

Selbstorganisierende Klassifikation von 2D-Laserscans zur Navigation eines AMR

Joachim Weber, Alexander Spieß, Klaus-Werner Jörg, Ewald von Puttkamer

Fachbereich Informatik • Arbeitsgruppe Robotik & Prozessrechentchnik
Universität Kaiserslautern, Postfach 3049, D-67653 Kaiserslautern
email: {jweber, puttkam}@informatik.uni-kl.de, a_spiess@gmx.de, kw.joerg@androtec.de

Zusammenfassung. Der vorliegende Beitrag stellt einen neuronalen Interpretationsansatz für Laserscans vor, der eine zur Klassifikation von Kamerabildern entwickelte Technik einsetzt. Besonderes Merkmal dieses Ansatzes ist die Fähigkeit, sich in einer unüberwachten Lernphase selbstständig auf die besonderen Merkmale der Einsatzumgebung einzustellen. Die Zuordnung der verarbeiteten Laserscans zu eigenständig ausgebildeten Szenenklassen legt den Einsatz dieser Technik in topologischen Navigationsverfahren nahe.

1 Einleitung

Für die Wiedererkennung von Positionen eines Serviceroboters in der Einsatzumgebung mittels Laser-Entfernungsbildern werden derzeit eine Reihe von Techniken eingesetzt, die sich vorwiegend auf die Extraktion bestimmter Merkmale stützen. Dies macht die explizite oder implizite Modellierung relevanter Umgebungseigenschaften notwendig. So werden bei Cox [1] ausschließlich Linien extrahiert und mit einer vorhandenen Karte verglichen. Der von Weiß & v. Puttkamer [2] vorgestellte Ansatz korreliert X- und Y-Histogramme winkelnormalisierter Laserscans, was dieses Verfahren von einer vorwiegend polygonalen Einsatzumgebung abhängig macht. Auch in [3] müssen für verschiedene Typen charakteristischer Umgebungsmerkmale gewisse Modellparameter angegeben werden, da Relationen zwischen Merkmalspositionen zur Suche passender Referenzscans genutzt werden. Diese und andere landmarkenbasierte Verfahren versagen, sobald die implizit oder explizit verwendeten, a priori definierten Merkmalsmodelle nicht in ausreichendem Maße in den Laserscans repräsentiert sind.

Die wenigen, auf der Verwendung reiner Rohdaten basierenden Verfahren weisen andere Einschränkungen auf, z. B. das notwendige Vorhandensein ungefährer Positionsannahmen bei dem iterativen Algorithmus von Lu & Milios [4]. Die auf Hauptkomponentenanalyse basierende Technik bei Crowley, Wallner & Schiele [5] setzt hingegen ein umfangreiches Scan-Trainingsset von exakt bekannten Positionen voraus, während die in [6] von Fox et al. vorgestellte Monte-Carlo-Lokalisation künstliche Samples an zufälligen Positionen auswertet. Damit basieren die beiden letztgenannten Verfahren auf metrisch korrekten Weltmodellen, welche die Möglichkeit zur *absoluten* Lokalisierung bieten, d. h. die genaue Positionsangabe von (x, y, ϕ) in ei-

nem globalen Weltkoordinatensystem. Allerdings ist der für die korrekte Erzeugung und Konsistenzerhaltung der metrischen Karte notwendige Aufwand in vielen Anwendungsbereichen der Service-Robotik unnötig und unerwünscht.

Wenn nicht zu jedem Zeitpunkt eine absolute Lokalisierung notwendig ist, bildet topologische Navigation eine Alternative zur Verwendung metrischer Weltmodelle. In diesem Fall können Techniken eingesetzt werden, die Gebiete der Einsatzumgebung wiedererkennen können, ohne den aufwändigen Prozess der exakten relativen Positionsbestimmung durchführen zu müssen.

Gegenstand dieses Beitrags ist ein Verfahren, das eine selbstorganisierende Wiedererkennung von Positionen durch die Klassifikation von Laserscans durchführt. Dabei zeichnet sich diese ursprünglich zur Verarbeitung von Videobildern entwickelte Technik durch die Selbstadaption an die signifikanten Merkmale der Einsatzumgebung aus.

