

Maschinelles Sehen zur Steuerung autonomer mobiler Systeme

Volker Graefe und Rainer Bischoff

Universität der Bundeswehr München
Institut für Meßtechnik
Werner-Heisenberg-Weg 39, 85577 Neubiberg
E-Mail: {Graefe | Rainer.Bischoff}@unibw-muenchen.de

Kurzfassung

Die Erforschung und Realisierung „sehender“ Maschinen und „intelligenter“ Roboter bildet seit 1977 einen Arbeitsschwerpunkt des Instituts für Meßtechnik. Ziel ist zum einen ein grundlegendes Verständnis des Sehens, der Autonomie und der Intelligenz von technischen Systemen, und zum anderen die Realisierung intelligenter sehender Roboter. In diesem Beitrag beschreiben wir drei unserer autonomen mobilen Systeme, die das maschinelle Sehen als Hauptsensormodalität einsetzen, um sich in ihrer jeweiligen Umgebung mit angemessener Geschwindigkeit fortbewegen und selbständig vielfältige Aufgaben erfüllen zu können. Mit ihrer Hilfe haben wir u.a. die Problem-bereiche der autonomen Mobilität im öffentlichen Straßenverkehr, die Navigation mobiler Roboter in bekannten und unbekanntem Umgebungen, das maschinelle Lernen und die Steuerung von Robotern ohne Kenntnis quantitativer Modelle untersucht. Im folgenden geben wir einen Überblick über die Konzeption und Realisierung dieser autonomen Systeme und stellen einige mit ihnen gewonnene Erkenntnisse vor.

1 Einleitung

Ausgehend von der Hypothese, daß sich die Intelligenz von Lebewesen im Laufe der Evolution aus dem Zusammenwirken von Sehen, Bewegungssteuerung und Adaption an die Umwelt entwickelt hat, erforschen wir die Grundlagen eines solchen Zusammenwirkens bei autonomen Robotern. Unser Fernziel ist die Realisierung von Robotern mit einer praktischen Intelligenz von der Art, wie sie auch Lebewesen zum Überleben in ihrer jeweiligen Umwelt befähigt.

Höher entwickelte intelligente Lebewesen sind zumeist mit einem Sehsinn ausgestattet, der es ihnen erlaubt, sich in ihrer Umwelt zu orientieren und sich zweckmäßig und zielorientiert zu bewegen, z.B. zur Nahrungssuche oder zur Vermeidung von Gefahren. Nur in den Bereichen, in denen aufgrund des fehlenden Sonnenlichtes keine visuelle Erfassung der Umgebung möglich ist (z.B. in Höhlen oder im Meer), haben sich andere Sensorsysteme als leistungsfähiger erwiesen, beispielsweise die Verwendung von Ultraschall bei Fledermäusen und Delphinen. Wir sind daher der Auffassung, daß zur Entwicklung und Realisierung intelligenter Roboter Sichtsysteme von entscheidender Bedeutung sind, da sie es den Robotern ermöglichen, komplexe und dynamisch veränderliche Situationen in Echtzeit zu erkennen. Neben der Fähigkeit,

mit Sensoren Situationen in einer dynamischen Umwelt zu erkennen, benötigt ein intelligenter Roboter insbesondere auch die Fähigkeit des Lernens, also die Fähigkeit, durch Interaktion mit der Umwelt Kenntnisse zu erwerben bzw. zu erweitern und Verhaltensweisen zu erlernen bzw. zu verbessern. Es ist zu erwarten, daß Roboter, die eine derartige Intelligenz besitzen, in ihrer jeweiligen Umgebung eine ähnliche Adaptionsfähigkeit aufweisen werden wie Lebewesen.

