

Entwicklung des Gesichtssinns für automatisiertes Fahren

Der 4-D Ansatz brachte 1987 den Durchbruch

Dr.-Ing. **Ernst D. Dickmanns**, Prof. i.R.; UniBw München, Neubiberg

Kurzfassung

Nach kurzen Skizzen der zeitlichen Entwicklung von Bodenfahrzeugen und digitalen Prozessrechnern wird der Durchbruch zu dynamisch sehenden Fahrzeugen in den beiden letzten Jahrzehnten des vorigen Jahrhunderts etwas eingehender behandelt. Das Projekt „ProArt“ im Rahmen des EUREKA-Vorhabens ‚PROMETHEUS‘ (ab 1987) brachte Europa an die Spitze der Entwicklung sehender Straßenfahrzeuge; der demonstrierte Stand an der UniBw München am Projektende 1994 wird beschrieben. Danach erfolgte eine zweimalige Aufspaltung der Entwicklungslinien, die thematisiert werden. Den Abschluss bildet ein Ausblick mit der Frage, ob langfristig sehende technische Subjekte auch aktive Augen haben werden wie die meisten hochentwickelten biologischen Systeme.

Abstract

Following short sketches of the temporal development of ground vehicles and digital micro-processors, the breakthrough to vehicles capable of dynamic vision in the last two decades of the previous century is treated to some more detail. The project “ProArt” in the framework of the EUREKA-initiative ‘PROMETHEUS‘ (starting 1987) has brought Europe in a leading position in the development of ground vehicles capable of real-time vision; the demonstrated state of development at UniBw Munich at the end of the project in 1994 is described. After that, two bifurcations in the line of development occurred that are discussed. An outlook with the question ‘Will technical agents with the sense of vision have active eyes like most highly developed biological species’ forms the conclusion.

1. Technischer Hintergrund bei Fahrzeugen, Elektronik und Methodik

Zur Einleitung sei ein kurzer Überblick über die zeitliche Entwicklung von Bodenfahrzeugen und Elektronik (Sensoren und digitalen Mikroprozessoren) mit einigem Abstand gestattet, der den abschließenden Ausblick motiviert. Bei der Methodik kamen die unterschiedlichen Hintergründe in der Informatik und dem Gebiet der Systemdynamik zum Tragen, die kurz beleuchtet werden.

1.1 Bodenfahrzeuge

Bild 1 gibt einen gerafften Überblick über die Entwicklung von Radfahrzeugen in den letzten etwa sechs Jahrtausenden. Von einer in Polen gefundenen Keramik kann man schließen, dass schon vor fünfeinhalb tausend Jahren Radfahrzeuge bekannt waren (oben links in Bild 1). Älteste gefundene Reste von Rädern reichen gut fünftausend Jahre zurück (mit Speichen unten links, gut dreitausend [1]). Die alten Römer hatten schon ein weit ausgebautes Netz von Straßen über ganz Westeuropa. Lebewesen mussten über viele Jahrtausende die Antriebsleistung liefern. Im 18. Jahrhundert kamen die ersten Dampfmaschinen zur Erprobung, aber der Durchbruch zur Motorisierung von Straßenfahrzeugen erfolgte erst gegen Ende des 19. Jahrhunderts, in Deutschland vor allem durch Carl Benz (rechter Bildteil). Die Serienfertigung von Autos am Fließband ab 1932 durch Henry Ford brachte zwar einen starken Anstieg der verkauften Automobile weltweit, aber der massive Anstieg der Bodenfahrzeuge in Deutschland erfolgte erst nach dem 2. Weltkrieg (Bild 2, gestrichelte Kurven, untere nur PKW). Die stückweise durchgezogene Kurve zeigt den Anstieg der Zahl der Unfalltoten in Deutschland, die im Jahr 1970 einen Höchststand erreichte; danach sank sie durch die verschiedensten Maßnahmen kontinuierlich ab. Seit mehreren Jahren liegt sie im wiedervereinigten Deutschland zwischen drei und vier tausend. Weltweit beträgt die Zahl



Bild 1: Skizze der zeitlichen Entwicklung von Radfahrzeugen bis ~ 1900



Bild 2: Kraftfahrzeuge als Massenprodukt und jährliche Zahl der Unfalltoten.

der jährlichen Verkehrstoten heute etwa $1\frac{1}{4}$ Millionen, davon etwa ein Viertel Kinder. Da menschliches Versagen die häufigste Ursache ist, liegt heute auf dem Gebiet der Unterstützung von menschlichen Fähigkeiten der größte Aufwand zum Fortschritt in Richtung null Verkehrstote. Deswegen sollen seine Sinnesfähigkeiten durch Elektronik ergänzt und erweitert, langfristig möglicherweise ersetzt werden. In [2] wird ein Überblick über die Geschichte sehender Straßenfahrzeuge gegeben.

1.2 Elektronik

Nach der Erfindung des Transistors 1947 und des Mikroprozessors (μP) in den 1970ern ist bisher eine ständige Weiterentwicklung der μP zu beobachten (Erhöhung der Taktraten, Verkleinerung der Breiten der Leiterbahnen, dichtere Packung etc.), die zu einer Steigerung der Rechenleistung um einen Faktor zehn alle vier bis fünf Jahre geführt hat. Bild 3 zeigt im Zentrum die Zahl der Transistoren für geläufige μP über der Zeit. Sie liegt 2017 bei etwa zehn Milliarden; Mitte der 80er Jahre lag sie bei ~ 1 Million. Das menschliche Gehirn hat etwa 100 Milliarden Neuronen (Bildmitte oben) mit je bis zu zehntausend Querverbindungen; die Zahl der Elemente, die in ihren Eigenschaften allerdings sehr unterschiedlich sind, wird von den technischen μP etwa in der ersten Hälfte der 20er Jahre erreicht werden, wenn die