Abschnitt 2 behandelt das Verfahren, während Abschnitt 3 experimentelle Ergebnisse vorstellt. In Abschnitt 4 wird auf den Einsatz der Klassifikation in Selbstlokalisationsarchitekturen eingegangen und Abschnitt 5 bietet einige Schlussbetrachtungen.

2 Selbstorganisierende Klassifikation

In [7] stellen v. Wichert & Kleiner ein Verfahren für die unüberwacht und selbstorganisierend lernende visuelle Szenenerkennung vor. Das Verfahren ordnet ähnliche Kamerabilder durch eine hierarchisch strukturierte neuronale Bildanalyse entsprechenden Klassen zu, wobei weder die Anzahl der Bildklassen, noch die von einer Klasse repräsentierten Merkmalskombinationen a priori vom Benutzer vorgegeben wird. Die Besonderheit dieses Ansatzes liegt also in der selbstständigen Anpassung an die besonderen Eigenschaften der Einsatzumgebung, was die Notwendigkeit der manuellen Modellierung relevanter Umgebungsaspekte beseitigt. Im ersten Verarbeitungsschritt wird für jedes Videobild eine n-kanalige Merkmalsfilterung durchgeführt, z. B. durch Texturfilter, Kantenfilter, usw.. Zur Stabilisierung der Bilddarstellung werden die n resultierenden Merkmalsbilder einer pixelweisen Vektorquantisierung durch eine selbstorganisierende Merkmalskarte unterzogen. Die unterschiedlichen Merkmalsbilder werden also in diesem Schritt auf *ein* diskretes Klassenbild reduziert. Die eigentliche Zuordnung des Videobildes zu Klassen ähnlicher Bilder übernimmt ein weiteres neuronales Netz nach Überführung des Klassenbilds in eine vektorielle Darstellung. Beide selbstorganisierende Merkmalskarten werden in einem Offline-Trainingsschritt mit typischen Kameraansichten der Einsatzumgebung trainiert.

Diese Selbstadaptivität an die Charakteristik und das Erscheinungsbild der Einsatzumgebung ist eine für die Interpretation auch andersgearteter Sensordaten sehr nützliche Eigenschaft. Dies wird nachfolgend anhand der Klassifikation von 180°-Laserscans gezeigt. Analog zu [7] werden die Laserscans gemäß dem in Abbildung 1 gezeigten Ablauf verarbeitet. In einem ersten Schritt wird der Laserscan einer n-kanaligen Merkmalsfilterung unterzogen, d. h. jedem Entfernungswert (*Pixel*) p_i des Laserscans wird ein n-dimensionaler Vektor zugeordnet, dessen Komponenten jeweils die Ausgabewerte der zugehörigen Filter für das Pixel p_i repräsentieren. Filteropera-

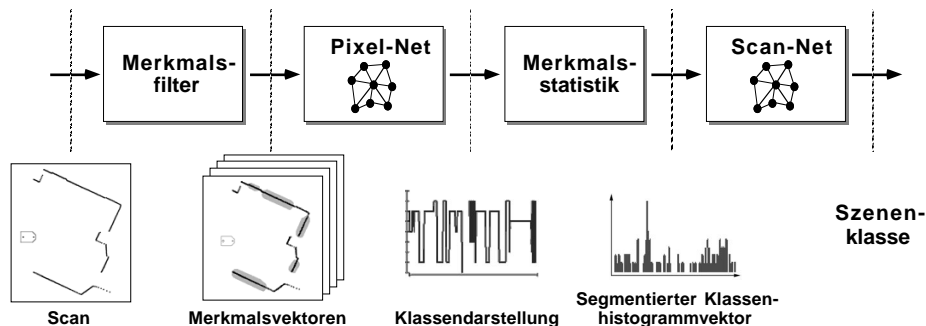


Abbildung 1: Ablaufdiagramm der Scanklassifikation

tionen betreffen z. B. die Extraktion von Kanten, Linien, etc.