Die Ausrichtung unserer Arbeiten wurde von Anfang an wesentlich durch den Grundsatz bestimmt, daß alle Forschungsergebnisse zum Nachweis ihrer Gültigkeit in praxisnahen Experimenten in der realen Umwelt überprüft und demonstriert werden müssen. Diese Arbeitsweise ist sehr aufwendig, hat aber gegenüber reinen Simulationen den gravierenden Vorteil, zu weitaus solideren und in der Praxis belastbareren Ergebnissen zu führen. Grundvoraussetzung für unsere Arbeiten sind daher Roboter und leistungsfähige Echtzeit-Bildverarbeitungssysteme, die wir zu Beginn selbst entwerfen und bauen mußten, da entsprechende Hardware auf dem Markt nicht verfügbar war. Einige der damit durchgeführten Experimente (u.a. Andockmanöver eines Luftkissenfahrzeugs, Landeanflug eines Flugzeugs und autonome Straßenfahrt) sind in Zusammenarbeit mit dem Institut für Systemdynamik und Flugmechanik der Universität der Bundeswehr realisiert worden [Dickmanns, Graefe 1988]. Im Rahmen des europäischen Verbundprojekts PROMETHEUS, das die Steigerung der Verkehrssicherheit und die Entlastung des Autofahrers zum Ziel hatte, erforschten wir die Grundlagen einer Verkehrssituationserkennung (Abschnitt 2). Danach stand die Navigation mobiler Systeme in bekannten und unbekanntem Innenraumumgebungen im Vordergrund, für die wir das Konzept der situationsabhängigen verhaltensbasierten Navigation entwickelten (Abschnitt 3). Das neueste von uns entwickelte System ist ein mobiler Manipulator, mit dem wir die sichtgesteuerte Handhabung verschiedener Objekte an unterschiedlichen Orten einer weitläufigen Einsatzumgebung untersuchen (Abschnitt 4).

2 Autonome Mobilität im Straßenverkehr

2.1 Echtzeit-Bildverarbeitungssysteme

Alle unsere Roboter setzen digitale Bildverarbeitung als leistungsfähige Sensorik ein. Die von Videokameras gelieferten Bilder werden in Echtzeit von Mehrprozessorsystemen ausgewertet. Der von uns gewählte gerätetechnische Ansatz war von Anfang an darauf ausgerichtet, eine hohe Systemleistung durch eine optimal an die Aufgabenstruktur des Rechnersehens angepaßte Systemarchitektur zu erzielen und nicht durch extrem aufwendige, aber unangepaßte Höchstleistungsrechner. Unsere Mehrprozessorsysteme sind aus lose gekoppelten Standard-Mikroprozessoren aufgebaut und deshalb leicht in Hochsprachen zu programmieren. Ein breitbandiger Videobus und ein unabhängiger Systembus (bzw. Link-Verbindungen bei Transputer- und DSP-Systemen) ermöglichen ein effizientes gleichzeitiges Arbeiten aller Prozessoren im System ohne Verzögerungen durch Kommunikationsengpässe.

Diese Hardware-Struktur bildet eine ausgezeichnete Grundlage für die Realisierung besonders leistungsfähiger „objektorientierter“ Sichtsysteme [Graefe 1989]. Grundgedanke ist dabei, daß das Gesamtsystem (Hardware und Software) entsprechend den sichtbaren Objekten strukturiert wird, soweit diese für die Aufgabe des sehenden Roboters relevant sind.

Fünf Generationen solcher Echtzeit-Bildverarbeitungssysteme wurden bisher entwickelt. Die erste, 1977 konzipiert und noch mit 8-Bit-Prozessoren ausgestattet, war gleichwohl schon in der Lage, ein instabiles mechanisches System (inverses Pendel) allein durch Rechnersehen zu steuern [Haas 1982]. Die dritte, mit 32-Bit-Prozessoren, weist etwa die tausendfache Leistungsfähigkeit auf und ermöglichte so die gesamte Bildverarbeitung für ein autonom im Autobahnverkehr mitfahrendes Auto [Graefe, Kuhnert 1992]. Während die ersten vier Generationen von Bildverarbeitungssystemen vollständig an unserem Institut entwickelt worden sind, setzen wir nun aus Wirtschaftlichkeitsgründen Industrieprodukte ein. Zur Zeit wird zur Bildverarbeitung und Robotersteuerung ein Mehrprozessorsystem bestehend aus digitalen Signalprozessoren des Typs TMS 320C40 von Texas Instruments eingesetzt, das in einem Standard-PC Platz findet [Bischoff 1997].

