

**Forschungsberichte der  
Technischen Fakultät  
Abteilung Informationstechnik**

**Autonomie für situierte Robotersysteme  
– Stand und Entwicklungslinien –**

Alois Knoll

Report 99-06

**Impressum:**

Herausgeber:

Robert Giegerich, Alois Knoll, Helge Ritter, Gerhard Sagerer,  
Ipke Wachsmuth

Anschrift:

Technische Fakultät der Universität Bielefeld,  
Abteilung Informationstechnik, Postfach 10 01 31, 33501 Bielefeld

ISSN 0946-7831

---

## Zusammenfassung

Für zukünftige Handhabungs-, Service- und humanoide Roboter ist die Verbindung der motorischen Komponente mit der Kommunikationskomponente (dem „Situieren Künstlichen Kommunikator“, *SKK*) eine wesentliche Voraussetzung, um diese komplexen Systeme bedienbar zu machen. Wir beschreiben mögliche Entwicklungslinien im Rahmen des Baufix-Szenarios des Sonderforschungsbereichs 360, entlang derer schrittweise die kognitiven Fähigkeiten *Erkennen*, *Lernen* und *Planen* verbessert werden können.

---

# 1 Gegenstandsbereich

Das reibungslose, intuitiv vom menschlichen Instrukteur steuerbare Ineinandergreifen von Anweisungsfolgen, Montagesequenzen und Umweltwahrnehmung erfordert auf der Seite des „Situiereten Künstlichen Kommunikators“ *SKK* ein ausgesprochen komplexes Zusammenspiel sensorischer, kognitiver und manipulativer Fähigkeiten. Der Instrukteur erwartet zum einen, daß der *SKK* die gewünschten Operationen schritthaltend mit den Anweisungen ausführt, daß also der Einsatz von Ressourcen (Sensoren, Aktoren, Prozessoren) mit Blick auf rechtzeitige (Ergebnis-)Verfügbarkeit geplant wird, einschließlich situativ adäquater Reaktion im Fehlerfall. Zum anderen sollen auch längere Sequenzen, die der Mensch nach seiner eigenen Erfahrung als Einheit empfindet, vollständig autonom ablaufen können (ohne ständige Rückfragen oder die Erwartung weiterer Instruktionen).

Bei der Montage größerer Einheiten (Aggregate) ist es weder wünschenswert noch möglich, daß ein menschlicher Instrukteur sämtliche sensorischen und motorischen Einzeloperationen durch natürlichsprachliche Anweisungen einzeln steuert. Statt einer objektreferenzierenden natürlichsprachlichen Anweisungsfolge wie „Nimm die Dreilochleiste“, „Greif die blaue Schraube mit dem anderen Arm“, „Greif jetzt um“, „Steck sie jetzt durch das mittlere Loch“, „Zieh sie fest an“ . . . sollte also einfach eine aggregatreferenzierende Instruktion der Form „Bau mir ein *Leitwerk* mit blauer Schraube und gelbem Würfel“ gegeben werden können. Die Ausführung dieser komplexen Anweisung sollte ohne Rückgriff auf bereits bestehende oder im voraus zu entwerfende Modelle der Aggregate möglich sein. Dazu muß das (Teil-)Aggregat im Dialog von Mensch und Situieretem Künstlichen Kommunikator (*SKK*) zu einem früheren Zeitpunkt konstruiert und durch den Menschen benannt worden sein. Eine veränderte Ausgangssituation (Lage der Teile, sensorische Ausstattung, zur Verfügung stehende Aktoren) soll zu keiner Einschränkung der Autonomie führen. Sehr wünschenswert ist ferner die Fähigkeit des *SKK*, Bedeutungen neuer Handlungsverben und Objekte zu lernen.

Der *SKK* wird damit also ein System, das durch Erfahrung bei der interaktiven Montage sukzessive lernt und auf immer abstraktere Anweisungen situativ reagieren kann. Unabdingbare Voraussetzung dafür ist, daß die *innere Struktur* der Aggregate, d.h. die wesentlichen Parameter, die ihre Konstruierbarkeit bestimmen (Kräfte, Montagepunkte, Zugänglichkeiten, Perspektiven, . . . ), im Zuge der Ausführung der Montageoperationen mitgelernt wird. Das unterscheidet diesen Ansatz *qualitativ* von anderen Vorhaben, bei denen versucht wurde, allein durch Vormachen und (visuelles) Abschauen ein Robotersystem auf die Durchführung einer Montageoperation zu trainieren.

Wie eingangs angedeutet, wäre der menschliche Instrukteur schon während der unter Umständen recht langsam ablaufenden Lernphase einfacher Operationen des *SKK* vollkommen überfordert, wenn er das Zusammenspiel der sensorischen und aktorischen Fertigkeiten sowie die Planung der Bereitstellung von Verarbeitungskapazitäten in allen Details steuern sollte. Umso weniger kann er dies bei den komplexeren Instruktionen: nur der *SKK* selbst vermag mit dem Wissen um die jeweiligen Randbedingungen und um die situativen Erfordernisse eine verschränkte Planung von elementaren Montageschritten bzw. Betriebsmitteln (Aktoren, Sensoren, Prozessoren) vorzunehmen. Solche Randbedingungen bzw. Erfordernisse sind unter anderen:

- die Wirkungsmöglichkeiten der Aktoren im jeweiligen Szenenkontext,
- die sensorisch erfaßten Meßgrößen beim Montageprozeß und die zu erwartenden Störeinflüsse,
- die zu erwartende Qualität der Erfassung der verschiedenen Umwelt- und Montageprozeßkenngrößen,
- die Fähigkeiten der optischen Sensoren, aus ihren unterschiedlichen Sichten Umweltausschnitte bzw. -ereignisse aus einer bestimmten räumlichen Lage heraus zu registrieren,
- die augenblickliche Verfügbarkeit kognitiver Verarbeitungsleistung und weiterer Module (Fehlererkennung, Kollisionsvermeidung, . . . ) bzw. deren Priorisierung.

Die komplexen Abläufe des Multi-Sensor/Multi-Aktor-Systems müssen geplant und zur Laufzeit in Abhängigkeit vom situativen Kontext kontrolliert werden können. Die lernfähige Aktivitäts- und Ereignissteuerung muß also

- eine aufgabenbezogene Ressourcenplanung und -steuerung unter Einbezug aller situativen Einflußgrößen (mechanische Randbedingungen, Sensorrauschen, Verdeckungen, . . . ) für die Aktoren vornehmen können,
- damit verbunden über den Einsatzort und die aufgabengerechte *Informationsfusion* für alle sensorischen Einheiten entscheiden,
- den impliziten und explizit geäußerten Intentionen des Instrukteurs Folge leisten und seine Wünsche inferieren können (Bedeutungsverstehen und vorausschauende Planung),
- im Fehlerfall eine situierte Reaktion des *SKK* zur Behebung ermöglichen (wiederaufgreifende Planung für nicht vorhersehbare Aktor-/Objektkonstellationen),

- konstraintuitives Verhalten des *SKK* einschränken (Erzielung von *Aktionsvernunft*) und dabei gleichzeitig kritische Systemparameter berücksichtigen, z.B. Vermeidung kinematischer Singularitäten, Kollision und Ausgleich von Sensorbegrenzungen (Genauigkeit, räumlicher Blickwinkel) durch Sensordatenfusion,
- für den Fall, daß sich der Instrukteur mit dem *SKK* in eine Sackgasse begeben hat, nach Rückfrage eine gezielte Demontage bis zu definierten Wiederaufsetzpunkten vornehmen,
- zur Prävention solcher Sackgassen eine simulierende, dem eigentlichen Montagegeschehen vorgeschaltete Planung durch Emulation bzw. Probehandeln samt kognitiv adäquater Präsentation für den Instrukteur ausführen.

Angestrebt wird also, den *SKK* dazu zu befähigen *autonom lernend* auf der Basis

- einer genauen Umweltbeobachtung und Objektlageerkennung,
- der vielfältigen Interaktionen mit dem Instrukteur,
- der Nutzung generischen und domänenspezifischen Planungswissens sowie
- der Aufzeichnung und Analyse der konkreten Durchführung der Montageoperationen

über immer *längere Strecken*, für eine immer *größere Zahl von Aggregaten*, für immer *generellere Ausgangs- und Zwischensituationen* und für einen immer *größeren Satz von Handlungsverben* autonom kognitive Leistungen zu erbringen und zu handeln. Schließlich sollte er in die Lage versetzt werden, sich durch Lernen in Interaktion mit dem Instrukteur schrittweise an neue Domänen mit zusätzlichen Operationen, Montageobjekten und ihre Ansichten, korrespondierenden Handlungsverben und (Fehler-)Situationen anzupassen.

Mit anderen Worten: die Planungs-, Datenfusions- und Lernkomponenten machen aus dem auf sprachliche Anweisung hin nur atomare Instruktionen<sup>1</sup> ausführenden Konstrukteur einen „mit- und vorausdenkenden“ *Handlungspartner, der aus Erfahrung und Fehlern lernen und dazu Nach-Sehen, Nach-Denken, Nach-Fragen kann*. Instruktionen können dabei auf allen Abstraktionsniveaus gegeben werden: auf der höchsten Ebene („Bau ein Flugzeug“) entsprechen sie einem komplexen, autonom ausführbaren Programm des *SKK* (analog zum „High-Level-Programming“ gemäß [LR99]), auf unterster Ebene („Zieh ein wenig stärker“) kontrollieren sie die autonome Umsetzung reaktiver Aktionsmuster. Man kann also von der Erzielung von *Autonomie des SKK auf allen Ebenen* sprechen (*any-level-autonomy*), deren „Reichweite“ durch die Interaktion im Handlungsdialog angestoßen bzw. gesteuert wird.

Der *SKK* selbst muß zur reibungslosen Annahme von Instruktionen auf allen diesen Ebenen bestimmen,

- welche Module zur Planung zu aktivieren sind,
- welche Betriebsmittel angefordert werden müssen,
- wann die menschliche Führung zu akzeptieren, zu ergänzen oder temporär zu ignorieren ist,
- in welchen Situationen Rückfragen erforderlich sind – vor allem aber, wie sie so zu stellen sind, damit der Instrukteur möglichst viel von seinen Absichten preisgibt, die Kommunikation also minimiert werden kann.

Außerdem sollte er selbsttätig vom „Ausführungsmodus“ in den „Lernmodus“ wechseln können, d.h. er sollte in der Lage sein, zu erkennen, wann eine ihm dargebotene Information in einer bestimmten Situation „Innovationscharakter“ hat, d.h. zur Erweiterung des Wissens beitragen kann.

Offensichtlich ist die Erzielung von Autonomie über längere Strecken eine große Herausforderung. Wir stellen im folgenden den Stand der Forschung auf den Gebieten dar, deren Beherrschung Voraussetzung für die Realisierung eines solchen autonomen Systems ist. Am Schluß der Betrachtung steht jeweils eine kurze Kritik des Stands *unter unserer Aufgabenstellung*.

Danach wird anhand interessanter Beispiele aus dem SFB-Szenario aufgezeigt, welche Aufgabenstellungen aus den drei Bereichen Planung, Fusion und Lernen einen Schritt hin auf das angestrebte Ziel zu führen versprechen (siehe Abschnitt 3).

## 2 Stand der Forschung

Die Realisierung einer sensorbasierten Planungskomponente für stationäre kooperierende Montageroboter als Grundlage eines *SKK*, der die im vorigen Abschnitt genannten Anforderungen erfüllt, erfordert zunächst die Integration und Erweiterung klassischer Techniken zur Geometrie-basierten Montage-, Trajektorien- und Griffplanung für das zu bearbeitende Szenario. Im Mittelpunkt stehen die Themenbereiche Sensoreinsatzplanung, Datenfusion, Lernen von diskreten Montagesequenzen, Generalisierung von Lernergebnissen sowie Transfer von Fertigkeiten vom Menschen auf den Roboter bzw. von einem Roboter auf einen anderen im Sinne einer Generalisierung.