Aus dieser vektoriiellen Darstellung wird mittels des *Pixel-Nets* eine diskrete Klassendarstellung des Laserscans erzeugt, indem das Netz dem Merkmalsvektor jedes Pixels eine dafür repräsentative Klasse zuweist. Das *Pixel-Net* - bei v. Wichert ein Kohonennetz, in der vorliegenden Arbeit und in [8] ein *Growing Neural Gas Net* [9] - ist zuvor in einer ausgezeichneten Offline-Trainingsphase mit den Merkmalsvektoren typischer Laserscans der Einsatzumgebung trainiert worden und hat dabei eigenständig die Klassen auftretender Merkmalskombinationen gebildet. In einfachen Umgebungen könnten dies z. B. Türkanten, rechtwinklige Ecken o. Ä. sein, oder auch für einen menschlichen Betrachter weniger auffällige Merkmalskombinationen. Dies stellt den angesprochenen Aspekt der selbstorganisierenden Adaption an die Eigenheiten der in der Einsatzumgebung auftretenden Objektmerkmale dar.

Die Wahl des Filtersatzes ist hingegen eine manuell vorzunehmende Einstellung auf die Umgebungscharakteristik. Wie aus den Ergebnissen in Abschnitt 3 ersichtlich ist, lassen sich aber bereits mit rudimentären Filtern gute Ergebnisse erzielen und die Parametrierung ist unkritisch.

Die Klassendarstellung eines Laserscans ist wie das Klassenbild in [7] stabil gegenüber mäßiger Translation und Rotation, d. h. die Form von Bereichen einer Pixelklasse und die relative Lage zu Bereichen anderer Pixelklassen bleibt nahezu unverändert. Insbesondere wird sich diese Darstellung, wie bei v. Wichert ausgeführt, bei stetiger Bewegung des Roboters ebenfalls nur stetig ändern, was für die abschließende Szenenklassifikation von besonderer Bedeutung ist.

Ziel dieser Klassifikation ist die Zuordnung der aktuellen Sensordaten zu typischen, i. Allg. allerdings nicht eindeutigen Situationen, z. B. Korridor, Raum mit runden Säulen, 'Blick' durch eine Tür in ein kleines Zimmer, usw.. Diese Zuordnungen sollen dabei invariant gegenüber kleineren Änderungen der Position und Orientierung sein und genau wie bei der Festlegung der typischen Merkmalskombinationen im *Pixel-Net* selbstständig von dem Verfahren ausgebildet, also nicht explizit vorgegeben werden.

Die Klassifikation der gescannten Szene wird durch das so genannte *Scan-Net* durchgeführt, ein Gasnetz, das ebenfalls in einer Offline-Trainingsphase unüberwacht mit den Klassendarstellungen von Laserscans der Einsatzumgebung trainiert wurde.

Analog zu [7] werden in der vorliegenden Arbeit aus der Klassendarstellung der Laserscans Histogrammvektoren - also Häufigkeitsanalysen der aufgetretenen Pixel-Net-Klassen - erzeugt und durch das Scan-Net quantifiziert. Um diesem abschließenden Vektorquantisierer die Erkennung der *relativen* räumlichen Merkmalsbeziehungen zu ermöglichen, wird dies wie bei v. Wichert durch die Erzeugung separater Histogrammvektoren für unterschiedliche, sich überlappende Bereiche des Eingaberaums gewährleistet.

3 Klassifikationsergebnisse

Die im folgenden vorgestellten Ergebnisse wurden in Offline-Verarbeitungsschritten mit aufgezeichneten Versuchsfahrten des Roboters *Phoenix* erzielt. Eine erste Aufzeichnung diente als Datenbasis für das Training von Pixel- und Scan-Netz mit 4 einfachen Filtern: Linien-, Abriss-, Ecken- und Distanzfilter.

Linien- und Abrissfilter sind rein binäre Filter, d.h. sie haben als Ausgabewerte lediglich 1 und 0. Der Linienfilter spricht in den Bereichen eines Laserscans an, in denen die Oberflächenwinkel eines gescannten Objekts einen konstanten Wert aufweisen, der Abrissfilter an Stellen mit großen Entfernungssprüngen. Der reellwertige Eckenfilter liefert an den Stellen zwischen benachbarten Linien unterschiedlicher Ausrichtung die entsprechende Winkeldifferenz, ansonsten den Wert 0. Der Distanzfilter realisiert die Identität, d. h. die Entfernungswerte des Scans werden ungefiltert durchgereicht. Dies ermöglicht der Scanklassifikation die Beachtung von Größenordnungen der erfassten Szene.