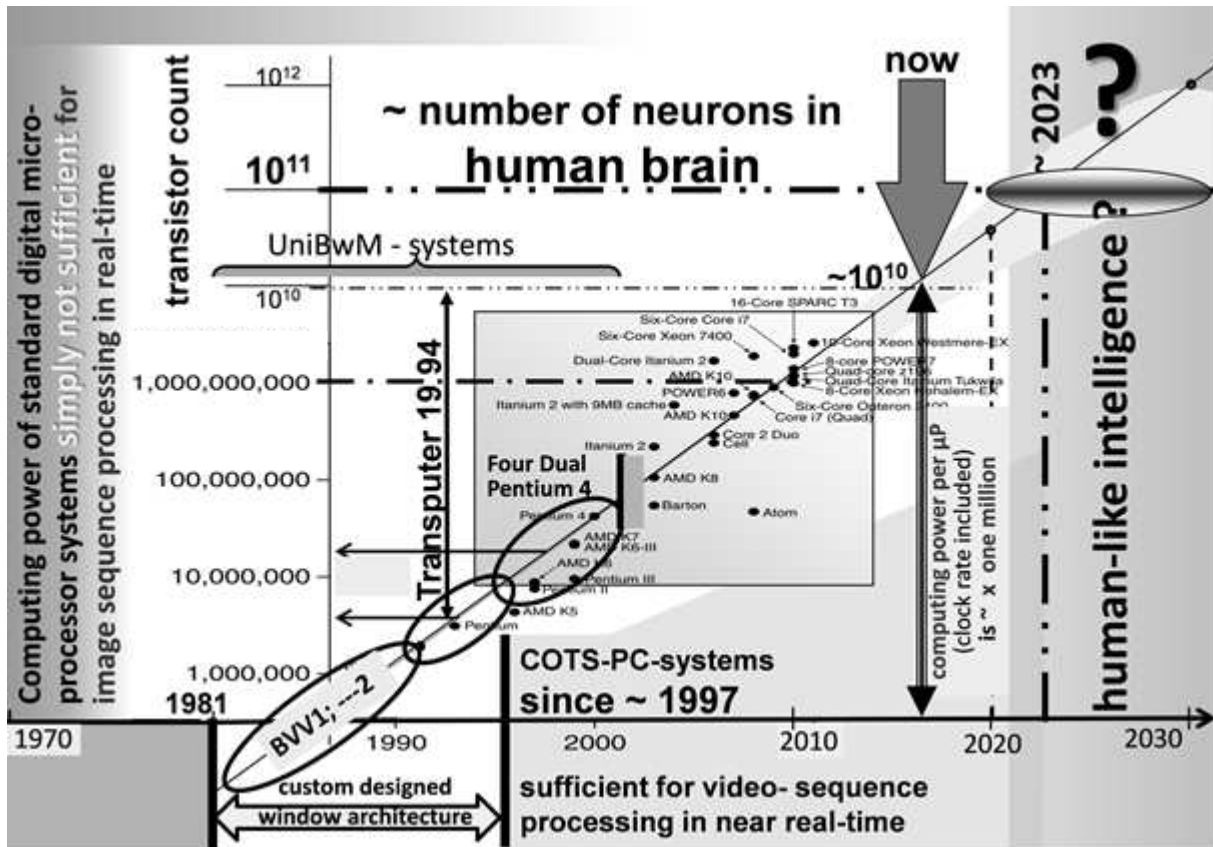


Bild 3: Skizze der Transistorzahlen pro μP mit gängigen Typen über der Zeit. Die gemittelte Steigung ergibt mit der parallelen Größenreduzierung den Leistungszuwachs von etwa einer Größenordnung alle vier bis fünf Jahre.

gemittelte Steigung im Bild (schwarze Diagonale) weiter gilt. Kann man dann mit menschenähnlicher Intelligenz bei technischen Systemen rechnen (rechts im Bild)? Der unten rechts schattierte Bereich gibt den Zeitraum an (seit ~ 1997), seit dem mit handelsüblichen Netzwerken und μP hinreichend leistungsfähige Echtzeit-Bildverarbeitungssysteme realisiert werden konnten. In der Lücke von ~ 1980 bis 97 mussten spezielle Konzepte eingesetzt werden, um das Ziel für spezifische Aufgabengebiete zu erreichen. Die US-amerikanische DARPA hatte 1982 eine große Entwicklung mit ‚Massiv Parallelen (Einfach-) Prozessoren‘ (MPP) angestoßen, die aber kaum zu einsetzbaren Ergebnissen geführt hat [3]. Man wollte in typischen Straßenumgebungen aller Art zunächst im Schrittempo fahren können.

Im Gegensatz dazu haben wir die Autobahnumgebung mit ihren strikt geregelten Fahrbahnparametern als Einstieg gewählt: Kreuzungsfreiheit, gute Oberflächen und Spur-Markierungen, unidirektionaler Verkehr etc. sowie begrenzte Arten von Verkehrsteilnehmern (nur schnelle Fahrzeuge, keine Radfahrer und Fußgänger etc.). Damit konnte die Zahl der verschiedenen notwendigen Extraktoren für Bildmerkmale klein gehalten werden. Bei

Verwendung von rekursiven Schätzverfahren aus der Systemdynamik und einem speziellen ‚Fensterkonzept‘ (s. unten) bei Parallelprozessoren, sollte Echtzeitbetrieb mit etwa 10 Hz Auswertefrequenz und konventionellen μP bei dieser Aufgabe schon Anfang der 1980er Jahre möglich sein.

1.3 Methodik

Wesentlich für die Systemleistung bei der Anwendung ist die eingesetzte Methodik samt Software. In allen damals üblichen Ansätzen aus den Fachgebieten ‚Künstliche Intelligenz‘ und ‚Computer Science‘ wurden die Bildfolge und die Lage der Objekte in konsekutiven Bildern als primäre Messgrößen gewählt. Nach Inversion der perspektivischen Abbildung und anschließender Differenzbildung zwischen den Resultaten aufeinander folgender Bilder wurde dann auf die zeitliche Bewegung geschlossen. Diese Differentiation vergrößert das Messrauschen und die Totzeit bis zum Erhalt der neuen, aktuellen Schätzwerte.

Im Gegensatz dazu haben wir von Anfang an mit dynamischen Modellen im 3-D Raum für den zu beobachtenden *Prozess* gearbeitet, also mit *räumlich-zeitlichen Strukturen*. Die Verbesserung der Schätzwerte erfolgt dann durch Rückkopplung von Vorhersagefehlern; die hier erfolgende zeitliche Integration mittels der Differenzgleichungen für den zu beobachtenden Prozess reduziert das Messrauschen und erlaubt, auch Ableitungsgrößen (wie Geschwindigkeitskomponenten) in einem Durchgang mit zu schätzen. Bei vorliegender Beobachtbarkeit können sogar nicht direkt messbare Größen wie die bei perspektivischer Abbildung verloren gegangene Entfernung rekonstruiert werden, ein wesentlicher Vorteil ! Dies führt zur Fähigkeit des monokularen Bewegungs-Stereosehens.

Bei von Anfang an hoher Auswertefrequenz ($1/4$ Videorate = 12,5 Hz, 80 ms Zykluszeit) bilden höhere Fahrgeschwindigkeiten keine besonderen Schwierigkeiten. Bei einer Sicht von 6 bis 30 m und einer Geschwindigkeit von 108 Km/h (30 m/s) werden alle Abschnitte der Fahrbahn achtmal nacheinander mit verschiedenen Auflösungen erfasst, was sowohl eine Erkennung der Krümmungsparameter sowie der eigenen seitlichen Ablage gestattet. Dies ist bei Hindernisfreiheit alles, was man zum sicheren Fahren braucht.