2.2 Objekterkennung in dynamischen Szenen

Beim ersten Schritt der Bildverarbeitung, der Merkmalsextraktion, stehen bei uns Verfahren im Vordergrund, die der Wirkungsweise rezeptiver Felder von biologischen Sichtsystemen ähnlich sind und die auf zeitaufwendige Ganzbildoperationen (z.B. Glättungs- und Segmentierungsverfahren) verzichten. Ein solches Verfahren zur Merkmalsextraktion, die gesteuerte Korrelation, wurde im Rahmen einer systematischen Untersuchung von Methoden zur schnellen, und vor allem robusten, Merkmalsextraktion am Institut für Meßtechnik entwickelt [Kuhnert 1986]. Es läßt sich außerordentlich effizient auf einem Mikrorechner implementieren und ermöglicht die Analyse von Bildfolgen in Echtzeit. Die flexible und robuste, aber normalerweise sehr rechenzeitintensive Korrelation wird – auch ohne Verwendung von Spezialhardware – dadurch echtzeitfähig, daß sie zum einen nur in den jeweils relevanten Teilen des Bildes berechnet wird und daß zum anderen als diskrete Vergleichsfunktion eine dünn besetzte ternäre Maske dient. Eine entscheidende Bewährungsprobe bestand das Verfahren der gesteuerten Korrelation 1987 bei den „Weltrekordfahrten“ eines fahrerlosen Straßenfahrzeugs (96 km/h auf Autobahnen und ca. 40 km/h im Universitätsgelände).

Entsprechend dem objektorientierten Grundansatz [Graefe 1989] wird zur Erzielung höchster Robustheit die Merkmalsextraktion von den höheren Systemebenen aus gesteuert und gestützt. Selbst unter erschwerten Bedingungen (Schatten, Reflexionen auf nasser Fahrbahn, texturierter Hintergrund etc.) lassen sich damit gute Ergebnisse

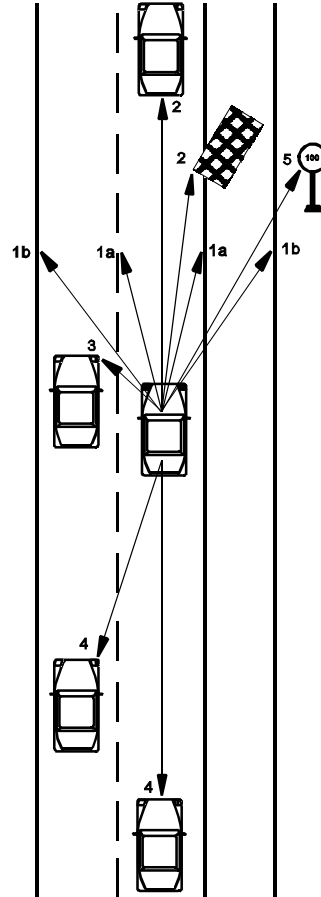


Abbildung 1: Typische Verkehrssituation mit verschiedenen Objekten in einer komplexen dynamisch veränderlichen Szene, die vom vorgestellten Bildverarbeitungssystem in Echtzeit erkannt wird.

erzielen. Primär werden dazu wegen ihrer Vielseitigkeit und Effizienz sogenannte 2D-Objektmodelle eingesetzt, die das Erscheinungsbild des betreffenden physischen Objekts modellieren. Dies können einfache Formmodelle sein; in komplexeren Szenen haben sich daneben Modelle auf der Grundlage von Symmetrien sowie von Statistiken über Merkmale im Bild besonders gut bewährt [Regensburger 1994].

Auch Modellwissen über die Entfernungabhängigkeit der optischen Abbildung wurde eingesetzt, und zwar bei der Erkennung von hinter dem eigenen Fahrzeug fahrenden anderen Fahrzeugen auf Autobahnen. Hierfür wurde ein einfaches Verfahren zur größennormierten 2D-Objektrepräsentation auf der Grundlage einer entfernungsabhängigen Unterabtastung von Bildern entwickelt. In Verbindung mit einer entfernungsabhängigen Kontrastanpassung bei der Merkmalsextraktion erleichtert es die Objekterkennung sehr. Zugleich ermöglicht es sehr kurze Ausführungszeiten von weniger als 1 ms [Efenberger 1996].

Die Hinderniserkennung beruht auf einem mehrstufigen Verfahren mit einem Entdeckungsprozeß und einem – unter Umständen mehrfach vorhandenen – Verfolgungs- und Klassifizierungsprozeß, denen noch weitere Prozesse, z.B. zur Schätzung des Bewegungszustands des Hindernisses, nachgeschaltet sein können. Entdeckungsentfernungen von über 300 m für kleinere und ca. 700 m für größere Objekte wurden erreicht [Solder 1992]. Abbildung 1 gibt einen Überblick über die von uns bearbeiteten Module zur Verkehrssituationserkennung, die erfolgreich in Autobahnszenen in Echtzeit getestet wurden [Graefe 1992, 1993].