<sup>1</sup>Instruktionen sind atomar bzw. Handlungsprimitiva, wenn sie aus nur einem Handlungsverb bestehen und sich nur auf freistehende Einzelobjekte beziehen. Für den menschlichen Instrukteur bilden sie damit eine Einheit, obwohl sie aus Sicht des *SKK* bereits eine umfangreiche Abfolge sensorischer, aktorischer oder sensomotorischer Einzelaktionen darstellen, die ihrerseits ebenfalls eine Planung erforderlich machen. Beispiel: „Schraub“ ist eine Sequenz aus Lochfinden, Einschnidtest, Verklemmungsprüfung, Eindrehen, Umgreifen, fortgesetztem Eindrehen, Prüfung des Anzugsmoments, Loslassen.

## 2.1 Planung für die Aktorik

Die Planungsschritte, die dem Zusammenbau eines Aggregats vorausgehen (oder sie begleiten), sind verschiedenen Abstraktionsebenen zuzuordnen. Wir gehen kurz auf den Stand der Forschung auf diesen Ebenen ein, wobei wir mit der abstraktesten „roboterfernsten“ Ebene, der Montagesequenzplanung, beginnen<sup>2</sup>:

### 2.1.1 Montagesequenzplanung.

Aufgabe der Montagesequenzplanung ist es, festzulegen, wann welche Operation mit welchem Objekt durchgeführt wird. Dazu werden Unterziele festgelegt und zur Erfüllung jedes Unterziels bestimmt, ob seine Vorbedingungen erfüllt sind. Solche Vorbedingungen sind beispielsweise die Erreichbarkeit eines zu montierenden Objekts durch den Manipulator, die korrekte Lage des Teilaggregats, an die ein Objekt anmontiert werden soll, die Verfügbarkeit eines geeigneten Manipulators, usw. Die ersten Arbeiten in diesem Bereich stammen aus den frühen siebziger Jahren und hatten typischerweise Aufgaben aus der *blocks-world* zum Gegenstand; STRIPS [FN71] ist sicher der bekannteste dieser Planungsformalisten. Innerhalb dieser Domäne waren Aufgaben zu lösen, die durch eine eindeutig vorgegebene Ausgangssituation und ein ebenso eindeutig spezifiziertes Ziel gekennzeichnet sind (Erzeugung eines Plans zum Aufeinanderstapeln von unterscheidbaren Würfeln in eine gewünschte Konfiguration). Obwohl zunächst von geringer praktischer Relevanz, waren diese Ansätze die Grundlage für verschiedene plangenerierende Systeme für den Fertigungsbereich.<sup>3</sup>

In heutigen CAD-Systemen wäre es im Prinzip möglich, die Montagesequenzplanung<sup>4</sup> trotz ihrer hohen algorithmischen Komplexität automatisch vorzunehmen, weil die Geometrie aller Teile samt ihrer Materialbeschaffenheit vollständig modelliert ist. Mit diesem Objektwissen wäre zudem eine Planung von Annäherungsbewegungen beim Aggregatbau, die Simulation dynamischer Kräfte bei der Ausführung von Bewegungen, Einbezug von Toleranzen und sogar die Erzeugung von Programmen zur Robotersteuerung möglich. Dennoch ist festzuhalten, daß dies alles nicht oder nur in geringem Maße geschieht: auch in Industriebetrieben, die die Konstruktion vollständig über CAD-Systeme abwickeln, wird die Erzeugung von Montagesequenzen, ja sogar die Codierung der entsprechenden Roboterprogramme bis heute vom Menschen vorgenommen.<sup>5</sup>

Dies hat im wesentlichen drei Gründe: Erstens bestehen die Steuerungsprogramme für die Roboter nur zum geringeren Teil aus den eigentlichen Bewegungskommandos; den größeren Teil nimmt die Generierung von Prozeß- und Überwachungsdaten ein, und hierfür fehlen den CAD-Systemen die nötigen Erzeugungs- bzw. Integrationskomponenten. Wesentlich wichtiger ist jedoch ein zweiter Grund, der darin besteht, daß im CAD-System keine *benannte* Zielspezifikation für das endgültige Aggregat möglich ist, d.h. der gewünschte Endzustand kann nicht anders spezifiziert werden, als durch das „manuelle“ Fügen verschiedener Teile zum gewünschten Aggregat<sup>6</sup>. Wenn dieses Zusammenfügen aber die unabdingbare Voraussetzung zur Beschreibung des Aggregats ist, dann entfällt natürlich die Notwendigkeit und die Möglichkeit, die Montagesequenz automatisch zu finden. Drittens gibt es keine Möglichkeit, Konstruktionswissen<sup>7</sup> auf neue Konstellationen zu übertragen und damit in neuen Konstruktionssituationen zumindest Vorschläge für Sequenzen machen zu können.

Bei den reinen *Planungssystemen* sind im Zusammenhang mit der Montageplanung verschiedene Ausprägungen zu unterscheiden:

- Systeme zur Fertigungsplanung für miteinander zu synchronisierende Maschinen auf hoher Abstraktionsebene, also z.B. ganze Fabrikhallen. Zum Einsatz kommt hier die Theorie der *discrete event systems*<sup>8</sup> operationalisiert in Form von Zustandsstabellen, endlichen Automaten oder Petri-Netzen samt der dazugehörigen Werkzeuge zur Manipulation (ein typisches Beispiel gibt [Zha99a]). Während diese Ansätze sich hauptsächlich zur deliberativen Planung „im voraus“ eignen, können für dynamische Fälle (sich ändernde Maschinenstatus, Ausfall von Maschinen mit Umleitungsmöglichkeit, Einphasen neuer Bearbeitungszentren im laufenden Betrieb) Multi-Agenten-Systeme mit Kommunikation über das Vertragsnetzprotokoll

<sup>2</sup>In der industriellen Anwendung liegt über der Montagesequenzplanung noch die Planung für die Materialverteilung und Einsatzplanung für alle zusammenarbeitenden Aktoren einer größeren Produktionseinheit (also z.B. einer Fabrikhalle), siehe etwa [LB91, Rev99]. Wir gehen auf die hier entstehenden Fragestellungen nicht ein.

<sup>3</sup>Wobei am Rande bemerkt sei, daß hier auch die schon früh entwickelten Techniken zur Planung in hierarchisch untergliederten Suchräumen (beginnend mit ABSTRIPS [Sac74]), zum Planen mit Einschränkungen und *least commitments* ([Wei94], implementiert etwa in MOLGEN [Ste81]) zum Einsatz kommen. Interessant ist in diesem Zusammenhang auch [Wil88].

<sup>4</sup>Dies betrifft mindestens jedoch die Erzeugung sogenannter Montagevorranggraphen, die die Relationen von Objekten repräsentieren und ggf. über unterschiedliche Traversierungen alternative Wege zu einem vollständigen Aggregat aufzeigen. Das bislang leistungsfähigste System, welches auf einfachen Planungstechniken unter Verwendung von Einschränkungen ein Objekt mit 472 Einzelteilen plant, beschreibt [JWC98] – man vergleiche diese Zahl mit der Menge von Objekten, aus denen das Baufix-Flugzeug besteht.

<sup>5</sup>Über erste Ansätze, die Sequenzplanung und die Programmierung zu verknüpfen, wird in [YG99] berichtet; [Zha99b] präsentiert einen Ansatz für ein hybrides Planungs- und Robotersteuerungssystem auf der Basis von Petri-Netzen als *discrete-event-system* zur Spezifikation der Montagesequenz, verknüpft mit einem Steuerungssystem mit kontinuierlichem Zustandsraum.

<sup>6</sup>In gewisser Weise ist das in der Robotik oft umgekehrt: Man weiß, was man montiert haben möchte, aber die Einzelteile sind nur ungenügend beschrieben oder sensorisch erfäßbar. Siehe dazu die programmatischen Ausführungen in [Smi92].

<sup>7</sup>Erfahrungswissen des Menschen oder konzeptuelles/situatives Wissen, welches das System bei vorherigen Konstruktionsvorgängen extrahiert hat.

<sup>8</sup>Einführende Überblicke sind z.B. [RW89, Ho92].

[SD81, Fox81] vorteilhaft eingesetzt werden [Bak88, Par89, OHB<sup>+</sup>99, BA99]. In beiden Fällen kann aufgrund des hohen Abstraktionsgrads Sensorik nur in der Form des Auftretens eines bestimmten Ereignisses („Sensor hat Teil  $X$  erkannt“) einbezogen werden.

- Systeme zur Planung von Aktionen einzelner Manipulatoren unter Unsicherheit, d.h. Berücksichtigung des dynamischen Umweltzustandes. Die meisten Planungssysteme gehen davon aus, daß der Ausgangszustand vollständig bekannt ist, daß die umgebende Welt statisch ist und daß nur der Planausführer Zustandsänderungen durchführen kann. Systeme zum Einbezug unsicherer Information, dynamischer Umwelten und nur zeitweise zur Verfügung stehender Betriebsmittel in Planungsprozesse sind erst seit kurzem Gegenstand der Forschung. Grundlagen sind einfache probabilistische Modellierungen von Zuständen in einer unsicheren Domäne [DHW93] oder aber stochastische Prozesse (partiell beobachtbare Markowsche Entscheidungsprozesse [Ast65, DL95]). Mit letzteren lassen sich sowohl Unsicherheit bezüglich des Ergebnisses, das eine Handlung hat, als auch die nur unvollständige Beobachtbarkeit des Umweltzustandes modellieren. Das Planungsproblem wird als Kostenfunktion modelliert, wobei jeder Planungsschritt zwar Kosten verursacht, aber gleichzeitig einen Wegfortschritt in Richtung auf einen Zielzustand bringt (eine Übersicht gibt [BDH99]). Hingewiesen sei auf [LL98]; hier wird untersucht, welcher Endzustand bei einer gegebenen Menge minimaler Aktionen, die jeweils ein binäres sensorisches Ergebnis liefern, von einem Startzustand ausgehend erreicht werden kann. Schließlich sind auch für diese Aufgabenstellung Versuche mit dem Vertragsnetzprotokoll gemacht worden [FPWK99].
- Systeme zur Erfüllung von Echtzeitforderungen. Die Einhaltung von Zeitvorgaben kann durch Reduktion des Suchraums über die Abstraktion von Details erfolgen, durch inkrementelle Planung, d.h. Beschränkung auf die Auswahl und direkte Ausführung von Handlung<sup>9</sup> auf Basis einer domänenspezifischen Heuristik [BLG97], oder durch Kopplung von schnellen, reaktiven Planern mit langsameren, deliberativen Subsystemen, wobei das deliberative Subsystem die Parameter des reaktiven im langsamen Zeittakt verändert [GMKB97, BCG<sup>+</sup>99, ZK99b].

**Kritik.** Planungssysteme, die Pläne für einzelne Manipulatoren unter einem gewissen Maß an Unsicherheit erzeugen können, sind im Prinzip verfügbar. Es bleibt aber festzuhalten, daß bei dem hier interessierenden Thema der Planung unter Unsicherheit und bei beschränkter Zeit für Teams von Agenten (MAP: Multi-Agent-Planning), bei denen ein *Zwang zur Zusammenarbeit* vorliegt, weil sie nur *gemeinsam* Ziele erreichen können, nahezu keine Arbeiten vorliegen<sup>10</sup>. Eine Ausnahme ist [BB97]; hier wird eine STRIPS-artige Repräsentation für nebenläufige, verkettete Aktionen eingeführt und auf ein einfaches Beispiel angewandt. Im übrigen fällt auf, daß die meisten Arbeiten zur dynamischen Planung in der Robotik der Wegplanung für mobile Roboter, speziell unter Vermeidung von Kollisionen, gewidmet sind. Eine Übertragung auf die Problematik der Montage ist nur sehr begrenzt möglich. Unter Fehlerbehandlung wird üblicherweise die erneute Ausführung ab einem Wiederaufsetzpunkt bzw. Umleitung verstanden (siehe z.B. [Wu99]), keinesfalls jedoch die gezielte Demontage mit der streckenweisen Vertauschung von Ausgangs- und Zielzustand<sup>11</sup>. Schließlich gibt es bislang überhaupt keine implementierten Systeme, die unter den oben genannten Forderungen arbeiten und zusätzlich noch das Eigenverhalten eines Roboters – etwa über ein geometrisches oder gar dynamisches Selbstmodell – mit ins Kalkül ziehen, um Emulation oder Probehandeln wirklichkeitsnah<sup>12</sup> abbilden zu können.