Die Ansprechbereiche des nur punktuell wirksamen Abriss- und Eckenfilters wurden dabei künstlich auf einen definierten Nachbarbereich ausgedehnt, um die Erzeugung von Merkmalskombinationen mit dem ansonsten dazu disjunkten Linienfilter zuzulassen. Im *Pixel-Net* wurden für diese Trainingsfahrt selbstorganisierend 8 relevante Merkmalsklassen (Neuronen) ausgebildet. So repräsentiert beispielsweise Klasse 0 90°-Grad-Ecken in mittlerer Entfernung, Klasse 1 Abrisskanten in kurzer Entfernung, usw..

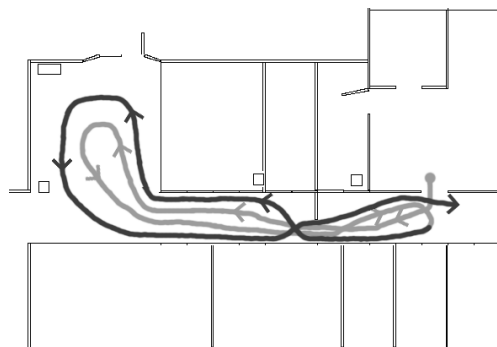


Abbildung 2: Roboterfahrstrecke: 1. Teilfahrt hellgrau, Wiederholung dunkelgrau

Um den Einfluss der Neuronenanzahl im *Scan-Net* auf die Stabilität der Wiedererkennung zu überprüfen, wurden für das *Scan-Net* verschieden parametrisierte Trainingsläufe mit einer unterschiedlichen Anzahl resultierender Neuronen durchgeführt. Für eine robuste Selbstlokalisierung ist primär die Reproduzierbarkeit der Klassifizierung von Bedeutung. Dementsprechend wurde zur Überprüfung der Wiedererkennungszuverlässigkeit in einer zweiten Aufzeichnungsfahrt die gleiche Strecke zweimal mit zum Teil erheblichem Versatz befahren (siehe Abbildung 2; 1. Teilfahrt: hellgrau, 2. Teilfahrt: dunkelgrau). Die Klassifikationsergebnisse dieser Versuchsfahrt sind in Abbildung 3 für die verschiedenen *Scan-Net*-Varianten dargestellt. Dabei sind auf der x-Achse die insgesamt 1129 Scans der gesamten Doppelfahrt, auf der y-Achse die Neuronen des jeweiligen *Scan-Netzes* aufgetragen. Für jeden Scan ist durch einen Punkt gekennzeichnet, welches der Neuronen von ihm aktiviert wurde. Die vertikale Trennung in der Mitte der X-Achse kennzeichnet den Beginn der Fahrtwiederholung. Ähnliche Aktivierungssequenzen in der 1. und 2. Teilfahrt wurden jeweils flächig zur besseren Sichtbarkeit markiert.

Das Klassifikationsergebnis macht im wesentlichen drei Sachverhalte deutlich:

- Die Zuordnung von Position und Klasse ist erwartungsgemäß nicht eindeutig - gleiche Klassen erscheinen an verschiedenen Positionen.
- Die Klassifikation weist die geforderte Stabilität über geringe Änderungen von Position und Orientierung auf; es lassen sich *Cluster* von gleichen Klassenzuordnungen bei benachbarten Scans detektieren.
- Für die Fahrtwiederholung lassen sich für alle *Scan-Net*-Ausprägungen korrespondierende Clustersequenzen zur 1. Teilfahrt beobachten.

Für die in Abschnitt 4 besprochene Verwendung der Szenenklassifikation zur topologischen Navigation sind die beiden letzten Eigenschaften besonders bedeutsam, da sie die für eine robuste Navigation benötigte Stabilität der Szenenklassifikation nachweisen.