3 Autonome Mobilität in Innenräumen von Gebäuden

Mit den sichtgesteuerten mobilen Robotern *ATHENE I* und *II* wurden die Voraussetzungen für die autonome Navigation in Fabrikgebäuden und Büroumgebungen geschaffen. Sie basieren auf konventionellen Flurförderfahrzeugen (Abbildung 2) und können in weitläufigen Einsatzumgebungen verschiedene Transportaufträge übernehmen [Wershofen 1996]. Hierzu wurde das Konzept der situationsabhängigen verhaltensbasierten Navigation entwickelt [Wershofen, Graefe 1992]. Die Verhaltensauswahl des Roboters erfolgt in jedem Moment auf Grundlage einer ständig mitlaufenden Situationserkennung. Diese Situation ergibt sich im wesentlichen aus den wahrgenommenen Zuständen relevanter Objekte, den aktuellen Handlungsmöglichkeiten und den zu erledigenden Aufgaben. Dies erfordert reichhaltige Sensorinformationen, wie sie primär von einem dynamischen objektorientierten Sichtsystem geliefert werden können [Bischoff et al. 1996].



Abbildung 2: *ATHENE II*, ein intelligenter mobiler Roboter, basierend auf einem konventionellen Flurförderfahrzeug und ausgestattet mit einem objektorientiertem Sichtsystem

Zur Navigation verwenden die Roboter topologische Karten, die sie bei einer Erkundungsfahrt selbständig aufbauen und mit Attributen über geometrische Sachverhalte sowie Beschreibungen von sichtbaren Landmarken und Hinweisen auf bestimmte Verhaltensweisen versehen. Ein Bediener überwacht die Roboter dabei, hilft ihnen an kritischen Stellen bei der Verhaltensauswahl und teilt ihnen relevante Stationsnamen mit, so daß ihnen einfache Hol- und Bringdienste unter Verwendung allgemein bekannter Namen wie z.B. „Kopierer“ oder „Labor“ aufgetragen werden können. Die Kommunikation zwischen Bediener und Roboter wird durch die situationsorientierte verhaltensbasierte Systemarchitektur stark vereinfacht. Insbesondere kann die Kommunikation situationsabhängig und unter Bezugnahme auf Sachverhalte und Objekte der Umgebung erfolgen, wie sie für den Benutzer erkennbar sind, und ohne Bezugnahme auf Roboter-interne Gegebenheiten wie etwa Speicheradressen, Variablennamen, Prozeßbezeichnungen oder Koordinatensysteme [Graefe, Bischoff 1997].

4 Mobile Manipulatoren

Serviceroboter rationalisieren und automatisieren Transport-, Handhabungs- und Verarbeitungsaufgaben in den verschiedensten Dienstleistungsbereichen. Mögliche Aufgaben liegen in den Bereichen Reinigung, Überwachung, Baugewerbe, Rehabilitation und Pflege oder Wartung und Reparatur. So vielfältig die Dienstleistungen sind, die an Menschen oder für Einrichtungen erbracht werden können, so unterschiedlich werden die Serviceroboter sein, die an ihre jeweilige Aufgabe und die Umwelt angepaßt sein müssen. Bestimmte Teilsysteme werden jedoch viele Serviceroboter gemein haben, beispielsweise eine mobile Plattform, ein Manipulationssystem, Steuer- und Antriebstechnik, Bedienerschnittstellen und Sensoren. Die dazugehörige Software sollte modular gestaltet sein, damit die einzelnen Teilsysteme je nach Bedarf zu funktionsfähigen Robotern konfiguriert werden können. Es ist eine große technische Herausforderung, solche Serviceroboter zu entwickeln und zu bauen, da Kompetenzen u.a. in vielen Ingenieurwissenschaften und der Informatik gefordert sind.