2. Ein alternativer Ansatz zum Echtzeit-Sehen (aus der Systemdynamik)

Nimmt man im konventionellen deutschen Fernsehbild nur eines der Halbbilder (entweder nur alle geraden oder nur alle ungeraden Zeilen) so hat man zwar die vertikale Auflösung halbiert, aber die doppelte Zeit zur Bildinterpretation zur Verfügung. Dieses Halbbild (engl. video field) hatte typischerweise 360×120 Bildpunkte (Pixel). Bild 4 zeigt ein typisches Vollbild mit eingezeichneten Auswertefenstern. Ihre Lage und ihre Parameter können von

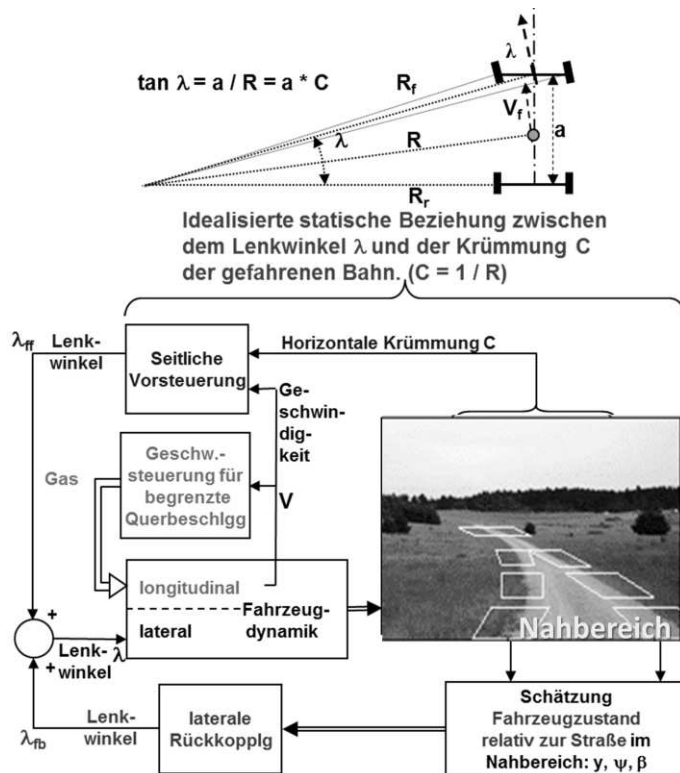


Bild 4: Skizze des Verfahrens zur rekursiven Fahrbahnerkennung und Fahrzeugsteuerung (1986)

geeignete Rückkopplung der Fehler korrigiert (λ_{fb} unten links). Bei dieser Vorgehensweise ist keine Sollbahn zu berechnen; die Dynamik des gewünschten Verhaltens kann mit den bewährten Methoden der Systemdynamik und Regelungstechnik festgelegt werden [4 bis 7].

2.1 Integrierte räumlich / zeitliche Modelle für den realen Prozess

Gemäß Bild 5 wird parallel zur Beobachtung des realen Prozesses in der Welt über Fernsehsignale (links im Bild) durch Merkmalvergleich im Auswerteprozess (rechtes Quader) eine Vorstellungswelt aufgebaut mit Prozessmodellen in Raum (je nach Bedarf bis zu drei Raumdimensionen) und Zeit. Die Initialisierung ist der kritische Schritt, bei dem aus der Kombination von Merkmalen in der Bildfolge Hypothesen von Objekten in der beobachteten Welt und deren Relativposition gebildet werden müssen. Mit den ebenfalls hypothetischen Prozessmodellen sowie den perspektivischen Abbildungsgesetzen werden dann die Lagen der Merkmale im nächsten Bild vorhergesagt; dies erlaubt die Einschränkung der horizontalen und vertikalen Suchbereiche im Bild. Mit den Vorhersagefehlern und der Jacobi'schen Matrix der Abbildungsgleichungen (Matrix der Ableitungen der Mess- nach den

einem zum nächsten Videozyklus der Situation angepasst werden. Ein Modell mit Variablen der Differentialgeometrie für den Krümmungsverlauf der Fahrbahn gestattet mit der Lenkgeometrie (oben im Bild) die Festlegung eines angemessenen Ausschlags für den niederfrequenten Lenkwinkel λ_{ff} (oben links) sowie der maximalen Geschwindigkeit, wenn eine Grenze in der Querbeschleunigung einzuhalten ist. Sie wird über die Stellglieder Gas und Bremse geregelt.

Im Nahbereich (unten im Bild) ist aus der Lage der Fahrbahnkanten die laterale Position des eigenen Fahrzeugs auf der Fahrbahn abschätzbar. Deren Abweichung vom gewünschten Sollwert wird durch eine

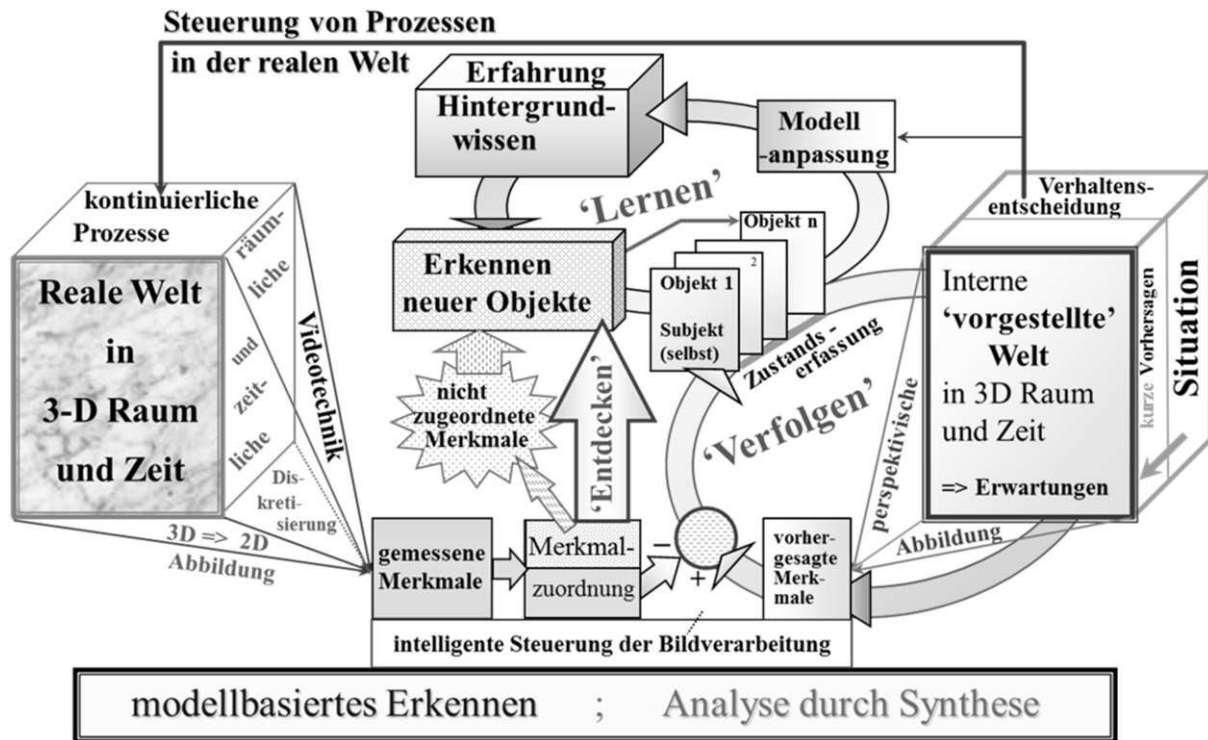


Bild 5: Aufbau einer Vorstellung über die beobachtete Welt durch Rückkopplung von Vorhersagefehlern für markante Merkmale und Objektpositionen; drei parallele Aktivitäten: Entdecken neuer Merkmale und Generierung von Objekthypothesen, verfolgen von Objekten, und Modellanpassung (Lernen).