### 2.1.2 Bewegungs- und Griffplanung.

Aufgabe der klassischen Bewegungsplanung für stationäre Roboter [FGL87] ist die Erreichung einer hohen *Armgeschwindigkeit* hin auf einen Zielpunkt bei gleichzeitig hoher *Bahntreue* und möglichst geringer mechanischer Belastung der Gelenkmotoren sowie die Vermeidung von Kollisionen mit statischen oder bewegten Objekten<sup>13</sup>. In jüngster Zeit ist mit einer stärkeren Einbeziehung von Sensoren und der daraus folgenden dynamischen Variabilität der Aufgaben darüber hinaus die Problematik der Singularitätenvermeidung immer stärker in den Vordergrund gerückt [Zha95, Llo98]. Deren Ziel ist, daß der Roboter auf seiner Bahn nicht über eine kinematisch unerreichbare oder aber über eine nicht eindeutig einer bestimmten Gelenkkonfiguration zuzuordnende Stellung geführt wird, weil dies unter Umständen unbegrenzte Gelenkgeschwindigkeiten bzw. -beschleunigungen zur Folge hätte.

Gegenwärtig konzentriert sich das Interesse auf die Bewegungsplanung mobiler Roboter unter Einbezug von Sensorik, der ein wesentlicher Teil der Beiträge zu den einschlägigen Konferenzen gewidmet ist; zunehmendes Interesse finden kooperierende

<sup>9</sup>Siehe zu diesem Problemkreis auch [Mae90, EHW<sup>+</sup>92, Nil94, PC96, KQW97, MS97].

<sup>10</sup>Nichtlineare Planung von (nebenläufigen) Aktionen allein reicht dazu nicht aus, etwa wie in [MR91], weil damit das Problem der Aufgabenaufteilung nicht gelöst ist. Ein Beispiel ist das gemeinsame Tragen eines Stabes bei gleichzeitiger Forderung nach permanenter horizontaler Ausrichtung (Standardbeispiel für kooperierende, kraftgeregelte Roboter).

<sup>11</sup>Ein erster Ansatz dazu wird in [LM99] vorgestellt; hier handelt es sich um das Lösen kraftschlüssiger Verbindungen. Die Demontage erfordert dort also eine zusätzliche Kraftanwendung.

<sup>12</sup>Um dem Instrukteur das „Hineindenken“ in das dynamische Verhalten der Roboterarme zu erleichtern, sollte angestrebt werden, deren dynamisches Verhalten dem menschlichen Vorbild anzunähern. Beobachtungen dieses Verhaltens und Vorschläge zu seiner technischen Umsetzung finden sich in [Hol90, Bur96, HSFS99].

<sup>13</sup>Die Standard-Einführung ist [Lat91], eine neuere Arbeit über die Bewegungsplanung für Industrieroboter ist [LC97]; in [YX96] wird ein Ansatz auf der Basis genetischer Algorithmen vorgeschlagen und simuliert.

mobile Roboter [PO98, ŠO98]. Dies erklärt sich damit, daß die Probleme der Bewegungsplanung und Kollisionsvermeidung bei stationären Robotern zwar nicht für alle denkbaren, aber doch für die wichtigsten Einsatzfälle gelöst sind. Es existieren frei verfügbare Software-Bibliotheken, die für die Praxis ausreichend stabil, schnell und anpassungsfähig sind [Mir97]. Für die in unserem Zusammenhang sehr wichtige Singularitätenvermeidung gibt es ein ebenfalls leistungsfähiges Software-Paket [Sch98]. Bei der *Griffplanung* ist festzuhalten, daß Planung im Sinne der Lösung eines kombinatorischen Problems nur bei mehrfingrigen Händen erforderlich ist (Bestimmung von Griffposturen und optimalen Griffpunkten, Berechnung von Fingerbewegungen), siehe dazu [MLS94, BFH99, Man99]. Bei den üblichen Zweifingergreifern lassen sich zwar auch einige interessante Aufgabenstellungen, speziell zur Auswahl optimaler Greifpunkte finden [SLG<sup>+</sup>99]. Bei nicht exakt bekanntem Objekt und der Notwendigkeit der sensorischen Beeinflussung der Bewegung benutzt man in der Technik des *visual servoing* einen rein reaktiven Ansatz [Cor93, HD98].

## 2.2 Sensoreinsatzplanung

Die Aufgabe der Sensoreinsatzplanung besteht in der Aktivierung von Sensoren und in der Einstellung ihrer Parameter mit dem Ziel der Lösung einer Erkennungsaufgabe bei Erreichung vorgeschriebener Erkennungsqualität und ggf. einer Obergrenze für entstehende Kosten. Solche Parameter sind der Sensorstandort bzw. der Sichtwinkel (bei bewegbaren bzw. mitbewegten Sensoren), die Empfindlichkeit und der fokussierte Bereich (Festlegung einer *region-of-interest*). Es geht hier darum, diese Parameter in der Wechselwirkung zwischen Erkennungsaufgabe, Objektstruktur und Sensorcharakteristik möglichst weitgehend automatisch zu bestimmen und dementsprechend einen Plan zu generieren, wann und wo welcher Sensor mit welchen Parametern zu aktivieren ist. Eine solche Planung ist die Voraussetzung für eine sinnvolle Fusion der redundanten Informationen aus den einzelnen Sensoren.

Die Anzahl von Arbeiten zu diesem Thema ist noch relativ klein, und sie beschränkt sich auf die Planung von optischen Sensoren, d.h. Laser-Entfernungsmesser und Kameras bzw. die dazugehörigen Lichtquellen. In der Übersicht [TAT95] werden als Eingangsgrößen für generische Sensorplanungssysteme Sensormodelle (einschließlich der Wirkung von Beleuchtung), Objektmodelle und die Beschreibung der Erkennungsaufgabe (Objekterkennung, Szenenrekonstruktion, Merkmalsextraktion) genannt. Ausgangsgrößen sind die Blickrichtung der Kamera, ihre optischen Parameter und die Strahlrichtung der Beleuchtungsquelle. In [AAT99] wird dies wieder aufgegriffen; auch hier geht die Systematik jedoch nicht über die Aufstellung einiger Heuristiken hinaus.

In [GI94] wird untersucht, welche Ansichten eines Passagierflugzeugs zu einem guten Erkennungsergebnis beitragen können. In [TTA95] wird ein implementiertes System zur Erkennung eines einfachen mechanischen Aggregats vorgestellt. Dazu werden verschiedene *constraints* definiert, die ein optimaler Satz von Kameraparametern erfüllen muß (Sichtbarkeit vs. Teilverdeckung, Auflösung, Entfernung für den Fokuspunkt, Öffnungswinkel) und danach wird dieser berechnet. Ein Zusammenhang mit einer Montageaufgabe, also die Planung unter der Randbedingung, daß eine Aktion zur sensorischen Erfassung einen möglichst großen Fortschritt im Hinblick auf die Steuerung einer Teiloperation bringt, wird nicht hergestellt. Letzteres ist jedoch das Ziel in [MI98]. Basierend auf dem Wissen um die für die Montage relevanten Kontaktflächen, auf dem Wissen um die Extraktionsfähigkeit des Sensors bezüglich der Objektmerkmale und einer Bewertungsfunktion für den zu erwarteten Erfolg einer Sensoroperation wird ein Plan erzeugt. Dieser Ansatz wird für einen Laser-Entfernungsmesser, der um eine Tischfläche rotiert und in seiner Höhe verändert werden kann, implementiert. In [CA99] wird dargestellt, wie allein basierend auf einer Objektansicht eine Folge von *visual tasks* erzeugt werden kann, die ein Roboter, an dessen Hand eine Kamera befestigt ist, dann ausführt.

Ein weiteres Einsatzgebiet ist die Überprüfung von Teilen auf Einhaltung von Toleranzen [TUWR97, YMC98, GMC99]. Solche Teile können im Prinzip ungeordnet und in unvorhersehbarer Position angeliefert werden. Interessant ist hier die Bewertungsfunktion für die beste Sicht; sie wird üblicherweise probabilistisch oder über eine Fuzzy-Regelbasis [KAF99] modelliert. In [LH91] wurde vorgeschlagen, sie so anzusetzen, daß diejenige Sicht des Objekts (bzw. derjenige Sensor) ausgewählt wird, die/der die größte Reduktion von Vieldeutigkeiten bei der Interpretation des interessierenden Bildausschnitts erzielt.

In [Erd95] wird argumentiert, daß es gar nicht erforderlich ist, daß ein Sensor den gesamten Zustand der Umgebung oder eines Objekts erfaßt, sondern daß es reicht, wenn er den Aktionsplaner in die Lage versetzt, über die jeweils nächste mögliche Aktion zu informieren. Die Sensoren werden als ideal angenommen, ihre Meßunsicherheit wird auf resultierende Ungenauigkeit der durch den Sensor gesteuerten aktorischen Operation abgebildet und deren Abfolge durch *backchaining* geplant.

**Kritik.** Es bleibt festzuhalten, daß es auf der Ebene der reinen Ortsplanung und auch für die Aktivierung von Sensoren in Abhängigkeit von einer Montageaufgabe Ansätze gibt. Die Beeinflussung der einzelnen Verarbeitungsschritte über die Wahl der optischen Parameter hinaus, d.h. Vorverarbeitung bis Bilddeutung, unter einer Bewertungsfunktion für die entstehenden Kosten (Bewegungszeit, Verarbeitungszeit, Sensorbelastung) ist in unserem Fall mit Blick auf die Datenfusion und die Verkettung mit der Handlungsplanung erforderlich, aber noch nicht untersucht worden. Abgesehen von ersten Ansätzen zum „Nachschauen“ zur Vermeidung von Verdeckungen [MH98] ist auch die Problematik der zielgerichteten Exploration (etwa das Herumgehen um ein Objekt, das kontrollierte Bewegen eines Aggregats vor einer oder mehreren Kameras) noch weitgehend unbearbeitet.

## 2.3 Datenfusion

Aufgabe der Sensorsteuerung ist es, für einen bestimmten interessierenden Weltausschnitt die zur Verfügung stehenden Informationen zu sammeln und so aufzubereiten, daß sie in ihrer Gesamtheit für die Lösung einer Erkennungs- oder Steuerungsaufgabe genutzt werden können. Das Ziel der Datenfusion<sup>14</sup> ist, Informationen aus verschiedenen Quellen (Teilansichten), die bezüglich ihres Gehalts redundant sind, zu einem einheitlichen Bild zu verschmelzen, d.h. Widersprüche aufzulösen und Lücken zu füllen.<sup>15</sup>

Gegenstand in unserem Zusammenhang ist beispielsweise die Zusammenfügung zweier Bilder, die beide einen begrenzten Blickwinkel wiedergeben, zu einem „Panoramablick“ mit dem summierten Blickwinkel. Hier ist für den Ausgleich der Kamera- und Perspektivenverzerrungen und die Anpassung an den beiden zusammenzufügenden Bildrändern zu sorgen. Eine komplexere Aufgabe der Fusion ist die Rekonstruktion von Objekten, die von mehreren Kameras an unterschiedlichen Standorten aufgenommen werden. Dabei sollen die Objektparameter (Position, Orientierung, Form, Farbe) genauer bestimmt werden, als das mit einer Kamera allein möglich wäre.