In Bezug auf die verschiedenen *Scan-Net*-Varianten zeigt sich, dass die Parametrierung des Trainings unkritisch ist, da sich für alle drei Varianten korrespondierende Aktivierungssequenzen in den beiden Teilfahrten identifizieren lassen. Allerdings nimmt aufgrund der feineren Diskretisierung mit steigender Neuronenanzahl natürlich die Rotations- und Positionsempfindlichkeit zu, was bei stärkeren Abweichungen die Wiedererkennung erschwert. So wurden z. B. in der hochauflösendsten Variante des *Scan-Nets* (32 Neuronen) einige Neuronen durch die Klassifikation der Testfahrt überhaupt nicht aktiviert.

4 Topologische Navigation

Metrikbasierte Navigationsarchitekturen weisen für den Einsatz entfernungsmessender Sensorik, z. B. Laserscannern oder Ultraschallsensoren, neben der sehr anschaulichen Umweltrepräsentation eine Reihe weiterer Vorzüge auf:

- Die explizite Hindernismodellierung aller statischen Objekte erlaubt den Einsatz einer vorausschauenden Bewegungsplanung des Roboters. Dies ist für viele Fahrzeugkinematiken in engen Umgebungen notwendig, um ein 'Festfahren' des Roboters in Engpässen zu vermeiden.

