich auf ebenem Terrain bewegt, und verwendet eine Kombination aus verschiedenen Merkmalen (optischer Fluss, Struktur, Farbe), um den mittels Kamera erfassten Bereich in Fahrtrichtung des Roboters in befahrbaren Untergrund und Hindernisse zu segmentieren. Voraussetzung für die Funktionsweise des Verfahrens ist, dass die Hindernisse genügend Struktur aufweisen, was zwar oft gegeben ist, aber auch zu prinzipiellen Einschränkungen für die robuste Funktion führt. Erste experimentelle Ergebnisse bestätigen die Eignung des Verfahrens für die reale Baumarkt-Umgebung [8]. Angestrebt wird eine kombinierte Nutzung der Sonarsensoren und der visuellen Untergrund-Hindernis-Segmentation mit dem Ziel, den Gegebenheiten der realen Einsatzumgebung besser Rechnung tragen zu können.

Globale Navigation - Kartenaufbau und Zustandsschätzung Hinsichtlich des Aufbaus konsistenter Umgebungskarten konnte ein Stand erreicht werden, der eine sichere und effiziente Handhabung gestattet. Genutzt wurde wiederum ein aus der Literatur bekannter probabilistischer Algorithmus [11], der die Belegtheitsinformation der lokalen Umgebungskarten zu einer konsistenten globalen

Umgebungskarte integrativ fusioniert. Der Einsatz einer visuell-basierten Odometriekorrektur, die auf der kamerabasierten Auswertung der Fußbodenstruktur der Einsatzumgebung beruht, sichert, dass der Odometriefehler klein genug gehalten werden kann, um eine konsistente Karte des gesamten Baumarktes aufzubauen. Der Vorteil dieser Vorgehensweise besteht aus pragmatischer Sicht darin, dass aufwendigere und konzeptionell mächtigere Algorithmen zum Kartenaufbau bei gleichzeitiger Schätzung der aktuellen Roboterpose bezüglich der gerade aufgebauten Karte (z.B. EM-Algorithmus) umgangen werden können.

Da letztlich jedoch keine Information über den tatsächlich aufgetretenen Odometriefehler vorliegt, kann mit diesem Verfahren kein Betrieb des Roboters über längere Zeit und damit längere Fahrstrecken bei hinreichend exakter Schätzung der Roboterpose realisiert werden. Deshalb wurde ein Verfahren zur visuell-basierten Posenschätzung entwickelt, das auf dem MCL-Algorithmus basiert, der als robustes und effizientes Werkzeug zur Zustandsschätzung zunehmend Verbreitung findet [3]. Dieses Verfahren wurde um den Einsatz omnidirektionaler Ansichten, die die sensorische Beobachtung des Roboters liefern, erweitert [4](siehe Abb. 2). Die experimentellen Ergebnisse zur visuell-basierten Zustandsschätzung sind soweit ermutigend, dass eine erfolgreiche Umsetzung dieses Verfahrens für das gesamte Baumarkt-Areal ohne prinzipielle methodische Probleme möglich erscheint [5].

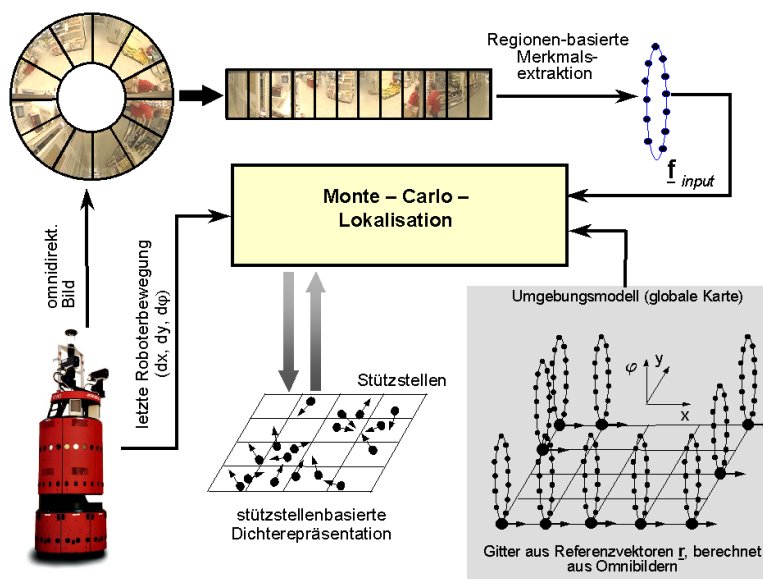


Abbildung 2: Architektur zur visuell-basierten Monte-Carlo-Lokalisation. Zugrunde gelegt wird eine Gitterrepräsentation der Umgebung, die durch die aus visuellen Panoramaansichten gewonnenen Merkmalsvektoren attribuiert wird.