Das Zusammenfügen von mehreren Bildern unterschiedlicher Weltausschnitte wird schon seit einiger Zeit untersucht, weniger allerdings im Zusammenhang mit der Robotik, vielmehr für Szenen im freien Gelände oder in Büroräumen (neuere Arbeiten sind hier [SK97, IA97, LWG97]). Die Problematik der Datenfusion ist in der Robotik beginnend mit den frühen Arbeiten [DW87] und [HM91] hauptsächlich unter dem Aspekt der statistischen Modellierung von Sensoren [BM98] und teilweise auch von Objekten [Por88] untersucht worden, inspiriert auch vom Problem der Verfolgung von beweglichen Objekten mit Sensoren an unterschiedlichen Standorten [BSF88]. Es kommen Kombinationen ganz unterschiedlicher Sensoren zum Einsatz: Kameras, Laser-Entfernungsmesser, Ultraschallsensoren, Formerkennung mit strukturiertem Licht, taktile Sensoren, usw. Methodisch wurde dabei auf Entscheidungstheorie [Ber85, Das93, KMM96], lineare und nichtlineare Kalman-Filter [MDW95, VB97, Mut98], die Kombinationsregel nach Dempster-Shafer [Bog87, HK89, Mur98], Bayes-Netze [CLZ96, Pea97] und Modelloptimierung [JS99] gesetzt. Es gibt aber auch Verfahren, die das Problem auf die Minimierung einer für die unterschiedlichen Sensoren zu definierenden Energiefunktion [CY90] oder auf die Repräsentation der Sensoreigenschaften in einer Fuzzy-Regelbasis abbilden [AG92, MBF98, ZSK98, KAF99, ZSK99]. Schließlich wurde im Zusammenhang mit der Steuerung mobiler Roboter bereits sehr früh damit begonnen, künstliche neuronale Netze, deren Eingangsschicht direkt mit einfachen eindimensionalen Sensoren verbunden war, auf die direkte Aktorsteuerung in Abhängigkeit von den unterschiedlichen Meßwerten zu trainieren [Pom93, BP97, HLSvS98]. Interessanterweise wurde jedoch das Problem der Verbesserung der Szenenrekonstruktion bei Vorhandensein mehrerer Kamerasichten aus unterschiedlichen Blickwinkeln und der dadurch tatsächlich erreichbare Fusionsergebnis kaum betrachtet (eine Ausnahme ist [LWG97], einen Ansatz für die Fusion von Entfernungsbildern präsentiert [DWJM98]), obwohl spätestens seit den Arbeiten [FH86, Aya88, AF88] das Problem formuliert und ein statistischer Ansatz in Form eines Kalman-Filters vorhanden war.

**Kritik.** Die algorithmische Seite der Datenfusion ist – auch durch Anpassung von Arbeiten aus anderen Gebieten, speziell der Statistik – bereits recht ausgereift. Bislang sind allerdings zwei Schlüsselprobleme für unseren Einsatzfall nur unbefriedigend gelöst: die *Sensormodellierung*, die Voraussetzung für eine genaue Bestimmung der Objektparameter ist und die die Erkennungsaufgabe bestmöglich lösende automatische Extraktion wichtiger Szeneneigenschaften (*Informationsbewertung*):

- Für die erfolgreiche Anwendung der statistischen Verfahren sind Modelle der Fähigkeiten von Sensoren erforderlich, die über eine einfache quantitative Angabe der vom Sensor zu erwartenden Unsicherheit hinausgehen (siehe [WI95]); es zeigt sich darüber hinaus, daß die Normalverteilungsannahme *in der Praxis* nur in den seltensten Fällen zu befriedigenden Ergebnissen führt. Einen interessanten Fortschritt in dieser Richtung stellt [JS95] dar. Die Autoren untersuchen, wie eine minimale statistische Beschreibung der verschiedenen Sensordaten aussehen kann, deren Unsicherheit insbesondere nicht unbedingt normalverteilt sein muß. Ihr Ansatz basiert auf einem Modell für die (möglichen) Beobachtungen der unterschiedlichen Sensoren, die verglichen werden mit den tatsächlichen Beobachtungen. Das Problem besteht allerdings auch hier darin, daß das Sensormodell inexakt sein kann und deshalb Korrespondenz und Modell gleichzeitig geschätzt werden müssen. Eine Lösung dieses Problems könnte die Verwendung von Techniken aus der Signalverarbeitung zur nichtparametrischen Systemidentifikation, namentlich unter Heranziehung von *blind signal separation* oder *independent component analysis* [Car98] bzw. des Prinzips der maximalen Entropie [Bur75, BS95, FP97] sein. Als weitere Lösung bietet sich an, das Sensormodell zu *lernen* und dieses Wissen in eine Form zu bringen, die dem Fusionsalgorithmus zugänglich ist. Voraussetzung ist eine geeignete Auswahl von Lernbeispielen.
- Offensichtlich kann die Fusion nicht generell auf der Ebene von Bildpunkten (auf der Signalebene) stattfinden, sie muß vielmehr auf der Ebene von Bildmerkmalen oder auch Objekteigenschaften durchgeführt werden [BBW<sup>+</sup>97]. Die automatische Auswahl der adäquaten Ebene, die Auswahl der Merkmale, die Einschätzung des Beitrags der einzelnen Sensoren *a priori*, insgesamt also die Extraktion der für die Erkennungsaufgabe wesentlichen Information über eine Bewertungsfunktion sind bislang noch wenig untersuchte Fragestellungen.

<sup>14</sup>Es herrscht ein gewisses Maß an semantischer Verwirrung um den Begriff der Datenfusion vs. Datenintegration, Datenkombination, Datenaggregation und weitere. Wir verweisen auf [CM99a, CM99b] für eine informelle Klärung der Begriffe und verwenden nur den Begriff Datenfusion.

<sup>15</sup>Bücher, die den Gegenstand im allgemeinen zum Inhalt haben, sind [WL90, Hal92, Sad96, BI97, CGS97, GMN97].

Insgesamt besteht bezüglich der beiden letzten Punkte noch erheblicher Forschungsbedarf, bevor das Potential der Datenfusion für die uns interessierenden Aufgabenstellungen genutzt werden kann.<sup>16</sup>

## 2.4 Erwerb und Generalisierung von Fertigkeiten

Die Programmierung von Robotern durch menschliches „Vormachen“ und „Abschauen“ ist seit langem Ziel der Robotikforschung. Roboter zum Sprühlackieren oder für bestimmte Schweißoperationen, die strukturell einfache Bewegungen machen müssen, bei denen aber Parameter wie Form oder Geschwindigkeitsprofil einzuhalten sind, konnten bereits Ende der siebziger Jahre durch einfache Führung gelehrt werden [Tod86]. Im Prinzip gilt dies auch für Anwendungen der Telerobotik [She92]. Für die Durchführung von Montageoperationen hingegen gibt es bis heute kein praktisch einsetzbares System, welches rein oder zum überwiegenden Teil durch Verfolgung der Bewegungen eines menschlichen Instruktors Handhabungsoperationen lernt (oder auch nur parametrisierte instantiiert).

Eine der ersten Versuche, optische erfaßte menschliche Operationen auf Roboter zu übertragen war [KII92]<sup>17</sup>: Ein Instrukteur machte auf einem Tisch einfache *pick-and-place*-Operationen mit kleinen Klötzchen vor, der Roboter sollte dieselben Manipulationen vornehmen. Ansätze allerdings, die kompliziertere Operationen durch reines Zuschauen erlernen wollten, sind bislang nicht sehr erfolgreich gewesen. Der Grund ist darin zu suchen, daß damit weder die „innere Struktur“ des zu montierenden Aggregats (geeignete Montageflächen, Griffpunkte, Schwerpunkte) noch die nicht-optischen Größen (Kraft, Moment) erkannt werden können. Gleiches gilt für Absichten des Instruktors bzw. sein in möglicherweise nicht sichtbaren Aktionen verstecktes Erfahrungswissen. Weitere Arbeiten zum rein optischen Vormachen sind [PB94]; ein Vorschlag für eine umfassende Architektur findet sich in [SNI<sup>+</sup>94]<sup>18</sup>; eine Möglichkeit, für diesen Lernvorgang Hidden-Markov-Modelle (HMM) zu trainieren, stellt [YXC94, XY95] vor.

Erfolgversprechender sind daher die Versuche, durch Führung des Roboters zu lehren. Dies geschieht z.B. mit Hilfe eines Kraftsensors [Kai97], durch Steuerung über manuelle Gesten/Anweisungen [SWR99, BKM<sup>+</sup>98], durch introspektive Beschreibung der Zieloperation aus Sicht des Instruktors [BU97] oder auch durch Vorlage eines gedruckten Bauplans [ATT99]. [Mye99] stellt einen Ansatz vor, bei dem Führung des Roboters im *master-slave*-Modus, durch seine Bewegung im *zero-gravity*-Modus<sup>19</sup> oder durch Abschauen ein Roboterprogramm in einer konventionellen Programmiersprache erzeugt wird, das schrittweise verfeinert werden kann. In [MTYN99] wird eine Entwicklung vorgestellt, bei der ein Fräs-Zentrum direkt durch die Verfolgung der Fingerbewegungen eines Instruktors gesteuert wird. Damit entfällt der Einsatz von Steuerknüppeln oder speziellen mehrdimensionalen Bewegungsaufnehmern. Komplexe Bewegungszusammenhänge und Aktionen sind hiermit allerdings kaum zu instruieren. Dies ist bei dem Ansatz nach [VK99] eher zu erwarten: hier werden vordefinierte Fertigkeiten durch bestimmte Gesten ausgelöst. Das Hauptproblem wird dabei als das des Findens von Übereinstimmungen identifiziert: Welche vordefinierte (Teil-)Fertigkeit des Roboters erreicht am ehesten das, was der menschliche Instrukteur mit seiner Geste erzielen wollte? Dabei hilft es dem Fluß der Kooperation zwischen Instrukteur und Roboter ganz wesentlich, wenn das Robotersystem über das reine Abschauen hinaus aus der Erfahrung, die es mit einem speziellen Instrukteur gemacht hat, bestimmte Aktionen extrapolieren kann. Eine derartige Vorausschau läßt sich zweckmäßig über HMMs realisieren [PL97, YXC97, YUDS99].

**Kritik.** Es bleibt festzuhalten, daß die Arbeiten zum Lernen durch Vormachen noch in der Anfangsphase stecken und speziell die Übertragung von solcherart gelernten Operationen auf andere Situationen (*skill abstraction*) oder gar die Nutzbarmachung bestimmter Teile von Operationen für andere Fertigkeiten (*skill transformation*) bislang nicht angegangen wurde.

## 3 Entwicklungslinien

Wie bereits in Abschnitt 1 dargestellt, ist es unser zentrales Ziel, den SKK mit einem möglichst hohen Maß an Autonomie auszustatten. Er soll nicht nur lokal und schrittweise elementare Aktionen abarbeiten können, sondern über eine Wegstrecke, deren Länge in der Interaktion mit dem Instrukteur bestimmt wird („Bau mir mal ein . . . “), selbständig komplexe Konstruktionshandlungen ausführen können. Er soll zu einem kooperierenden „mit- und vorausdenkenden“ Assistenten werden, wobei die jederzeitige Interaktions- und Interventionsmöglichkeit des Instruktors auf allen Ebenen erhalten bleibt. Anders ausgedrückt: Erreicht werden soll ein hohes Maß an Autonomie der Ablaufplanung und -ausführung, der Autonomie der sensorischen Erfassung, der Autonomie des (nebenläufigen) Wissenserwerbs samt Wissensanwendung und der Autonomie bei der Dialogführung. Wir stellen nun ein systematisches Arbeitsprogramm vor, dessen Befolgung die schrittweise Erreichung dieser Ziele mindestens für das Szenario des SFB 360 verspricht. Im einzelnen erscheint als Unterteilung zweckmäßig:

<sup>16</sup>Siehe auch die generellen Bemerkungen zur Frage: „was ist wann wie zu fusionieren“ in [DT99].

<sup>17</sup>Eine ähnliche Arbeit ist [IS94, KI95].

<sup>18</sup>Eine umfangreiche Referenz ist [IV97].

<sup>19</sup>Der Roboter gleicht die Gewichte seiner Armsegmente durch Gegensteuern der Gelenkmotoren gerade in dem Maße aus, wie es erforderlich ist, ihn mit (fast) verschwindender Kraft durch einen Menschen zu führen. Für eine geschmeidige Bewegung ist hierfür ein komplexes internes Modell der Roboterstatik und -dynamik erforderlich, darüber hinaus müssen die vom Menschen ausgeübten Führungskräfte durch einen empfindlichen Kraft-Momentensensor an der Roboterhand gemessen werden können.