Zustandsgrößen) können dann die Modellvorstellungen verbessert werden (Erweitertes Kalman Filter EKF). Einzelheiten sind zu finden in [4 bis 10], speziell [7, Kap. 6 und 7]. Drei Prozesse laufen stets parallel: 1. Das Entdecken neuer Merkmale und die Bildung neuer Objekthypothesen (Bildmitte), 2. Das Verfolgen von Objekten (rechts) mit bekannten Hypothesen, und 3. Eine Buchführung über interne Prozessabläufe, die zum Lernen verwendet werden kann (oben).

2.2 Das Fensterkonzept zur Echtzeit-Merkmalsextraktion in Bildfolgen

Wie aus Bild 3 ersichtlich, war vor 1997 Bildfolgenauswertung in Echtzeit mit kommerziellen μP nicht möglich. Das dynamische Fensterkonzept in Kombination mit dem 4-D Ansatz [4; 7, Kap. 5] wurde als einzige Möglichkeit gesehen, mit verfügbaren μP in gut strukturierten Szenen dieses Ziel zu erreichen. Fahren auf Schnellstraßen wurde als leichteste praktisch relevante Anwendung schon Ende der 1970er Jahre erkannt. Hierfür wurde ein ‚Hardware-In-the-Loop‘ (HIL) Simulationskreis aufgebaut, in dem das Fensterkonzept gemäß Bild 6

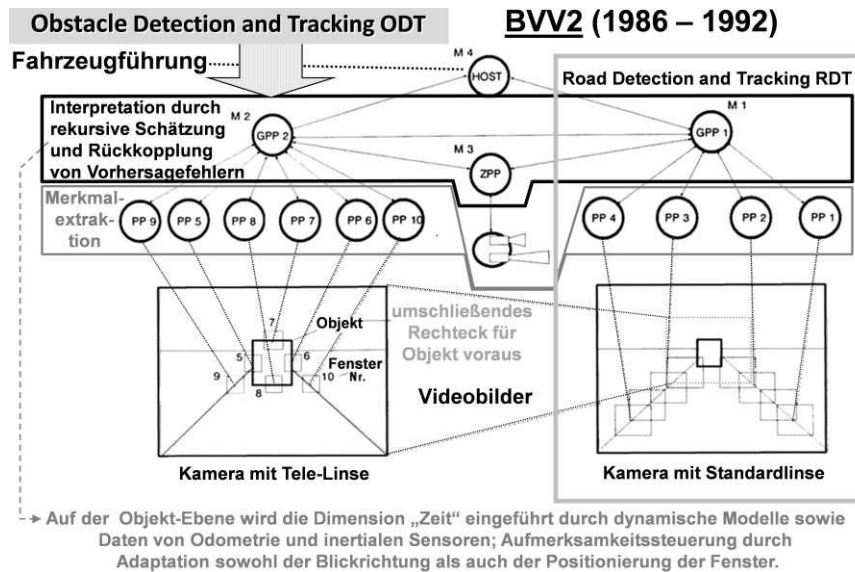


Bild 6: Das Fensterkonzept mit flexibel anpassbaren reduzierten Auswertebereichen für N Objekte parallel erlaubte frühes Echtzeit-Sehen auf Straßen ab 1985.

Objekt wurden Gruppen von Merkmalen definiert (s. rechte Hälfte für die Fahrbahn, linke für ein einzelnes Objekt in Bild 6). Bei Fahrzeugen als Hindernissen kamen nur vertikale und horizontale Kanten in Frage, da nur der untere Körperteil analysiert wurde; der obere Teil bei PKW mit Scheiben und polierten Flächen liefert in der Regel durch Spiegelung der Umgebung schnell variable, wenig verlässliche Bildmerkmale [7 bis 10].

3. Das Projekt „PROMETHEUS“-ProArt 1987 – 94

Nachdem mit VaMoRs 1986/87 auf alleinige Initiative der UniBwM die prinzipiell möglichen Fähigkeiten demonstriert worden waren, zeigte sich das mittlere Management der Daimler-Benz AG (DBAG) daran interessiert, gemeinsame Projekte über Drittmittel finanzieren zu lassen. Nach Einrüstung unseres BVV2 in einen dort vorhandenen ‚Spur Bus‘ im Rahmen des BMFT-Projektes ‚Autonom Mobile Systeme‘ mit erfolgreichem autonomem Anhalten vor einem stationären Hindernis aus 40 km/h auf der Versuchsstrecke Rastatt 1988 gelang es, die bis dahin im Prometheus-Projekt geplante laterale Führung durch vergrabene Kabel und induktive Felder auf Autobahnen durch maschinelles Sehen zu ersetzen. Der zunächst für die DBAG mit einer Kopie unseres BVV2 ausgerüstete 7t-Van ‚Vision Information Technology Application‘ VITA brachte auf der Halbzeitdemonstration Prometheus 1991 bei FIAT in Turin auch den Durchbruch beim oberen Management der Autoindustrie. Für die Abschlussdemonstration 1994 in Paris wurde ein sehr ehrgeiziges Ziel gesetzt: Man wollte

systematisch untersucht werden konnte. Jedem der parallelen μP wurden für jeden Auswertezyklus neu einige 32×32 Pixel große Bildbereiche vom Auswerteprozess mit günstigen Parametern für die Suche zugewiesen. Merkmale waren Kanten mit einem Winkelbereich für die Neigungen und für angrenzende mittlere Grauwerte. Für jedes