Gerade die Kombination von Mobilität und die Fähigkeit zur Manipulation stellen wesentliche Grundfunktionen der meisten Serviceroboter dar. So haben wir aufbauend auf unsere Erfahrungen zur situationsabhängigen verhaltensbasierten Navigation mit *ATHENE II* und der unkalibrierten Manipulation mit einem Knickarm-Roboter vom Typ Mitsubishi Movemaster [Graefe 1995] einen neuen vielseitig einsetzbaren Roboter entworfen und gebaut, den wir zur Erforschung und Lösung der noch offenen Probleme im Bereich der Servicerobotik nutzen wollen. Insbesondere bearbeiten wir mit diesem

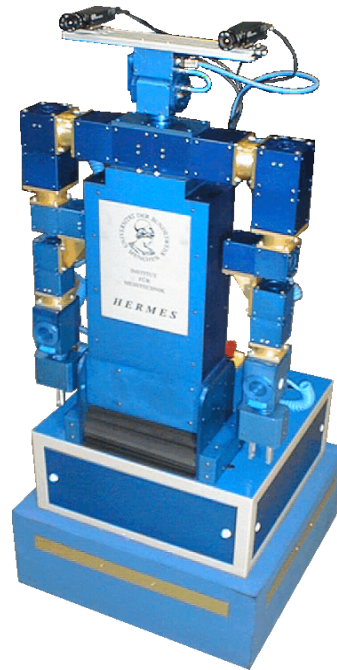


Abbildung 3:

Serviceroboter *HERMES* mit omnidirektionaler Basis, Manipulationssystem mit 13 Freiheitsgraden und Stereosichtsystem auf Schwenk-Neigeplattform (Abmessungen: 170 cm x 70 cm x 70 cm)

Roboter zunächst die autonome visuelle Erkundung einer weitläufigen Einsatzumgebung und die unkalibrierte Manipulation von verschiedenen Objekten an unterschiedlichen Orten innerhalb dieser Umgebung. Beim Entwurf und bei der Realisierung dieses mobilen „Manipulators“, den wir *HERMES* (“**H**umanoid **E**xperimental **R**obot for **M**obile **M**anipulation and **E**xploration **S**ervices”) genannt haben (Abbildung 3), richteten wir unser Hauptaugenmerk auf eine modulare und erweiterbare Struktur, sowohl bei der Hardware des Roboters (Antriebsmodule und Informationsverarbeitungssystem) als auch bei der Software [Bischoff 1997].

Die ersten Software-Module, die für *HERMES* geschrieben wurden, ermöglichen einem Bediener, den Roboter in seiner Einsatzumgebung an bestimmte Orte zu fahren und dort einfache Manipulationsaufgaben zu lehren und wiederholt durchführen zu lassen, z.B. das Aufheben eines Balles oder das Öffnen von Schubladen. Erste Implementierungen von sichtgekoppelten Verhaltensmustern erlauben *HERMES*, bekannte Objekte zu fixieren und sich Andockstationen (z.B. Beistelltischen) anzunähern [Graefe, Bischoff 1998].

5 Zusammenfassung und Ausblick

Maschinelles Sehen hat sich als sehr leistungsfähige Sensorik für autonome mobile Systeme erwiesen. Unser Ansatz des objektorientierten Sehens ermöglicht die Interpretation der reichhaltigen Sensorinformation einer Videokamera in Echtzeit und hat u.a. das autonome Mitfahren eines Straßenfahrzeugs im Autobahnverkehr selbst mit aus heutiger Sicht bescheidenen Mitteln und Anforderungen an die Hardware ermöglicht. Es ermöglichte ebenfalls die Analyse von Verkehrssituationen durch gleichzeitiges Entdecken, Klassifizieren und Verfolgen mehrerer Objekte.

Zur Steuerung mobiler Roboter in Fabrikgebäuden und Büroumgebungen wurde das Konzept der situationsabhängigen verhaltensbasierten Navigation auf Grundlage des objektorientierten Sehens entwickelt und mit dem mobilen Roboter *ATHENE II* in unserer Instituts Umgebung verifiziert. Es erlaubt einem Benutzer eine Art natürlicher Kommunikation mit dem Roboter, solange sich beide auf gemeinsam benannte und bekannte Objekte einer Einsatzumgebung beziehen. Die zukünftige Verbesserung der Wahrnehmungsleistung wird es dem Roboter ermöglichen, seine Einsatzumgebung ohne menschliche Überwachung zu erkunden.