Pfadplanung und Bewegungssteuerung Die eigentliche Pfadplanung beruht auf dem aus der Literatur bekannten Dijkstra-Algorithmus. Zur Gewährleistung eines ruhigen und glatten Fahrverhaltens bei sicherer Kollisionsvermeidung wurde die Hindernisvermeidung implizit in die Pfadplanung mit integriert, indem eine Abbildung des Hindernisabstandes auf die Weglänge des geplanten Pfades erfolgt. Die ursprünglich eingesetzte Potentialfeld-Methode zur Kollisionsvermeidung bzw. Bewegungssteuerung erwies sich für die vorliegende Einsatzumgebung als nur bedingt geeignet. Mit dem implementierten alternativen Verfahren auf der Basis von Vektorfeld-Histogrammen [2] konnten die Probleme der Potentialfeld-Methode überwunden werden, so dass letztlich ein effizientes Verfahren zur Bewegungssteuerung vorliegt.

3 Mensch-Roboter-Interaktion

Für die Akzeptanz von Servicerobotern, die mit beliebigen Personen interagieren sollen, wird der Aspekt der einfachen Benutzung in Verbindung mit einem möglichst intuitiven und natürlichen Gesamtverhalten solcher Systeme zukünftig eine zentrale Rolle spielen. Je besser es gelingt, den Roboter mit Verhaltensweisen auszustatten, die der Benutzer auch von einem menschlichen Interaktionspartner erwarten würde, desto weniger werden Berührungsängste und Zweifel an der Sinnfälligkeit dieser Systeme in den Vordergrund gelangen. Genau diese Zielstellung wurde mit der Entwicklung der in diesem Abschnitt vorgestellten Verhaltensmodule verfolgt. Abb. 3 skizziert den Ablauf eines Interaktionszyklus, wie er vom Shopping-Assistenten zum gegenwärtigen Zeitpunkt realisiert und demonstriert werden kann [1].

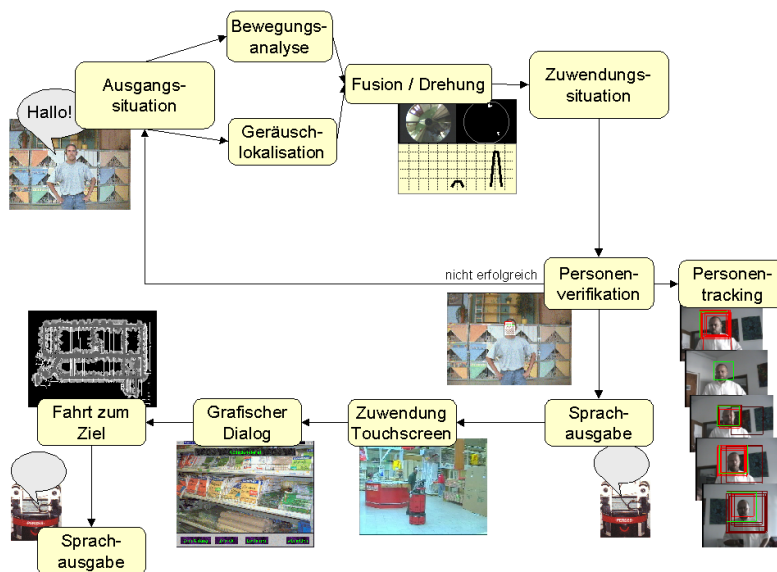


Abbildung 3: Überblick über den derzeit realisierten Interaktionszyklus zwischen dem Shopping-Assistenten PERSES und seinem Benutzer.

Aufmerksamkeitssteuerung Die Aufmerksamkeitssteuerung generiert Hypothesen über vorhandene Personen in der unmittelbaren Umgebung des Shopping-Assistenten. In einem multimodalen Ansatz wurden sowohl visuell-basierte als auch akustische Verfahren zu einem Gesamtsystem kombiniert. Diese Multimodalität einerseits und die Auswertung der Bewegungsrichtung der detektierten Bildregionen andererseits sichern ein natürliches und intuitiv verständliches Zuwendungsverhalten des Roboters. Personen in der Umgebung des Roboters können somit beobachten, dass sie vom Roboter aktiv wahrgenommen werden.