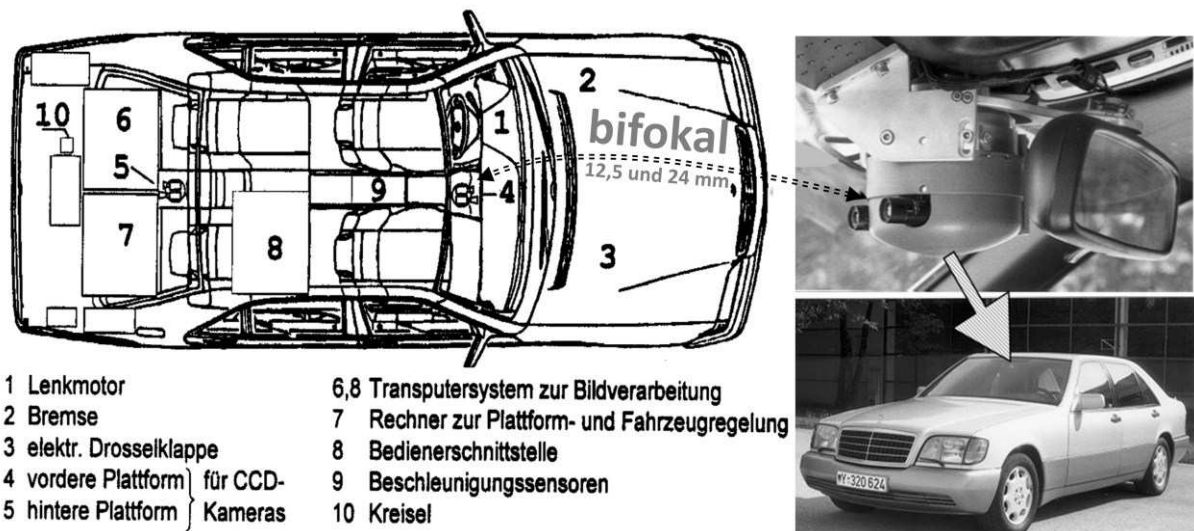


Bild 7: VaMoRs-PKW (kurz VaMP), das gemeinsam mit dem Zwillingfahrzeug der DBAG Mercedes 500-SEL (VITA2 als CED2) zur Abschlussdemonstration von Prometheus 1994 autonom im französischen dreispurigen Autoroute-Verkehr mit 1 Gast an Bord als CED 3 mitfuhr (bei VITA2 mit 3 Gästen war alle Elektronik im Kofferraum).

mit Gästen an Bord auf einer Autobahn am dortigen Flughafen CDG autonom im normalen dreispurigen Verkehr mitfahren können. Um das Ziel zu erreichen, wurden an der UniBwM zeitweise bis zu 20 Mitarbeiter vom Prometheus-Projekt finanziert. Folgende Leistungen wurden mit diesen damals einzigen dazu fähigen und weltweit schnellsten autonomen Fahrzeugen (Bild 7) mit dem Sehsystem der UniBwM auf Transputer-Basis erzielt [10 bis 15; und 7, Kap. 11]:

- + Erfassung der Parameter der eigenen Spur und der beiden direkten Nebenspuren,
- + Freies Spurfahren mit Geschwindigkeiten bis zu den erlaubten max. 130 km/h auf französischen Autobahnen,
- + Konvoi-Fahren mit geschwindigkeitsabhängigem Abstand (inklusive Anhalten),
- + Übergang in diesen Modus vom freien Spurfahren aus,
- + Vollautonome Spurwechsel und Überholen unter Beachtung von je bis zu fünf benachbarten Fahrzeugen in der vorderen und der hinteren Hemisphäre.

Die erfolgreichen Vorentwicklungen mit VaMoRs zur benötigten Technik führten dazu, dass uns vom zuständigen Sicherheitsbeauftragten der Bundeswehr erlaubt wurde, ab 1992 im öffentlichen Straßenverkehr autonom mitzufahren und zu testen, wenn drei Menschen zur Überwachung der Sicherheit an Bord waren; dies war ein wesentlicher Vorteil gegenüber

privaten Experimentatoren. Ohne dieses großzügige Entgegenkommen wäre das gesetzte Ziel wohl kaum erreichbar gewesen.

Parallel zu den Arbeiten im Prometheus-Projekt auf Autobahnen wurden im Bereich der allgemeinen Anwendungen auf Straßen niederer Ordnung mit VaMoRs die Fähigkeiten entwickelt, Querstraßen zu entdecken und deren Schnittparameter (Abstand, Breite, Winkel) zu schätzen [15; 7, Kap. 10]. Auch hierbei hat anfänglich der HIL-Simulationskreis eine wesentliche Rolle gespielt.

4. Zwei Gabelungen der Entwicklung des maschinellen Sehens (1997 und 2001)

In den Jahren 1995/96 wurde das entwickelte Sehsystem auf etwa um das Zehnfache leistungsfähigere μ P (Motorola Power-PC) im Rahmen eines EU-Projektes übertragen. Neben der Erhöhung der Bildauswerterate auf 25 Hz konnte die Zahl der benötigten COTS-Prozessoren für die Bildauswertung auf ein Fünftel reduziert werden [13; 14]. Mit diesem System wurde 1995 die Langstreckentestfahrt von München nach Odense zu einem Projekt-Treffen durchgeführt (Summe > 1600 km autonom gefahrene Teilstrecken) [14; und 7, Sektion 9.4.2.5].

Da Daimler wie die meisten anderen Industriefirmen anschließend nur an weniger anspruchsvollen Sehsystemen interessiert war, haben sich die Wege 1997 getrennt. Wegen der hohen Anfangskosten bei maschinellem Echtzeitsehen waren wir der Überzeugung, dass sich diese Systeme vor allem mit der Fähigkeit zum flexiblen Mehrfacheinsatz in der Praxis rechnen würden. Deshalb waren wir von vorne herein daran interessiert, Systeme zu entwickeln, die sich langfristig den Fähigkeiten des menschlichen Sehvermögens annähern können, d.h. die flexibel und lernfähig sind (oder zumindest werden können).

Industrie (geringe Systemkomplexität): 1. Bildauswertung für nur eine Aufgabenklasse, 2. Eine oder zwei Kameras körperfest montiert; 3. Niedrige aktuelle Kosten und erforderliche Rechenleistung 4. Kleine Wissensbasis, 5. Die Schätzung von Abständen erfolgt primär mit aktiven Sensoren (Radar oder Lidar).

Forschung (UniBwM: Erkenntnisgewinn mittels Sehen): 1. Langfrist Aspekte der flexiblen Mehrfachnutzung in Aufgabenbereichen dominieren. 2. Datenökonomie erfordert multiple Brennweiten parallel mit aktiver (zum Teil sehr schneller) Blickrichtungssteuerung (Sakkaden); 3. inertielle Stabilisierung mittels einfacher Drehratengeber auf einer Plattform; 4. Umfangreiche Wissensbasen auf verschiedenen Ebenen erlauben einfache Adaptation an Aufgaben und Situationen. 5. Auf aktive Sensoren wie Radar und Lidar wird analog zum menschlichen Wahrnehmen verzichtet.

Durch unsere Erfolge mit den Transputer-Systemen sowie den gezeigten Fahrleistungen nicht nur auf Autobahnen, die ständig auf internationalen Konferenzen präsentiert wurden, zeigte sich der amerikanische Verteidigungsbereich an einer Zusammenarbeit interessiert. In USA wurde Anfang der 90er Jahre die Entwicklung autonomer Bodenfahrzeuge von der DARPA weg und hin zum Army Research Laboratory (ARL), Aberdeen MD verlagert. Unter dessen Federführung wurde 1996/97 im Rahmen eines ‚Memorandum of Understanding‘ zwischen den Verteidigungsbereichen zusammen mit dem ‚National Institute of Standards and Technology‘ (NIST, Gaithersburgh, MD, J. Albus [19; 20]) sowie der Forschungsgruppe SRI (Princeton, NJ, P. Burt) und einem amerikanischen Industriepartner (General Dynamics, Robotik) das Projekt „AutoNav“ definiert. Auf deutscher Seite war die Fa. Dornier GmbH, Friedrichshafen beteiligt, die das mit VaMoRs zu entwickelnde System auf das Kettenfahrzeug Wiesel2 (einen digitalen Versuchsträger) portieren sollte.