Die Erweiterung der so gewonnenen Mobilität um die Fähigkeit, verschiedene Objekte zu handhaben, ist Ziel unserer gegenwärtigen Forschung. Dazu haben wir den modular aufgebauten Serviceroboter *HERMES* entwickelt, der bereits zahlreiche Eigenschaften besitzt, die von künftigen Servicerobotern erwartet werden, insbesondere was Modularität, Flexibilität und Erweiterbarkeit betrifft. Wir forschen an robusten Verfahren zur Steuerung mobiler Manipulatoren und haben bereits Verfahren entwickelt, die durch Vermeidung des Einsatzes von Modellwissen und durch den direkten Übergang von Bild-Sensordaten zu Motor-Steuerkommandos so robust sind, daß auf fortlaufende Kalibrierungs- und Wartungsarbeiten verzichtet werden kann. Damit eignet sie sich insbesondere für die Servicerobotik, in der Roboter einerseits ein hohes Maß an Robustheit im Umgang mit unerwarteten Situationen benötigen und andererseits Wartungspersonal nicht zur Verfügung steht.

Literatur

- Bischoff, R. (1997).** *HERMES - A Humanoid Mobile Manipulator for Service Tasks.* International Conference on Field and Service Robotics. Canberra, pp. 508-515.
- Bischoff, R.; Graefe, V.; Wershofen, K. P. (1996).** Combining Object-Oriented Vision and Behavior-Based Robot Control. Proc. Int. Conference on Robotics, Vision and Parallel Processing for Industrial Automation. Ipoh, Malaysia, pp 222-227.
- Dickmanns, E.D.; Graefe, V. (1988).** Applications of Dynamic Monocular Machine Vision. Int. J. Machine Vision and Applications. Vol. 1 (1988), pp 241-261.
- Efenberger, W. (1996).** Zur Objekterkennung für Fahrzeuge durch Echtzeit-Rechnersehen. Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München.
- Graefe, V. (1989).** Dynamic Vision Systems for Autonomous Mobile Robots. Proc. IEEE/RSJ Workshop on Intelligent Robots and Systems, IROS '89. Tsukuba, pp. 12-23.
- Graefe, V. (1992).** Visual Recognition of Traffic Situations by a Robot Car Driver. Proceedings, 25th ISATA; Conference on Mechatronics. Florence, pp 439-446.
- Graefe, V. (1993).** Vision for Intelligent Road Vehicles. Proceedings, IEEE Symposium on Intelligent Vehicles. Tokyo, pp 135-140.
- Graefe, V. (1995).** Object- and Behavior-oriented Stereo Vision for Robust and Adaptive Robot Control. Int. Symp. on Microsystems, Intelligent Materials, and Robots, Sendai, pp. 560-563.
- Graefe, V.; Bischoff, R. (1997).** A Human Interface for an Intelligent Mobile Robot. 6th IEEE Int. Workshop on Robot and Human Communication. Sendai, Japan, Sept. 1997, pp. 194-197.
- Graefe, V.; Bischoff, R. (1998).** Vision-Guided Intelligent Robots. Proc. Int. Conf. on Mechatronics and Machine Vision in Practice, Nanking, September 1998 (in Druck).
- Graefe, V.; Kuhnert, K.-D. (1992).** Vision-based Autonomous Road Vehicles. In I. Masaki (Ed.): Vision-based Vehicle Guidance. Springer-Verlag, pp 1-29.
- Haas, G. (1982).** Meßwertgewinnung durch Echtzeitauswertung von Bildfolgen. Dissertation, Fak. für Luft- und Raumfahrttechnik der Universität der Bw München.
- Kuhnert, K.-D. (1986).** A Model-driven Image Analysis System for Vehicle Guidance in Real Time. Proceedings, Second International Electronic Image Week. CESTA, Nice, pp 216-221.
- Regensburger, U. (1994).** Zur Erkennung von Hindernissen in der Bahn eines autonomen Straßenfahrzeugs durch maschinelles Echtzeitsehen. Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München.
- Solder, U. (1992).** Echtzeitfähige Entdeckung von Objekten in der weiten Vorausschau eines Straßenfahrzeugs. Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München.
- Wershofen, K. P. (1996).** Zur Navigation sehender mobiler Roboter in Wegenetzen von Gebäuden – Ein objektorientierter verhaltensbasierter Ansatz. Dissertation, Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München.
- Wershofen, K. P., Graefe, V. (1992).** An Intelligent Autonomous Vehicle Guided by Behavior-based Navigation. IFToMM-jc International Symposium on Theory of Machines and Mechanisms. Nagoya, pp 244-249.