- *Autonomie bei Abfolgeplanung und -ausführung*: Bereitstellung und Integration von Planungsverfahren zur Montage unter dem Gesichtspunkt der Echtzeitfähigkeit; Planung und Ausführung von Demontageoperationen,
- *Autonomie bei sensorischer Erfassung*: Verschränkte Planung von Sensor- und Aktoreinsatz unter Unsicherheit; Integration von Datenfusionsverfahren in diesen Prozeß,
- *Autonomie bei Wissenserwerb und -anwendung*: Interaktives Lernen neuer Montagesequenzen und dabei Anpassung an veränderte Ausgangssituationen,
- *Autonomie bei der Dialogführung*: Intervenierbarkeit; Abstraktion und Transformation von Fertigkeiten; interaktives Erlernen von Begriffen und Stellen von Rückfragen.

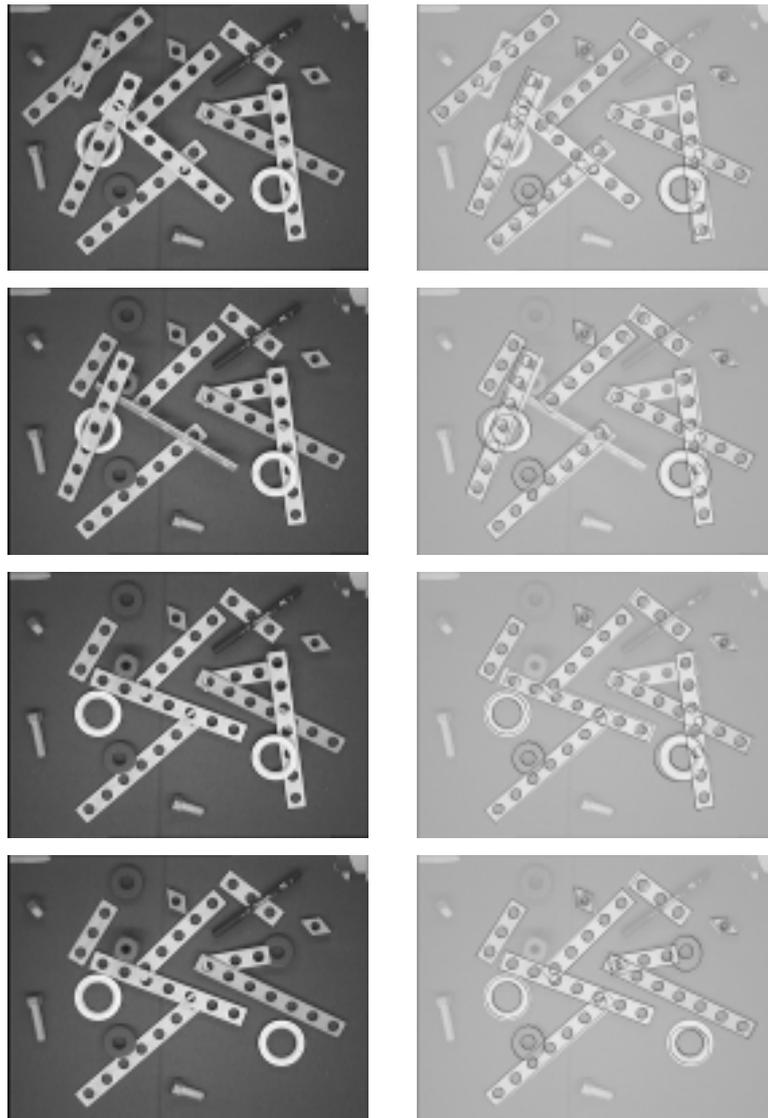


Abbildung 1: Folge von Szenen nach Manipulation durch den Roboter (von oben nach unten). Links Ausgangsbild, rechts visuelle Erkennung mit Kennzeichnung der rekonstruierten Lage. Nicht nur die Resultate der Bildverarbeitung ändern sich (Genauigkeit, Menge der erkannten Objekte), sondern es ändert sich auch die Objektlage: die Dreilochleiste rechts in Bildmitte wird erst nach dem Abheben der Siebenlochleiste (rechts, senkrecht) erkannt (Zeile 3 nach Zeile 4); die Siebenlochleiste in der Bildmitte fällt während des Abhebens der Fünflochleiste vom Würfel auf ihre schmale Seite.

### 3.1 Bereitstellung und Integration von Planungsverfahren

**Fragestellungen/Ziele.** Der erste Schritt zur Erzielung von Autonomie ist die Fähigkeit des *SKK*, automatisch ein Planungsgerüst für die Konstruktion komplexerer Aggregate bis hin zum fertigen Aggregat (z.B. dem Flugzeug) zu erstellen. Die Planung

findet auf der Grundlage *realer Daten* (nämlich einem – immer nur begrenzt korrekt arbeitenden – Objekterkennungs- und -rekonstruktionssystem) statt, die Planungsergebnisse werden nicht nur visualisiert, sondern ihre Korrektheit muß sich darüber hinaus in der realen Durchführung im Demonstrator bewähren. Deshalb müssen an das Planungssystem drei wesentliche als Randbedingung zu formulierende Forderungen gestellt werden, die bei einem weniger unmittelbar von den unvorhersehbaren Ereignissen in der Außenwelt abhängigen Planer eine eher untergeordnete Rolle spielen:

- Planung unter Unsicherheit. Unsicherheit resultiert aus der nicht eindeutigen Bestimmung der Objektlagen und damit der Vorbedingungen (Ist ein Objekt wirklich greifbar?) und der fehlerhaften Ausführung oder Erfolglosigkeit der auszuführenden Aktion.
- Planung unter Echtzeitanforderungen. Es ist den zeitlichen Erwartungen des Instruktors zu genügen.
- Planung mit geometrischen Randbedingungen. Verschränkung mit der Sensoreinsatzplanung und der Bahnplanung sowie der Griffplanung der kooperierenden Arme.

Die Bildfolge nach Abbildung 1 illustriert die Problematik der Unsicherheit: die einzelnen Szenen geben die Lage der Montageobjekte nach einem Griff durch den Roboter wieder; markiert sind auch die jeweils erkannten Objekte. Die Strategie des Roboters ist es, durch Abnahme des jeweils zuoberst liegenden Teils die Objektrelationen nicht zu ändern. Der Griff führt dennoch – verstärkt durch die Unsicherheiten bei der Erkennung – zu einer stets geänderten Ausgangssituation für die Planung der nächsten Schritte. Die dargestellte Situation ist durch die Vielfachüberdeckung besonders anspruchsvoll: auch der menschliche Beobachter hat hier Schwierigkeiten, Ursache, Wirkung und zweckmäßige Folgehandlung zu ermitteln (insbesondere, wenn ihm nur diese Perspektive zur Verfügung steht). Die Problematik tritt jedoch bereits im einfachsten Fall nebeneinanderliegender Objekte auf, wenn deren Lage zwecks besserer Greifbarkeit geändert werden muß. Dann ergeben sich nicht nur andere geometrische Relationen, es ist auch nicht sichergestellt, daß Objekte, die vorher erkannt wurden, weiterhin erkennbar sind.<sup>20</sup>

**Methodik.** Zurückgegriffen werden kann hier auf die ersten Ansätze zur Planung bei Mehrgagentensystemen [BB97]; sie sind zu adaptieren und auf ihre Leistungsfähigkeit hin zu untersuchen. Ihre Leistungsfähigkeit ist zu vergleichen mit Systemen, die lediglich eine *action-selection* in Abhängigkeit vom Systemzustand vornehmen, ohne einen klassischen Plan aufzubauen. Interessant ist im Sinne der Forderung nach Echtzeitfähigkeit auch der Ansatz des *analogical replay* [BV97], bei dem für das Planen unter Unsicherheit Teilpläne für bestimmte Situationen wiederverwendet und damit eine erhebliche Zeitersparnis erreicht wird.<sup>21</sup>

Vor allem soll, wo immer möglich, auf den reichen Methodenschatz, den die KI in den letzten Jahrzehnten aufgebaut hat, zurückgegriffen werden. Das gilt nicht nur für die Planungsverfahren, sondern natürlich insbesondere für die frei verfügbaren Planungssysteme. Interessant erscheinen uns die Systeme SATPLAN [KS96] und ASP [BLG97]. Diese Systeme müssen so erweitert werden, daß sie den Input des Sensorsubsystems akzeptieren und damit eine Planmodifikation bzw. Neuplanung und eine schritt haltende Planausführungsüberwachung möglich wird.

### 3.2 Planung und Ausführung von Demontageoperationen

**Fragestellungen/Ziele.** Es ist nicht damit zu rechnen, daß Instruktoren von vornherein über das notwendige Vorstellungsvermögen verfügen, um die Montagefolge komplexer oder auch einfacherer Aggregate mental zu planen und „aus dem Kopf“ in eine korrekte Anweisungsfolge umsetzen zu können. Wenn nicht vorher mittels Simulation und ggf. ein alle Beschränkungen einbeziehendes Probehandeln zwischen SKK und Instrukteur eine Übereinkunft über die vorzunehmenden Montageschritte erzielt wurde, dann wird es nicht ausbleiben, daß sich die beiden in Sackgassen bewegen. Mit anderen Worten: Der Instrukteur erkennt erst während oder nach der Ausführung der Aktion, daß sie nicht auf das vom ihm verfolgte Ziel hinführt. In einem solchen Fall muß nicht nur Intervenierbarkeit (siehe unten) sichergestellt sein, sondern es müssen auch die Primitiva für die Demontage zur Verfügung stehen (Schraube lösen, Schraube herausdrehen, Abziehen, Abnehmen, . . . ). Es muß ferner dafür gesorgt sein, daß während der Montageplanung Wiederaufsetzpunkte definiert werden, die bei einer Demontageplanung Zielpunkte sein können.

**Methodik.** Die eigentliche Demontageplanung soll mit denselben Techniken wie die Montageplanung, ergänzt um die erforderlichen Sensormeldungen, ausgeführt werden. Dabei ist der angestrebte Zielzustand durch einen Wiederaufsetzpunkt definiert, der Anfangszustand ist direkt der bis zum Interventionszeitpunkt erreichte Montagezustand. Wegen der weitgehenden Ähnlichkeit zwischen beiden ist zu erwarten, daß ihre Parameter über die Techniken zur Abstraktion von Fertigkeiten (Abschnitt 3.7) automatisch bestimmt werden können.

<sup>20</sup>Es sei bemerkt, daß an solchen Stellen Erkenntnisse über mentale Modelle und menschliches Planungsverhalten sinnvoll zu integrieren sind, siehe dazu beispielsweise [NRK99].

<sup>21</sup>Auch hier ergeben sich möglicherweise direkte Anknüpfungspunkte zu Erkenntnissen über das menschliche Problemlösen mit Analogieschlüssen [WHKH97].

### 3.3 Verschränkte Planung von Sensor- und Aktoreinsatz unter Unsicherheit

**Fragestellungen/Ziele.** Die beweglichen bzw. mitbewegten Sensoren des *SKK* sind so einzusetzen, daß sie die unter einer gegebenen Erfassungsaufgabe notwendige Information aus einer sich ständig verändernden Umwelt extrahieren. Erwünscht ist dabei die nahtlose Integration in die Planung der Aktorbewegungen, insbesondere natürlich derjenigen Aktoren, die einen mitbewegten Sensor (optisch, taktil) tragen. Die Folge nach Abbildung 1 illustriert das Problem für verschiedene Folgezustände einer Ausgangsszene, die durch die Ausführung bestimmter Aktionen durch den Roboter entstehen: Objekte ändern während des Ablaufs der einzelnen Manipulationen ihre Lage, das Erkennungssystem detektiert jeweils unterschiedliche Objekte (schlimmstenfalls werden solche, die vorher erkannt wurden, im weiteren Verlauf nicht mehr erkannt), und je nach Objektlage liefert der Erkenner unterschiedliche Genauigkeiten bei der Lagebestimmung.