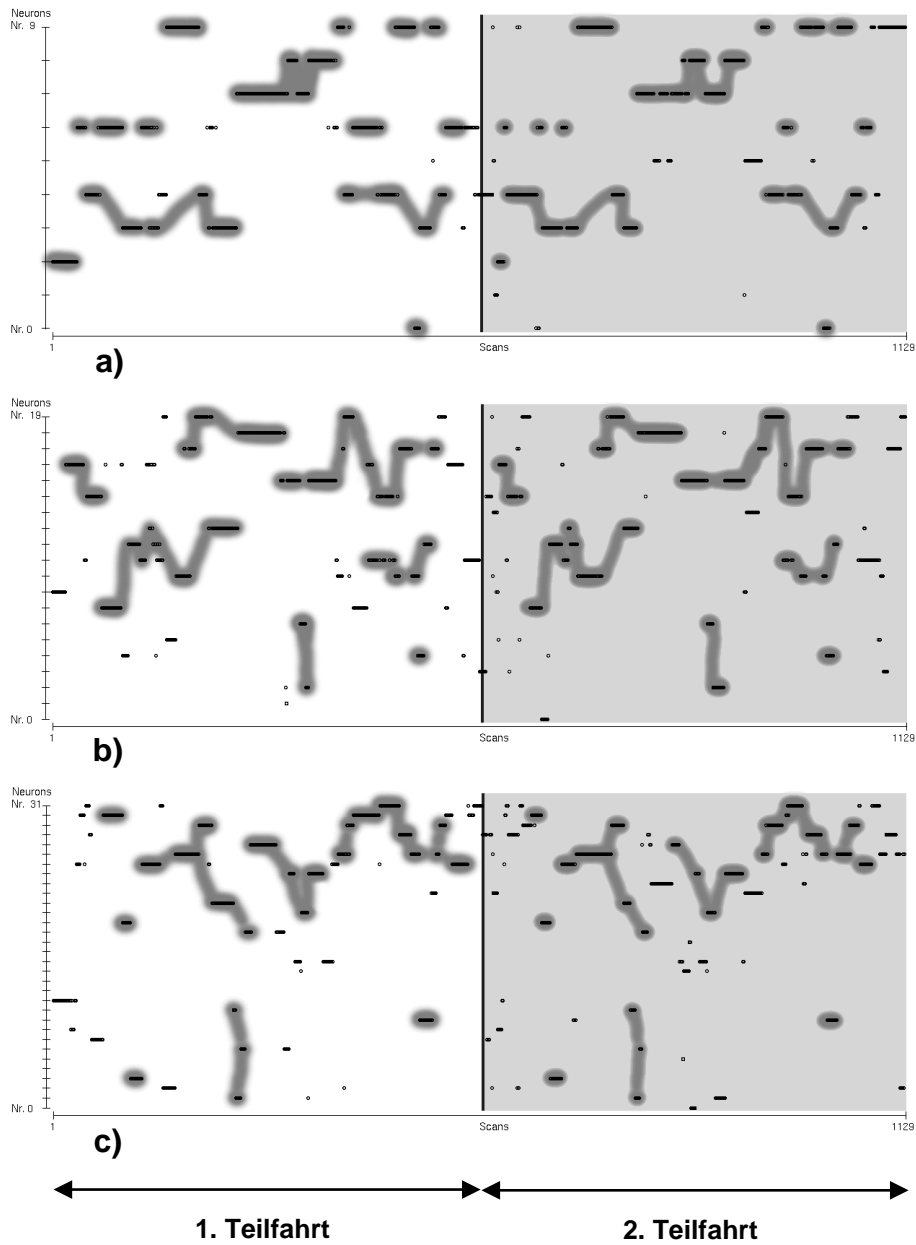


Abbildung 3: Klassifikationsergebnis der Testfahrt mit 10 (a), 20 (b) und 32 (c) Scan-Net-Neuronen für das gleiche Pixel-Net. Ähnlichkeiten zwischen 1. und 2. Teilfahrt sind zur besseren Sichtbarkeit flächig hervorgehoben.

- In vielen Anwendungsbereichen ist die Kenntnis der absoluten Roboterposition und -orientierung unerlässlich, beispielsweise bei fahrerlosen Transportsystemen, die sich lediglich auf vorgegebenen Bahnen bewegen dürfen. Auf metrischen Karten basierende Selbstlokalisierungstechniken, wie z. B. [4, 5, 6], erlauben die absolute Lokalisierung mit einer Präzision im Bereich des Modellierungsfehlers der Karte.

Der wesentliche Nachteil metrischer Weltmodelle liegt in dem zur konsistenten Kartenerstellung notwendigen Aufwand. Auch erlauben sie nicht ohne weiteres die Integration nichtmetrischer Sensordaten, wie z. B. von visuellen Klassifikationsergebnissen gemäß [7].

Besteht nicht die Notwendigkeit der absoluten Lokalisierung, bietet sich als Alternative die Verwendung topologischer Weltmodelle an. Diese repräsentieren die Umgebung als einen Graphen, in dem die Knoten charakteristische Stellen der Einsatzumgebung und die Kanten befahrbare Verbindungswege zwischen diesen Positionen repräsentieren. Sensordaten, wie Laser- oder Ultraschallscans, Kamerabilder oder relative gemessene Wegstrecken, werden den Knoten und Kanten als Attribute zugewiesen. Damit eignen sich topologische Weltmodelle sehr gut für die Integration verschiedenster Sensorik und reduzieren die globale Wegeplanung auf die Wegsuche im Graphen. Für den Kartenaufbau ist i. Allg. nur die topologische Korrektheit gefordert, die leichter zu gewährleisten ist, als das in Position und Winkel präzise Aneinanderfügen von Sensordaten bei metrischer Kartierung.

Die Ergebnisse der hier vorgestellten laserbasierten Klassifizierung könnten direkt in probabilistische Lokalisationsarchitekturen wie [10, 12] zur Identifikation von Knotenumgebungen integriert werden. Der Einsatz wahrscheinlichkeitsbasierter Verfahren ist dabei insbesondere aufgrund der Nichteindeutigkeit der Klassifikationsergebnisse erforderlich, aber auch zur Kompensation von Fehlklassifikationen wegen dynamischer Effekte in der Einsatzumgebung.

Die Scanklassifikation wird zur Zeit in den in [11] vorgestellten probabilistischen Ansatz zur Selbstlokalisierung integriert. Dabei dient das Referenzscan-Matching des APR-Algorithmus [3] zur Erzeugung von Positionshypothesen in einem topologischen Weltmodell, das um metrische Information erweitert wurde. Zur besseren Beurteilung der jeweiligen Hypothesenwahrscheinlichkeiten ist es sinnvoll, Bestätigung durch die Ergebnisse anderer Methoden zu suchen. Dies ist besonders in solchen Teilen der Einsatzumgebung wichtig, in denen das landmarkenbasierte APR aufgrund des Fehlens der gesuchten, benutzerspezifischen Merkmale keine Ergebnisse liefern kann und Positionshypothesen nur mittels Odometrie fortgeschrieben werden.