Ziel war die Entwicklung flexibler Fähigkeiten in visueller Wahrnehmung, Aktionsplanung und Missionsdurchführung, so dass bis 2001 kleinere Missionen auf Wegenetzen niedriger Ordnung einschließlich querfeldein-Anteilen mit der Vermeidung negativer Hindernisse (speziell eines Grabens) autonom durchgeführt werden konnten. Die genaue Geometrie des Wegenetzes und der Übergänge sollte nicht vorab als Wissen eingegeben werden müssen, sondern war bei der Durchführung der Mission vor Ort zu erkennen; dies bezeichnen wir heute als „Pfadfinder-Sehen“ (Scout-type Vision). Zur Lösung dieser Problemstellung wurde das „Erwartungsbasierte, Multifokale, Sakkadische“ Sehen (Expectation-based, Multi-focal, Saccadic- or EMS-Vision) entwickelt [17; 18a) bis 18f); weiter bis 25] und [7, Kap. 13/14]. Details finden sich auf der Website [27] sowie unter Youtube [28]. Bild 8 zeigt die Abschluss-Demonstration im Jahr 2001 in Neubiberg, bei der die angestrebten Ziele erreicht wurden.

Aufgrund ähnlicher Demonstrationen unserer amerikanischen Partner mit deren Fahrzeugen (HMMWV’s und XUV, **eX**perimental **U**nmaned **V**ehicle) in USA gelang es dort, in den Budgetberatungen des Verteidigungs-Haushalts 2001 vom US-Congress das Ziel schriftlich festzulegen, dass bis 2015 ein Drittel der Kampf-Bodenfahrzeuge der USA die Fähigkeit haben sollte, Missionen teilweise autonom durchzuführen.

Auf dieser Basis hat die DARPA dann zunächst für 2004 einen Preis in Höhe von 1 Million \$ für die Gruppe ausgeschrieben, die zu dem Termin zeigen konnte, dass ihr Fahrzeug **eine Versorgungsmission** über mehr als 200 km in der kürzesten Zeit autonom durchführen konnte. Allerdings wurden die Anforderungen dadurch erleichtert, dass eine genaue momentane Positionsvermessung durch GPS zur Verfügung stand und dass die Fahrstrecke durch eng gestaffelte GPS-Wegpunkte vorgegeben wurde. Negative Hindernisse (Gräben etc.) wurden ausgeschlossen, und zur Entdeckung von positiven Hindernissen (oberhalb der

Mission performance on network of minor roads

including offroad sections
with a negative obstacle

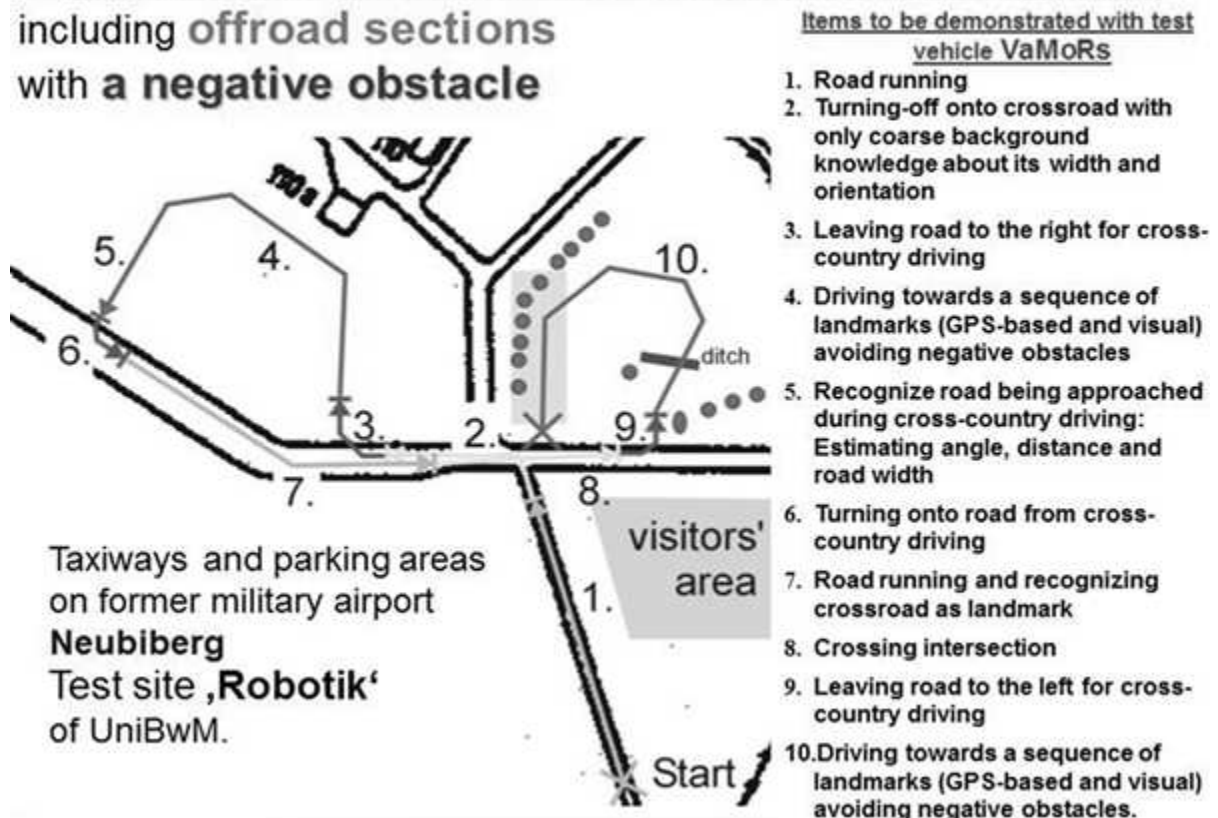


Bild 8: Autonome Missionsdurchführung in einem Wegenetz niederer Ordnung mit 2-maligem Verlassen zur Querfeldeinfahrt und wieder Auffahren auf die befestigte Straße. Im Schlusssektor 9-10 befindet sich ein Graben als negatives Hindernis an unbekannter Stelle.