Bei räumlich verteilter Erfassung durch Kameras ergibt sich darüber hinaus das Problem, daß korrespondierende Punkte eines Objekts gefunden werden müssen, obwohl sich die Ansichten von Sicht zu Sicht fundamental unterscheiden (Leiste mit Normalansicht vs. Leiste aus Seitenansicht) – verstärkt durch die Schwierigkeiten bei der Teilüberdeckung. Ziel muß es hier also sein, Sichten zu erzeugen, die es erlauben, von möglichst vielen Objekten korrespondierende Teilansichten zu erzeugen, die dann von der Datenfusionskomponente zu einem Gesamtbild zusammengefügt werden können.

Der Einsatz der Sensoren erfolgt gemischt parallel (für gleichzeitig arbeitende Sensoren) und sequentiell (für mitbewegte Sensoren). Für ihre zeit- und kostenoptimale Verwendung ist es also erforderlich, daß das Planungssystem folgenden Randbedingungen Rechnung trägt:

- Kenntnis des prinzipiell möglichen und jeweils aktuellen Erfassungsvermögens (z.B. Sichtwinkel, Standorte) der einzelnen Sensoren ,
- Kenntnis der Erfassungsaufgabe bzw. der aus der Handhabungsaufgabe folgenden Erkennungsaufgaben (z.B. genaue Bestimmung von Greifpunkten),
- Kenntnis der aktuellen Absichten des Einsatzes der Aktoren,
- Kenntnis der Anforderungen der Fusionskomponente,
- Bewertung des von einem bestimmten Sensor gelieferten bzw. des von ihm in einer bestimmten Situation zu erwartenden Ergebnisses.

Perspektivisch geht es hier darum, Erfassungsstrategien zum „Nachschauen“ zu entwickeln, die die eigene „Körperlichkeit“ des *SKK* mit in Rechnung stellen. Es ist eine Umgebung zu schaffen, die es dem *SKK* erlaubt, eine Selbstrepräsentation über seine sensorischen Fähigkeiten und deren Verwendung zur zielgerichteten Durchführung der Montageoperationen zu entwickeln (Aufbau eigener Repräsentationsschemata).

**Methodik.** Zunächst ist hier zu untersuchen, ob die Berechnung von Sensorstandorten mit Hilfe einfacher Heuristiken zur Bestimmung optimaler Greifpunkte [MI98] auch in unserem Einsatzfall zu brauchbaren Erfassungsergebnissen führt.

Danach ist ein Modell für die Sensoren und ihre Abbildungsfunktionen in Abhängigkeit vom Standort in der Sprache des Planungsmoduls für die Aktorik so zu entwickeln, daß ihre Einplanung mit maximalem Gewinn für die Lösung der Manipulationsaufgabe möglich wird. Die Bewertungsfunktion soll soweit wie möglich automatisch überwacht gelernt werden, und zwar durch Beobachtung einer exakt bezüglich der Lage der Bauteile bekannten Musterszene und den Vergleich des Meßergebnisses. Der Einbezug einer Bewertungsfunktion für das zu erwartende Erfassungsergebnis in den verschachtelten Planungsprozeß läßt es möglich erscheinen, daß eine Montagesequenz einen ganz anderen Verlauf nimmt als dies ohne den Einbezug sensorischer Unsicherheit der Fall wäre. Schließlich ist zu untersuchen, in welchen Situationen es sinnvoll ist, Aktorhandlungen so zu modifizieren, daß damit einerseits die mitbewegten Sensoren ohne zusätzliche Kosten Information beitragen können und andererseits muß erforscht werden, ob das Bewegen eines Objekts nur zum Zwecke der verbesserten Erfassungsmöglichkeit (z.B. durch eine ungezielte Vereinzelnungsoperation) sinnvoll ist.

Das eigentliche Planungsverfahren sollte soweit wie möglich in den für die Aktorplanung verwendeten Formalismus integriert werden; Sensormessungen entsprechen dann genauso auf das Ziel hinführenden Handlungen und verursachen ebenso Kosten, wie die zustandsändernden Aktionen der Roboter. Besonders vielversprechend erscheint uns hier die Formulierung der Planung als Problem der stochastischen dynamischen Programmierung auf der Basis von Markowschen Entscheidungsprozessen, ähnlich wie dies in [HM99] für ein als *discrete-event-system* modelliertes Sensorsubsystem eines mobilen Roboters implementiert wurde. Auf der Ebene der programmtechnischen Umsetzung ist es denkbar, die Planungsaufgabe zu partitionieren und auf ein *team of experts* abzubilden, die im Wettstreit miteinander unter Berücksichtigung der zu erwartenden Kosten und Ergebnisse über die jeweils nächste Sensoraktion entscheiden, ihren Erfolg eigenständig bewerten und dieses Wissen für die nächste Entscheidung heranziehen.

### 3.4 Datenfusion

**Fragestellungen/Ziele.** Nachdem über eine leistungsfähige Sensoreinsatzplanung die zur Lösung einer Aufgabe erforderlichen Informationsbestandteile extrahiert wurden, müssen diese Teile nun zu einer plausiblen Gesamtsicht integriert werden. Wie unter Abschnitt 2.3 bereits ausgeführt, sind die statistischen Verfahren zur Datenfusion bereits gut entwickelt, allerdings sind sie in unserem Zusammenhang nur bedingt einsetzbar. Normalverteiltes Sensorrauschen ist damit zwar gut behandelbar, es führt jedoch bei optischen Sensoren gegenüber den nichtlinearen Verzerrungen (Linsenfehler, Perspektive, Farbwerte) nur zu einem vernachlässigbaren Fehler. Um mit diesen Sensoren in vertretbarer Zeit zu einem möglichst exakten Abbild der wahren Objektparameter zu kommen, sind Techniken zu entwickeln zur

- Modellierung des Verhaltens der Sensoren (Abbildungsverhalten in Abhängigkeit von der Lage, von Umwelteinflüssen, von aufgewendeter Zeit) über die rein stochastische Charakterisierung hinaus unter Einbezug z.B. von Erfahrungswissen;
- Aufstellung einer Bewertungsfunktion zur Auswahl der Merkmale, deren Extraktion verspricht, für die Erkennungsaufgabe wesentliche Information zu liefern (Steuerung des Fusionsprozesses auf unterer Ebene);
- Steuerung des Fusionsprozesses auf höherer Ebene (Fusionsfolgeoptimierung), also die Auswahl von nacheinander angebotenen Sensorergebnissen zu einem Gesamtbild.

Wenn eine automatisch trainierbare Bewertungsfunktion für die spezifischen Leistungen der Einzelsensoren zur Verfügung steht, können wirklich redundante Sensoren praktisch implementiert werden. Ein solcher Sensor wird an das Sensornetzwerk angeschlossen und dann durch aktives Training unter Mitbenutzung seiner eigenen Beweglichkeit und bei Hilfestellung der Aktorik zum Vormachen vollständig autonom eingebunden. Mit anderen Worten: wenn ein komplexer Sensor hinzugefügt wird, wird er automatisch „eingelernt“, die zeitaufwendige explizite Programmierung und Kalibration entfällt vollständig.

**Methodik.** Zur *Modellierung des Verhaltens* der Sensoren eignet sich nach unserer Einschätzung insbesondere der in [ZK98b] vorgestellte Neuro-Fuzzy-Ansatz auf der Basis von B-Splines mit seiner exzellenten Fähigkeit zur Approximation nichtlinearer Funktionen [ZK99a, ZK98a, ZKS99, ZBK99]. Damit können nicht nur die Übertragungsfunktionen (von Kraftsensoren, der einzelnen Farbkanäle von Kameras, usw.) repräsentiert werden; es ist auch möglich, diese in Abhängigkeit von bestimmten Situationen (Beleuchtung, Perspektive, . . . ) zu modifizieren. Ferner sind Vorkenntnisse in Form von Funktionen, Tabellen, Wahrscheinlichkeitsverteilungen leicht in diese Repräsentation zu überführen. Um mehr oder weniger scharfes Erfahrungswissen über die Sensoren ausnutzen zu können, können Ansätze für Fuzzy-Quantoren dazu eingesetzt werden, natürlichsprachlich formulierte Bewertungen der Sensorleistungen in das Sensormodell einbauen zu können.

Für das Training und die Anwendung von *Bewertungsfunktionen* werden zweckmäßigerweise Fuzzy-Regelbasen verwendet. Sie sollen anhand von Beispielen mit Hilfe überwachten Lernens trainiert und dabei so effizient implementiert werden, daß sie als Auswahlfunktion für das Vertragsnetzprotokoll zur Laufzeit auswertbar sind. Damit ergibt sich gleichzeitig eine Teambildung bzw. eine *Fusionsfolgeoptimierung*<sup>22</sup> durch die dezentrale Ausschreibung von Folgeaufträgen, die aus den über die Bewertungsfunktion abzulesenden Beiträgen der einzelnen Sensoren für einen Auftrag resultieren.

### 3.5 Interaktives Lernen neuer Montagesequenzen

**Fragestellungen/Ziele.** Grundlage für die eigenständige Planung und nachfolgende Ausführung von längeren Ablauffolgen zur Konstruktion von Aggregaten ist die prototypische Durchführung einer Sequenz und die Benennung des Ergebnisses durch den Instrukteur. Während dieser Durchführung werden die Aktionsfolgen und die korrespondierenden Resultate der Umwelt- und Objekterkennung in einer geeigneten Repräsentationsform aufgezeichnet. Zunächst ist es sicher zweckmäßig, nur Lernvorgänge mit dem Ziel der selbständigen Wiederausführung auf der Ebene von Teilaggregaten, etwa des Flugzeugs, anzustreben (Leitwerk, Rumpf, Fahrwerk, Propeller, . . . ). Sobald diese erfolgreich gelernt werden können, sollen die Teilaggregate – zunächst ausgehend von einer standardisierten Konfiguration für die einzelnen Bauelemente auf dem Montagetisch – robust und völlig autonom gebaut werden. Um die Flexibilität des Ansatzes zu validieren, könnte danach beispielsweise der Aufbau einiger Aggregate des Baufix-Rollers gelernt werden.<sup>23</sup>

Ein sinnvolles Ziel ist ferner die Verknüpfung mit dem Modul zur Montagefolgeplanung (siehe Abschnitt 3.1) in der Weise, daß später auch *modifizierende Anweisungen* gegeben werden können. In einfacheren Fällen soll dies keinen weiteren Lernvorgang erfordern, bei aufwendigeren Modifikationen sollte zumindest erkannt werden können, welche der bereits gelernten alternativen Anweisungsfolge durchzuführen ist. Beispiele:

> „Bau mir ein Leitwerk, aber mit *grünem* Würfel“ (gelernt wurde für einen blauen),

<sup>22</sup>Die Fusionsfolgeoptimierung steht offensichtlich in enger Beziehung zur Sensoreinsatzplanung und setzt im Falle mitbewegter Sensoren darüber hinaus die Kenntnis der Aktorplanung voraus. Entsprechende Kostenfunktionen als Grundlage einer trilateralen Abstimmung sind zu entwickeln.

<sup>23</sup>Der Roller ist mit derselben Menge von Baufix-Elementen wie das Flugzeug aufzubauen, siehe dazu den Bauplan des Baufix-Kastens Nr. 1120.

- > „Bau mir die Tragflächen, aber nimm *statt der Siebenlochleiste die fünflöchrige*“ (gelernt wurde nur für die Siebenlochleiste),
- > „Bau jetzt ein Fahrwerk, aber für das *große Flugzeug*“ (gelernt wurde für das große und das kleine Flugzeug).

Eingeführt werden sollten schließlich spezielle Anweisungen (als eine Klasse von *Metainstruktionen*), die ausschließlich für den Lerndialog Anwendung finden, sogenannte *bedingte Instruktionen*. Damit wird es möglich, Aktionsfolgen von vorhergehenden Aktionsfolgen, vom Systemzustand und vom sensorischen Input abhängig zu machen. Beispiel:

- > „*Wenn Du jetzt eine Schraube und eine Dreilochleiste siehst, kannst Du sie auf den grünen Würfel schrauben. Sonst nimmst Du eine Fünflochleiste und schraubst sie an die Seite und sagst mir Bescheid. Danach fragst Du nach der Dreilochleiste*“.

Es sollte untersucht werden, welche dieser Anweisungen zweckmäßig eingeführt werden, um den Lernvorgang zu beschleunigen und welche die Abstraktionsfähigkeit des *SKK* erhöhen<sup>24</sup> (siehe Abschnitt 3.7).