Da der topologische Graph aus [11] eine vergleichsweise geringe Knotendichte besitzt - APR benötigt relativ wenige Referenzscans - ist hinsichtlich der hier vorgestellten Scanklassifikation eine feinere Diskretisierung im Graphen notwendig. Dazu wird die Aktivierungssequenz auf dem Weg zwischen zwei Knoten als gerichtetes Profil in der Verbindungskante gespeichert. Der Vergleich des beobachteten Aktivierungsverlaufs mit dem gespeicherten modifiziert als zusätzliche Informationsquelle die Berechnung der Hypothesenwahrscheinlichkeiten.

5 Zusammenfassung

Die Ergebnisse der vorliegenden Arbeit bestätigen die Leistungsfähigkeit des in [7] vorgestellten Ansatzes zur hierarchischen, selbstorganisierenden Verarbeitung von Sensordaten und zeigt seine Anwendbarkeit auf unterschiedliche Sensordomänen. Die Klassifikationsergebnisse von Laserscans weisen die zur topologischen Navigation erforderliche statistische Stabilität nach, wobei sich das System selbstständig auf die Merkmale der Umgebung einstellt. Die entfallende Notwendigkeit manueller Modellierung relevanter Umgebungsaspekte trägt erheblich zur Flexibilisierung, und damit Robustheit des Einsatzes von Servicerobotern in vielfältigen Einsatzumgebungen bei. Die Versuche zur vorliegenden Arbeit wurden im Rahmen des CAROL-Forschungsprojekts auf der mobilen Versuchsplattform *Phoenix* durchgeführt.

Literatur

- [1] I. J. Cox: *Blanche - An Experiment in Guidance and Navigation of an Autonomous Robot Vehicle*, IEEE Transactions on Robotics and Automation, Vol 7 (2), pp. 193-204, 1991
- [2] G. Weiß, E. v. Puttkamer: *A Map Based on Laserscans Without Geometric Interpretation*, Proceedings Intelligent Autonomous Systems (IAS-4), 1995
- [3] J. Weber, K.-W. Jörg, E. v. Puttkamer: *APR - Global Scan Matching Using Anchor Point Relationships*, Proceedings Intelligent Autonomous Systems (IAS-6), pp. 471-478, 2000
- [4] F. Lu, E. E. Milios: *Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans*, Proceedings IEEE Conference on Computer Vision and Pattern Recognition, pp. 935-938, 1994
- [5] J. L. Crowley, F. Wallner, B. Schiele: *Position Estimation Using Principal Components of Range Data*, Proceedings 1998 IEEE International Conference on Robotics & Automation, pp. 3121-3128, 1998
- [6] D. Fox, W. Burgard, F. Dellaert, S. Thrun: *Monte Carlo Localization: Efficient Position Estimation for Mobile Robots*, Proceedings 16th National Conference on Artificial Intelligence (AAAI-99), 1999
- [7] G. v. Wichert, K. Kleiner: *Selbstorganisierende Bildanalyse für die Navigation autonomer mobiler Roboter*, Autonome Mobile Systeme (AMS'95), Springer, pp. 152-161, 1995
- [8] H. Klein: *Ein Verfahren zur Szenenklassifikation anhand von Videobildern*, Diplomarbeit im Fachbereich Informatik, Universität Kaiserslautern, 1996
- [9] B. Fritzke: *A growing neural gas network learns topologies*, Advances in Neural Information Processing Systems, 7, 1995
- [10] R. Simmons, S. Koenig: *Probabilistic Navigation in Partially Observable Environments*, Proceedings International Joint Conference on Artificial Intelligence, 1995
- [11] J. Weber, L. Franken, K.-W. Jörg, E. v. Puttkamer: *Globale Selbstlokalisierung mittels Referenzscan-Matching*, Autonome Mobile Systeme (AMS'2000), Springer, 2000
- [12] D. Kortenkamp, T. Weymouth: *Topological mapping for mobile robots using a combination of sonar and vision sensing*, Proceedings 12th National Conference on Artificial Intelligence (AAAI-94), pp. 979-984, 1994