Personenverifikation Um festzustellen, ob sich in der durch die Aufmerksamkeitssteuerung vorgegebenen Richtung tatsächlich ein potentieller menschlicher Interaktionspartner befindet, wird eine anschließende Personenverifikation durchgeführt. Diese nutzt ausschließlich den visuellen Kanal und basiert auf der Verwendung verschiedener personentypischer Cues (Hautfarbe, Gesichtsstruktur, Kopf-Schulter-Silhouette), die in einem dreidimensionalen Feld dynamischer Neuronen so fusioniert werden können, dass eine sehr robuste Verifikation in einem definierten Entfernungsbereich erfolgen kann. Das Hauptaugenmerk bei der Entwicklung dieses Moduls lag auf der Vermeidung von falsch-positiven Verifikationen, um eine Kontaktaufnahme des Roboters mit unbelebten Objekten möglichst

auszuschließen.

Personentracking Für die Auswahl der Methode zum Personentracking war maßgeblich, dass das Verfahren in der Lage ist, auch multiple Hypothesen über die momentane Position des Objekts zu behandeln. Eingesetzt wird wiederum der Condensation-Algorithmus, der bereits für die visuell-basierte Zustandsschätzung Verwendung fand. Für das visuelle Personentracking kommen die Merkmale (Haut)farbe und Kopf-Schulter-Silhouette zum Einsatz, die auch für die Personenverifikation verwendet wurden. Durch die Hinzunahme von sonarbasierter Abstandsinformation in Richtung der zu verfolgenden Person konnte eine erhebliche Erhöhung der Robustheit des Trackingverfahrens erreicht werden, was den Einsatz multimodaler Trackingsysteme für reale Einsatzszenarien im allgemeinen motiviert (siehe [12]).

Klassische Interaktionsformen Alle Bemühungen in Richtung natürlicher und intuitiv verständlicher Mensch-Roboter-Interaktion dürfen nicht darüber hinwegtäuschen, dass an eine vollständige Substitution klassischer Eingabegeräte wie Tastatur oder Touchscreen in absehbarer Zeit noch nicht zu denken ist. Unser Ansatz versucht vielmehr, nach wie vor notwendige klassische Medien zur unmittelbaren MRI in ein intuitiv und natürlich wirkendes Gesamtverhalten einzubetten. Das entwickelte grafische Interaktionstool hat derzeit noch eher Demonstrationscharakter, kann aber dennoch als Prototyp für die unmittelbare Nutzer-Roboter-Interaktion im realen Einsatzfall dienen.

Roboterartikulation Mit dem Einsatz eines Moduls zur Sprachausgabe wird versucht, den aktuellen Interaktionsstatus des Roboters gegenüber seinem menschlichen Interaktionspartner möglichst einfach zu artikulieren. Gegenüber grafischen Ausgaben auf dem Display bietet dies den Vorteil, dass die Äußerungen des Roboters weitgehend unabhängig von der aktuellen Position des Benutzers wahrgenommen werden können.

Ein Robotergesicht dient neben der Vermittlung einer „smarten“ Gesamterscheinung des Roboters auch dazu, einfach verständliche „Emotionen“ des Roboters auf natürliche Weise zu verdeutlichen.

4 Zusammenfassung und Ausblick

Der vorliegende Beitrag beschreibt den aktuellen Realisierungsstand bei der Entwicklung eines interaktiven mobilen Shopping-Assistenten als Prototyp für eine Klasse von Servicerobotern, die als Informationskioske in verschiedenen Applikationen zum Einsatz kommen können. Ein Schwerpunkt liegt dabei auf der Nutzung visuell-basierter Verfahren, die der Komplexität der Einsatzumgebung Rechnung tragen und in Kombination mit weiteren Sensormodalitäten (ultraschall-basiert, akustisch) sinnvoll einsetzbar sind und zu einer höheren Robustheit des Gesamtsystems beitragen.