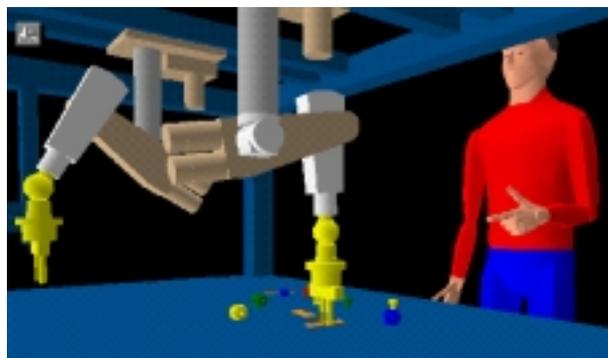


Abbildung 2: Kombinierte Interaktion über Sprache für die gewünschte Operation und Zeigen für die räumliche Relation.

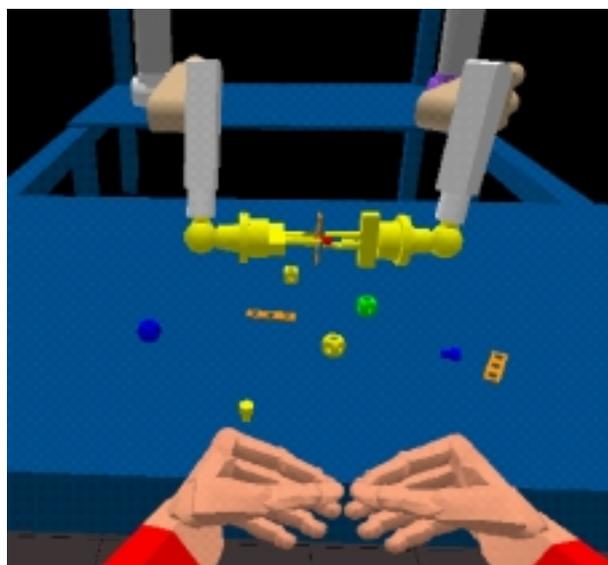


Abbildung 3: Zeigen einer Bewegungsrichtung mit den Händen des Instrukteurs zusätzlich zur Sprache.

**Methodik.** Es sind hier Ansätze für zwei Bereiche zu entwickeln und zu integrieren:

1. Lernen von diskreten Ereignisabläufen in Abhängigkeit von Sensormustern<sup>25</sup> bzw. Umweltkonstellationen unter Einbezug von automatisch zu erzeugenden Unterzielen und der Integration der reaktiven Planung,

<sup>24</sup>Dazu gehören beispielsweise die Verwendung von Vorankündigungen: „Wir wollen jetzt ein *Flugzeug* bauen ...“.

<sup>25</sup>Hierbei sind Sensormuster antizipatorisch einzubeziehen; das dazu notwendige Wissen soll aus dem Vergleich von Sensormuster und bekannter, zugrundeliegender Objektkonstellation abgebildet werden (Ableitung der Abbildungsfunktion des Sensors; zur Sensormodellierung siehe Abschnitt 3.4).

2. Ableitung und Erfragen von *Erklärungen* über den Prozeßverlauf, um kausales Montagewissen zu erwerben und dieses Wissen bei analogen Aufgaben erneut einzusetzen,
3. Verhaltensbeobachtung eines bestimmten Instruktors während der Führung des Lerndialogs.

Bei der *Verhaltensbeobachtung* des Instruktors sind seine sprachlichen und gestischen Anweisungen zu analysieren, die dazu dienen, den Lernvorgang zu steuern. Zum einen beobachtet der SKK den Instruktore und erkennt dabei dynamische Ereignisse, die wiederholt werden können. Auf der Ebene der Anweisungen zur Steuerung des Lernvorgangs sind neben der rein sprachlichen weitere Erfassungsmodalitäten zu kombinieren. Sie sollen integriert werden, sobald sie robust einsetzbar sind. Wir geben sie dennoch hier an, um die Richtung anzudeuten:

- Zeigeposition. Über die relativ einfache Auswertung von Zeigegesten mit gestreckten Zeigefinger oder mit einem speziellen Stift können in Verbindung mit der Spracheingabe nicht nur Objekte disambiguiert, sondern der Roboter kann auch zu einem bestimmten Objekt dirigiert werden (siehe Abbildung 2).
- Instruktionsqualifikation. Über komplexere Gesten können Anweisungen zur Modifikation bzw. zu exakterer Spezifikation sprachlicher Instruktionen gegeben werden: Reichen, Schrauben, Führen, Zusammen-Tragen; siehe Abbildung 3.
- Sprecherlokalisierung. Eine wesentliche Erleichterung kann es für den Prozeß der Intentionserkennung sein, wenn der Ort des Sprechers bekannt ist. Dies gilt speziell für die Einbeziehung räumlicher Relationen („Links von *mir*“; „Links von *Dir*“).
- Blickrichtung. Die ungefähre Abschätzung des Blickwinkels des Instruktors kann ganz offensichtlich ebenfalls die Erkennung seiner Intention wesentlich erleichtern.

Für alle diese Modalitäten sind Erfassungstechniken zu entwickeln bzw. anzupassen. Es ist außerdem zu studieren, welche Zeigegesten, Erläuterungsgesten und Blickwinkel in einem Montagedialog eine Rolle spielen, weil der Instruktore damit seine Absichten unterlegt.

Es soll in diesem Zusammenhang nicht unerwähnt bleiben, daß es sehr reizvoll wäre, wenn der Instruktore mit dem SKK in Verbindung mit einer geeigneten Visualisierung über den Bauplan des Flugzeugs zunächst rein simulativ eine Operationenfolge aushandeln könnte; damit ließe sich die Ausführung des Dialogs während der realen Montage sicher deutlich abkürzen. Das gleiche gilt, wenn während der Ausführung des Dialogs eine Prädiktion des Verhaltens des Instruktors erreicht werden könnte. Für solche Zwecke liegen bereits Vorarbeiten in Form von HMMs vor [YXC97, PL97]. Möglicherweise eignen sich auch die ersten Ansätze zur Planerkennung bei Kooperationen mehrerer Agenten [DR98] dazu, die Intentionserkennung zu verbessern.



Abbildung 4: Unterschiedliche Ausgangssituation am Beginn der Montage (aus denen unterschiedliche Montagefolgen resultieren). Die Abbildungen zeigen jeweils dieselbe Objektmenge.

Für das eigentliche *Lernen* der Folgen eignet sich unserer Erfahrung nach das Verstärkungslernen (ggf. in Kombination mit anderen Techniken zum überwachten Lernen) besonders gut, hier ergeben sich zusätzlich die Möglichkeiten der direkten Bewertung einer Aktion durch den Instruktore durch eine weitere Klasse von Metainstruktionen (Instruktore: „So<sup>26</sup> hast Du es *richtig* gemacht!“). Außerdem gibt es einen engen Querbezug zum Planen auf der Basis von Markow-Prozessen ([ZTS99], siehe Abschnitt 2.1.1). Als Repräsentation für die gelernten Aktionsfolgen sind Skripte, wie sie die Grundlage von OPERA bilden [ZvCK99], sinnvoll. Diese können dann einerseits in OPERA nachbearbeitet werden oder aber auch so transformiert werden, daß zu Kontrollzwecken durch den Instruktore das *Playback* zunächst rein virtuell durch ein Simulationssystem erfolgt. Alternative Repräsentationsformen sind UND/ODER-Graphen/Petri-Netze, Listen von konditionierten Aufgaben-Operatoren oder Vorrang-Graphen.

<sup>26</sup>So bedingt das Vorliegen einer entsprechenden Typologie der Ereignisse.

Ein weiterer wesentlicher Punkt ist die Abstraktion von der konkreten Lage der Baufix-Teile für die Durchführbarkeit der Montagefolge. Es sollen automatisch Operationen des Manipulators eingeplant werden können, die die Objekte in eine greifbare Position bringen, um auch bei ganz verschiedenen Ausgangssituationen zu einer erfolgreichen Montagefolge zu kommen (siehe Abbildung 4). Der Zusammenhang mit einer leistungsfähigen Objekterkennung ist klar, denn es muß wenigstens entschieden werden können, ob die erforderliche Objektmenge vorliegt. Im Prinzip kann mit einem solchen Abstraktionsvermögen sogar erreicht werden, daß der *SKK* von sich aus die Initiative ergreift, wenn man ihm eine bestimmte Menge von Bauteilen vorlegt: „Das sieht ja so ähnlich aus wie vorgestern, sollen wir also genauso vorgehen . . . ?“ oder „Damit können wir ja ein . . . bauen!“

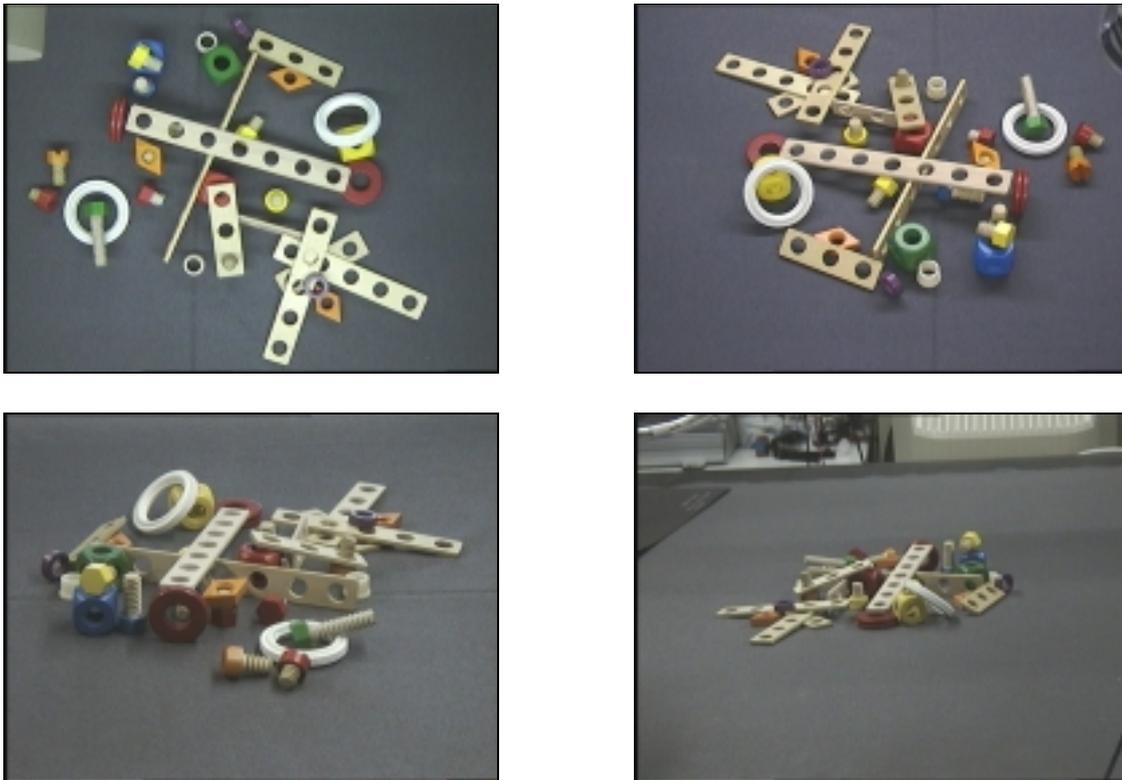


Abbildung 5: Unterschiedliche Ansichten derselben Szene, wie sie sich aus verschiedenen Blickpunkten mit unterschiedlichen Erfassungswinkeln ergeben. Hier hat auch ein menschlicher Betrachter Schwierigkeiten, korrespondierende Objekte sofort zu identifizieren.