Fahrebene) standen leistungsfähige Laser-Range-Finder zur Verfügung. Maschinelles Sehen spielt unter diesen Umständen keine wesentliche Rolle und wurde von einigen Gruppen auch ignoriert. Da genaue Wegekarten zur Verfügung standen, wurde visuelle Wahrnehmung vor allem dafür eingesetzt, die in der aktuellen Wissensbasis über die momentane Umgebung vorliegende Information über (markante) stationäre Objekte zu bestätigen; aus diesem Grund wird diese reduzierte Sehfähigkeit hier als „Bestätigungs-Sehen“ (Confirmation-type Vision) bezeichnet. Diese neue Entwicklungslinie mit geringeren Anforderungen an die visuellen Wahrnehmungsfähigkeiten (vor allem die erforderlichen Wissensbasen) erfreut sich seither in großen Teilen der Forschung und der Industrie großer Beliebtheit. Dies ist die zweite Gabelung der Entwicklungslinien, die mit den DARPA-Challenges 2004++ begann. Für anspruchsvollere autonome Fahrzeuge, die sich mit ihren Wahrnehmungsfähigkeiten der menschlichen integrierten Erfassung von Raum und Zeit

annähern sollen, ist ‚Scout-type Vision‘ nötig. Die beiden Typen sind folgendermaßen zu charakterisieren:

Pfadfinder-Typ (Scout-type): Die Fähigkeit, in unbekannter Umgebung ohne weitere Hilfen zu fahren, erfordert einen hinreichend großen Satz von Bildmerkmalen (hohe Auflösung, Farbe, Texturen) sowie große spezifische Wissensbasen über: a) Erzeugung und Test von Objekthypothesen aus Bildmerkmalen (4-D Ansatz); b) klassenspezifische Fähigkeiten zur Manöverdurchführung, c) Domänen-spezifische Situationen zu deren aktueller Erfassung und zur Entscheidungsfindung. Multifokales Sehen mit hoher zentraler Auflösung und aktiver Blickrichtungssteuerung sind nötig für bessere Dateneffizienz. Farbsehen erleichtert die Erkennung komplexer Situationen. Inertiale Blickstabilisierung ist wichtig für das Fahren auf unebenem Untergrund und für die Blickfixierung auf (sich bewegende) Objekte. Diese Systeme haben großes Wachstumspotential und gute Lernfähigkeiten wegen der großen Signalvielfalt und der spezifischen sowie allgemeinen Wissensbasen (Farben, Texturen).

Bestätigungssehen (Confirmation-type): Fahren ist nur in vorabgeklärten Umgebungen mit GPS und genauen Landkarten möglich. Ein reduzierter Satz an Bildauflösungen und Merkmalen kann hinreichen. Dieser Typ erfordert große, möglichst aktuelle lokale Datenbasen sowie GPS und hochgenaue Landkarten mit stationären Objekten; er kommt mit bescheideneren Wissensbasen aus, hat aber hohe laufende Kosten wegen der erforderlichen Aktualität der Umweltdaten. Bis heute erfordert er 360°-Laser (mit meist mäßigen Auflösungen in größerer Entfernung $> \sim 60\text{m}$); für diese ist eine aufwendige Synchronisation erforderlich; körperfeste Montage der Sensoren und einfache Kameras reichen meist. Dieser Typ ist prädestiniert für ebenes Gelände und gute Lichtbedingungen; er hat vergleichsweise wenig technologisches Wachstumspotential.

Bewegte Objekte und Subjekte erfordern ähnliche Fähigkeiten in beiden Bereichen.

5. Ausblick

Seit den DARPA Grand- und Urban Challenges 2004 bis 2007 und im Anschluss daran lag das Schwergewicht der weltweiten Entwicklung autonomer Straßenfahrzeuge auf ‚Confirmation-type‘ Vision mit 360° Rundum-Laser-Scannern auf dem Dach. Die verfügbaren Rechenleistungen erlaubten hier mit relativ geringem Aufwand in der Software-Entwicklung akzeptable Lösungen für einfache Anwendungen im normalen Straßenverkehr (s. z. B. [29]). Seither ist das Gebiet ‚Autonomes Fahren‘ auch dank der Entwicklungen in den Bereichen Radar- und Laser-Technologie weltweit explosionsartig ausgedehnt worden. Nach den jetzt schon verfügbaren Fahrer-Assistenzsystemen werden bald auch die ersten Fahrzeuge mit voll autonomen Fähigkeiten auf Schnellstraßen auf dem Markt angeboten werden.

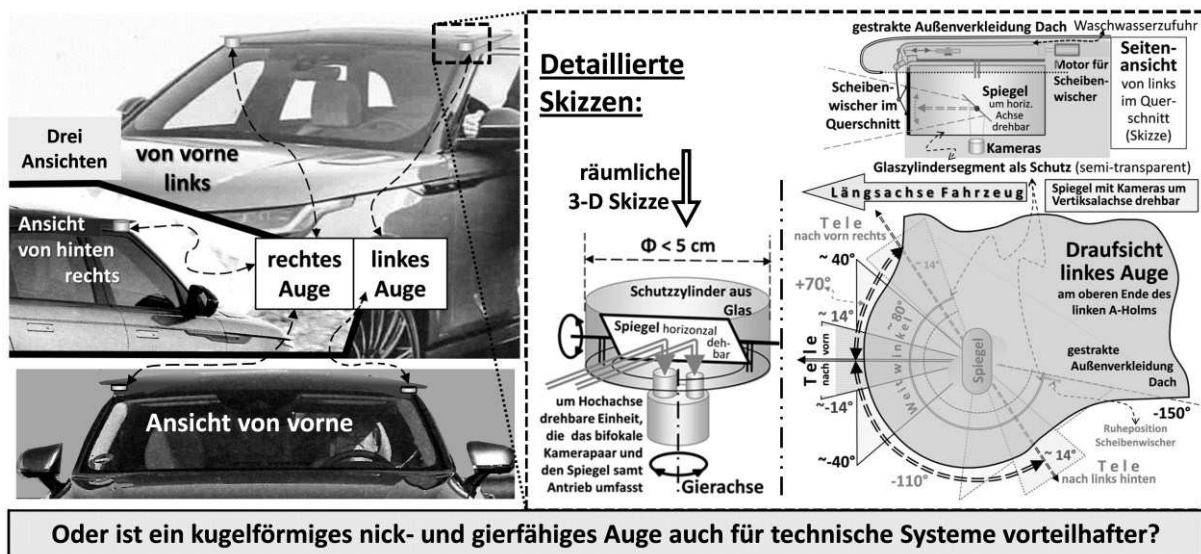


Bild 9: Kann so das langfristige Entwicklungsergebnis sehender Bodenfahrzeuge mit der Leistungsfähigkeit des menschlichen Gesichtssinns aussehen? Redundanz durch zwei Augen, dadurch auch binokulares Stereosehen, Blickstabilisierung durch inertielle Rückkopplung, Fähigkeit zur hochauflösenden aktiven Objektverfolgung.

In Anbetracht der höchstentwickelten Sehsysteme im biologischen Bereich drängt sich die Frage auf, ob nicht langfristig auch maschinelle Sehsysteme vom Pfadfinder-Typ eine leistungsfähigere und wirtschaftlichere Alternative gegenüber den bisher favorisierten vom Typ Bestätigungssehen sein könnten. Bild 9 zeigt als grobe Skizze eine von vielen möglichen Ausformungen. Eine Abschätzung, wie viele Jahrzehnte wir von einer solchen Entwicklungsstufe entfernt sind, ist derzeit wohl kaum möglich. Einige Untersuchungen in diese Richtung werden bei [30] durchgeführt.