Trotz aller Fortschritte bezüglich der angestrebten robusten Hinderniswahrnehmung konnte dieser Gesamtkomplex noch nicht so weit entwickelt werden, dass ein unbeaufsichtigtes Agieren des Shopping-Assistenten im Baumarkt sichergestellt werden kann. Dies liegt vorrangig in der noch fehlenden Integration des visuellen Verfahrens zur Hinderniswahrnehmung begründet, die Gegenstand aktueller Arbeiten ist, so dass mit experimentellen Ergebnissen zur Validierung des Gesamtsystems zur Hinderniswahrnehmung in den kommenden Monaten gerechnet werden kann. Weiterhin wird zum gegenwärtigen Zeitpunkt versucht, das Verfahren zur visuellen MCL auf den gesamten Baumarkt zu skalieren.

Bezüglich MRI konnten mit den entwickelten Methoden die informationstechnischen Voraussetzungen für eine möglichst intuitiv und natürlich wirkende Interaktion geschaffen werden. Schwerpunkte der künftigen Arbeiten werden hier die Weiterentwicklung der Methoden zum Personentracking sowie zur Dialogmodellierung bilden. Mit Ausnahme des grafischen Interaktionstools, das sich stark an den Anforderungen der vorliegenden Applikation Shopping-Assistent im Baumarkt orientiert, lassen sich alle dargestellten Verfahren prinzipiell auf eine Vielzahl weiterer Anwendungen von Servicerobotern übertragen.

Literatur

- [1] Boehme, H.-J., Wilhelm, T., Key, J., Schroeter, Ch., Hempel, T., and Gross, H.-M. An Approach to Multimodal Human-Machine Interaktion for Intelligent Service Robots. In *EURO-BOT'01 - the fourth Euromicro Workshop on Advanced Mobile Robots*, volume 86 of *Lund University Cognitive Studies*, pages 17–24. Lund University, 2001.
- [2] Borenstein, J. and Koren, Y. The Vector Field Histogram – Fast Obstacle-Avoidance for Mobile Robots. *IEEE Journal of Robotics and Automation*, 7(3):278–288, 1991.
- [3] Fox, D., Delleart, F., Burgard, W., and Thrun, S. Monte Carlo Localization: Efficient Position Estimation for Mobile Robots. In *Proceedings 16th National Conference on Artificial Intelligence (AAAI-99)*, 1999.
- [4] Gross, H.-M. and Boehme, H.-J. PERSES - a Vision-based Interactive Mobile Shopping Assistant. In *IEEE Internat. Conference on Systems, Man and Cybernetics (IEEE-SMC 2000)*, pages 80–85. IEEE/Omnipress, 2000.
- [5] König, A., Gross, H.-M., and Böhme, H.-J. Visuell basierte Monte-Carlo Lokalisation für mobile Roboter mit omnidirektionalen Kameras. Technical report, Technische Universität Ilmenau, Schriftenreihe des Fachgebiets Neuroinformatik, 2001. Report 2/01, ISSN 0945-7518.
- [6] Krabbes, M., Weber, S., Boehme, H.-J., and Gross, H.-M. Monokulare visuelle Hindernisdetektion auf Basis merkmalsbasierter Bildsegmentierung. In *AMS'98 - 14. Fachgespräch Autonome Mobile Systeme*, Informatik aktuell, pages 85–92. Springer-Verlag, 1998.
- [7] Mallot, H.P., Bühlhoff, H.H., Little, J.J., and Bohrer, S. Inverse Perspective Mapping Simplifies Optical Flow Computation and Obstacle Detection. *Biological Cybernetics*, 64:177–185, 1991.
- [8] Martin, C. and Böhme, H.-J. Monokulare Roboternavigation zur visuellen Hindernisvermeidung. Technical report, Technische Universität Ilmenau, Schriftenreihe des Fachgebiets Neuroinformatik, 2001. Report 1/01, ISSN 0945-7518.
- [9] Moravec, H.P. and Elfes, A. High resolution maps from wide angle sonar. In *International Conference on Robotics and Automation (ICRA)*, pages 116–121, 1985.
- [10] Thrun, S. Exploration and Model Building in Mobile Robot Domains. In *Proceedings of the IEEE International Conference on Neural Networks*, pages 175–180, 1993.
- [11] Thrun, S. Learning Maps for Indoor Mobile Robot Navigation. *Artificial Intelligence*, 99(1):21–71, 1999.
- [12] Wilhelm, T., Böhme, H.-J., and Groß, H.-M. Wechselseitige Unterstützung von sonarbasierendem und visuellem Personentracking. In *Robotik 2002*. VDI Verlag, 2002.