Literatur:

- [1] Wikipedia: Wheels: History
- [2] Dickmanns E.D.: Vision for Ground Vehicles – History and Prospects. Int. J. of 'Vehicle Autonomous Systems' (IJVAS), Vol.1, No.1, 2002, pp. 1 – 44.
- [3] Roland, A., Shiman, P.: Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983–1993. MIT Press 2002
- [4] Dickmanns, E.D.: 4-D-Dynamic Scene Analysis with Integral Spatio-Temporal Models. 4th International Symposium on Robotics Research, Santa Cruz 1987. In: Bolles R.C.; Roth B.: Robotics Research. MIT Press, Cambridge 1988, S. 311-318.

- [5] Wünsche H.-J.: Erfassung und Steuerung von Bewegungen durch Rechnersehen. UniBw München, LRT Diss. 1988
- [6] Zapp, A.: Automatische Straßenfahrzeugführung durch Rechnersehen. UniBw München, LRT Diss. 1988
- [7] Dickmanns, E.D.: Dynamic Vision for Perception and Control of Motion. Springer 2007
- [8] Dickmanns E.D.; Graefe V. 1988: a) Dynamic monocular machine vision. Machine Vision and Applications, Springer International, Vol. 1, pp 223-240. b) Applications of dynamic monocular machine vision. (ibid), pp 241-261
- [9] Mysliwetz B.: Parallelrechnerbasierte Bildfolgeninterpretation zur autonomen Fahrzeugführung. UniBw München, LRT Diss. 1990
- [10] Vier Beiträge in Masaki (ed): Proc. of Int. Symp. on Intell. Vehicles'94, Paris, Oct. 1994:
 - a) Dickmanns E.D.; Behringer R.; Dickmanns D.; Hildebrandt T.; Maurer M.; Thomanek F.; Schiehlen J.: The Seeing Passenger Car 'VaMoRs-P'. pp 68-73 Abstract , pdf
 - b) Thomanek F.; Dickmanns E.D.; Dickmanns D.: Multiple Object Recognition and Scene Interpretation for Autonomous Road Vehicle Guidance. pp. 231-236 Abstract
 - c) Von Holt: Tracking and Classification of Overtaking Vehicles on Autobahnen. pp 314-319
 - d) Schiehlen J.; Dickmanns E.D.: A Camera Platform for Intelligent Vehicles. pp 393-398
- [11] Brüdigam C.: Intelligente Fahrmanöver sehender autonomer Fahrzeuge in autobahn-ähnlicher Umgebung. UniBw München, LRT Diss. 1994
- [12] Schiehlen J.: Kameraplattformen für aktiv sehende Fahrzeuge. UniBw München, LRT Diss. 1995
- [13] Thomanek F.: Visuelle Erkennung und Zustandsschätzung von mehreren Straßenfahrzeugen zur autonomen Fahrzeugführung. UniBw München, LRT Diss. 1996
- [14] Behringer R.: Visuelle Erkennung und Interpretation des Fahrspurverlaufes durch Rechnersehen für ein autonomes Straßenfahrzeug. UniBw München, LRT Diss. 1996
- [15] Müller N. 1996: Autonomes Manövrieren und Navigieren mit einem sehenden Fahrzeug. UniBw München, LRT Diss. 1996
- [16] Dickmanns Dirk: Rahmensystem für die visuelle Wahrnehmung veränderlicher Szenen durch Computer. UniBw München, Fak. Informatik, Diss. 1997
- [17] Dickmanns E.D.: An Expectation-based, Multi-focal, Saccadic (EMS) Vision System for Vehicle Guidance. In Hollerbach and Koditschek (eds.): ‚Robotics Research‘ (The Ninth Symposium), Springer-Verlag, 1999
- [18] Sechs Beiträge in zu EMS-Vision in den Proceedings of the Internat. Symp. on Intell. Vehicles (IV'2000), Dearborn, (MI, USA), 2000 Oct. 4-5: {DED = Dickmanns, E.D.}

- a) Gregor, R., Lützeler, M., Pellkofer, M., Siedersberger, K.H., DED: EMS-Vision: A Perceptual System for Autonomous Vehicles. pp 52-57
 - b) Pellkofer, M., DED: EMS-Vision: Gaze Control in Autonomous Vehicles. pp 296-301
 - c) Lützeler, M., DED: EMS-Vision: Recognition of Intersections on Unmarked Road Networks. pp 302-307
 - d) Gregor, R., DED: EMS-Vision: Mission Performance on Road Networks. pp 468-473;
 - e) Hofmann, U.; Rieder, A., DED: EMS-Vision: Application to Hybrid Adaptive Cruise Control. pp 468-473
 - f) Siedersberger K.-H., DED: EMS-Vision: Enhanced Abilities for Locomotion.
- [19] Albus J.S.: 4-D/RCS reference model architecture for unmanned ground vehicles. Proc. of the International Conference on Robotics and Automation, San Francisco, 2000, April 24-27
- [20] Albus J.S., Meystel A.M.: Engineering of Mind. – An Introduction to the Science of Intelligent Systems. Wiley Series on Intelligent Systems, 2001
- [21] Lützeler M.: Fahrbahnerkennung zum Manövrieren auf Wegenetzen mit aktivem Sehen. UniBw München, LRT Diss. 2002
- [22] Gregor R.: Fähigkeiten zur Missionsdurchführung und Landmarkennavigation. UniBw München, LRT Diss. 2002
- [23] Pellkofer M.: Verhaltensentscheidung für autonome Fahrzeuge mit Blickrichtungssteuerung. UniBw München, LRT Diss. 2003
- [24] Von Holt V.: Integrale Multisensorielle Fahrumgebungserfassung nach dem 4-D Ansatz. UniBw München, LRT Diss. 2004
- [25] Siedersberger K.H.: Komponenten zur automatischen Fahrzeugführung in sehenden (semi-) autonomen Fahrzeugen. UniBw München Diss. 2004
- [26] Hofmann U.: Zur visuellen Umfeldwahrnehmung autonomer Fahrzeuge. UniBw München, LRT Diss. 2004
- [27] **www.dyna-vision.de**
- [28] Dickmanns E.D.: **<http://www.youtube.com/watch?v=YZ6nPhUG2i0>** ; a one-hour slide- and video-clip talk under the heading “Dynamic Vision as Key Element for Artificial General Intelligence” (Tech-Talk 2011 at AGI, Mountain View, CA)
- [29] Jährliche Konferenz: ‚International Symposium on Intelligent Vehicles‘, (kurz IV’{letzte_zwei_Ziffern_der_Jahreszahl}: IV’xy; seit IV’92 für das Startjahr 1992)
- [30] **www.unibw.de/tas**, Website TAS, LRT8, UniBw München