### 3.6 Intervenierbarkeit und Fehlerbehandlung

**Fragestellungen/Ziele.** Speziell bei der Ausführung längerer autonomer Handlungssequenzen ist die jederzeitige Unterbrechbarkeit und Modifizierbarkeit des Ablaufs unter Erreichung eines definierten sicheren Wartezustands unabdingbar für ein praktisch einsetzbares System. Asynchrone, d.h. nicht einem festen Raster zuzuordnende Unterbrechungen/Modifikationen können dabei sowohl vom Nutzer durch imperative Instruktionen („Halt“, „Zurück“, „Stärker“, „Schneller“)<sup>27</sup> gegeben, als auch durch nicht vorhersehbare, nicht zu kontrollierende detektierbare exogene Ereignisse in der Umwelt (Verrutschen, Verlieren, Verkanten, Kollidieren, . . . ).

Nach Beseitigung des Fehlers bzw. nach Wegfall der Ursache für die Intervention muß ein Wiederaufsetzen ebenfalls durch Zuruf erfolgen können („Weiter“), ohne daß die eigentlich intendierte Instruktion wiederholt wird, d.h. die Planungskomponente muß über ein entsprechendes Zustandskonzept verfügen. Schließlich soll der *SKK* von sich aus eine Fehlerbeseitigung vornehmen oder aber Wünsche gegenüber dem Instrukteur äußern können („Soll ich *es* nochmal versuchen?“, „Kannst Du den roten Würfel wieder etwas näher heranrücken? Ich kann ihn nicht erreichen!“)

**Methodik.** In der üblichen Weise wird unterschieden zwischen synchronen Ereignissen, die nur an bestimmten Punkten der Ausführung auftreten können (Schraube schneidet nicht ein, Würfel haftet nicht, . . . ) und asynchronen Ereignissen, die jederzeit und unvorhergesehen auftreten können (Hardwarefehler, Interventionen des Instrukteurs, die oben erwähnten exogenen Ein-

<sup>27</sup>Wobei der Bezug jeweils durch den *SKK* zu inferieren ist; für Rückfragen bleibt keine Zeit.

flüsse). Ob ein synchrones Ereignis ausgelöst werden muß, wird nur an fest definierten Punkten während der Montage überprüft. Um die asynchronen Events zu erkennen, werden mehrere aktive Sensoren (visuelle, taktile, Kraftsensoren) eingesetzt. Nach Feststellung des Auftretens eines Ereignisses werden beide Arten gleich behandelt, indem das Planungsmodul um eine geeignete, vorher definierte Ausnahmebehandlung ersucht wird. Interessant erscheint es uns, über den Mechanismus zum interaktiven Lernen neuer Montagefolgen in den vorgesehene bedingten Anweisungen auch Fehlerbeseitigungsstrategien gleich mitzulernen. Dabei soll es zwei Klassen geben: antizipatorische („Wenn Dir jetzt der Würfel herunterfallen sollte, dann . . .“) und rückgreifende („Wenn das<sup>28</sup> nochmal passiert, dann versuche . . .“). Auch hierzu kann ein zweckdienlicher Simulationsdialog mit dem Ziel der Durchführung von *Probearbeiten* konzipiert werden.

### 3.7 Abstraktion und Transformation von Fertigkeiten

**Fragestellungen/Ziele.** Hier geht es um die Extraktion der für die generischen Klassen von Handhabungsvorgängen wesentlichen Kenngrößen; langfristig kann auch die Abbildung des Vorgangs auf bestimmte physikalische Grundgesetze (Hooksches Gesetz, Reibungskegel, . . .) ins Auge gefaßt werden.

Die zentrale wissenschaftliche Fragestellung ist die Aufstellung minimaler expliziter Repräsentationen für die wichtigen Parameter der Handhabungsoperationen durch den *SKK* selbst. Es steht dies in unmittelbarer Analogie zum Begriffslernen (Abschnitt 3.8): so wie dort angestrebt wird, das Konzept „Schraube“ unabhängig davon, ob es sich um eine Baufix-Holzschraube oder eine kleine Metallschraube handelt, mit einer Semantik zu füllen (die ohne weitere Erklärungen auch auf neue Schraubentypen paßt), ist es hier das Ziel, Handlungen bei neuen Objekten unter Kenntnis der wesentlichen Objektcharakteristika *ausführen* zu können, ohne daß vorher ein erneuter Trainingsvorgang erforderlich wird, mit ihnen also eine allgemeine Handlungssemantik und -pragmatik zu assoziieren (und nach Möglichkeit auch, ohne dabei vorher eine große Menge unterschiedlicher Fälle bearbeitet zu haben [LKW97] bzw. eine ständige Rückkopplung mit dem Menschen zu benötigen [NYM99]).<sup>29</sup>

Mit dieser Abstraktion soll zweierlei erreicht werden: zum einen soll die Anpassungsfähigkeit an strukturell neue Konstellationen<sup>30</sup> erhöht und der zeitliche Aufwand bis zu ihrer sicheren Handhabung verkürzt werden, zum anderen soll aber – längerfristig – ein grundsätzliches *Verständnis* des *SKK* über sein eigenes Vermögen, Montagehandlungen vornehmen zu können, erreicht werden. Mit diesem *Eigenmodell* und darauf zu definierenden Emulationshandlungen (Selbstrepräsentation) kann der *SKK* dann auch bei neu auftretenden Situationen sofort über die wesentlichen Kenngrößen einer intendierten Operation in Abhängigkeit von den Parametern des Objekts bestimmen und über ihre Ausführbarkeit entscheiden („So etwas kann ich nicht schrauben, dazu wird meine Kraft nicht reichen“).

**Methodik.** Um das langfristig angestrebte Ziel der selbständigen Aufstellung einer vollständigen Handlungssemantik zu erreichen, können Ansätze der „naiven Physik“ bzw. „qualitativen Mechanik“ [SDS97, Dah88] zur Grundlage der Repräsentation gemacht werden. In einem ersten Schritt können gelernte Fertigkeiten so verallgemeinert werden, daß sie weitgehend unabhängig von konkreten Objekten für eine Klasse von Operationen anwendbar sind. Das Vorgehen soll sich dabei an die in [ZFK99, ZKR99, ZK99c] vorgestellte Methodik anlehnen: Basis für die Generalisierung und Transformation ist die Synthese eines Fuzzy-Reglers, der nach Training eine konkrete Operation ausreichend robust ausführt. Dessen Regeln können mit dem noch zu entwickelnden Verfahren zusammengefaßt bzw. verallgemeinert werden.<sup>31</sup> Konkret sollen neben dem Übertragen des Schraubvorgangs auf andere Arten von Schrauben/Gewindekombinationen auch Steckoperationen verallgemeinert werden.

### 3.8 Interaktives Erlernen von Begriffen, Stellen von Rückfragen

**Fragestellungen/Ziele.** Es ist bei der Übertragung der antrainierten Fähigkeiten des *SKK* auf eine umfassendere Domäne als es die Baufix-Welt ist, aber auch schon bei der Einführung neuer Objekte und ggf. Handlungen in diese Domäne, wünschenswert, die zugehörigen Begriffe und die damit verbundene Semantik nicht explizit modellieren zu müssen. Sie soll vielmehr bezüglich aller für die Manipulationsdurchführung relevanten Bedeutungen durch Lernen im Dialog mit dem Instrukteur vermittelbar sein. Dieser Lernvorgang steht vor dem Lernen von Montagefolgen, die ein Einverständnis zwischen *SKK* und Instrukteur über Begriffe und Semantik voraussetzen.

Es wird davon ausgegangen, daß der Begriff im Lexikon der sprachverarbeitenden Komponente vorhanden, daß aber seine Semantik weitgehend leer ist (dynamisches Lexikon). Um diese Semantik aufzufüllen, sind die folgenden wesentlichen Eigenschaften und Fakten durch Zeigen/Vormachen/Mitteilen/Bewegen durch den *SKK* abzuleiten, und dazu sind auch die richtigen

<sup>28</sup>Das bedingt das Vorliegen einer entsprechenden Typologie der Ereignisse.

<sup>29</sup>Zu einer kompletten Handlungssemantik gehört natürlich noch die Angabe von Vorbedingungen und Resultaten in einer geeigneten Darstellungsform.

<sup>30</sup>Schrauben in langes Gewinde des Klotzes vs. Schrauben in kurzes Gewinde der Rautenmutter oder Schrauben einer Leiste auf einen Würfel (ein Gewinde) vs. Schrauben eines Würfels auf einen Würfel (zwei hintereinanderliegende Gewinde).

<sup>31</sup>Es gibt häufig mehrere Möglichkeiten, benachbarte Regeln zusammenzufassen und nicht jede mögliche Gruppierung ergibt letztlich eine minimale Regelbasis. Es stellt sich also die Frage, welche Regeln jeweils gruppiert werden müssen, um die Regelbasis so weit wie möglich zu reduzieren. Es zeigt sich, daß Regeln, die zusammengefaßt werden können, mehrere wichtige Eigenschaften besitzen: Ihre Konsequenzen sind gleich, sie sind in der Regeltabelle direkt benachbart und sie bilden in der Regeltabelle rechteckige Bereiche. Dieser Sachverhalt soll bei der Generalisierung ausgenutzt werden.

Rückfragen zu stellen, die in diesem Fall die Form von Anweisungen an den Instrukteur übernehmen (*SKK*: „Zeig mir bitte jetzt die Punkte, in die ich die Schraube stecken kann!“):

- Aussehen des Objekts aus verschiedenen Sichten (ein Beispiel zeigt Abbildung 6), um eine Grundlage für seine Wiedererkennung durch optische Sensoren zu haben,
- Griffpunkte am Objekt (die handlungsabhängig sein können),
- Handlungen, die *auf* dem Objekt oder *mit ihm* ausgeführt werden können sowie der geometrische Ort für diese Handlungsausführung (Montageports, ...).

Alle diese Eigenschaften sind im Lernvorgang festzulegen. Unmittelbar damit verknüpft ist neben dem Stellen von Rückfragen das Problem der Simulation (Probearbeiten), denn damit ergibt sich unter Umständen erst das Verständnis für den aktuellen Fortschritt beim Verstehensprozeß des *SKK* bzw. die Schwierigkeiten, die der *SKK* aufgrund seiner Eigenwahrnehmung antizipiert, die aber für den Instrukteur nicht unmittelbar erkennbar sind<sup>32</sup> (*SKK*: „Soll ich es *so* machen, wie ich es Dir *jetzt* am Bildschirm zeige?“, „Diesen Punkt kann ich nicht erreichen, weil ...“).

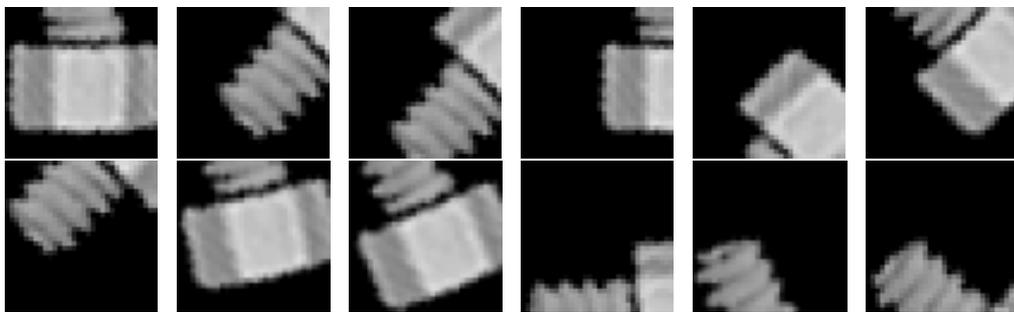


Abbildung 6: Das Wissen über eine Schraube beinhaltet neben den impliziten Funktionseigenschaften auch das *Aussehen* der (teilüberdeckten, nur teilweise sichtbaren) Schraube. Dazu sind eine ganze Reihe von Bildern als Trainingsmenge erforderlich.

**Methodik.** Es soll hier versucht werden, alle Komponenten des *SKK*, die dazu beitragen können, ein Objekt mit Begriff und Semantik zu versehen, so zu erweitern, daß sie den Lernprozeß steuern können und außerdem Repräsentationen erzeugen, die mit der Darstellung im Lexikon kompatibel sind (und umgekehrt wieder dazu verwendet werden können, sensorische und aktorische Prozesse bei der Ausführung der Montageoperationen zu steuern). Beispiel: der Bilderkenner muß über geeignete Techniken<sup>33</sup> entscheiden, welche Ansichten eines Objekts er für eine eindeutige Identifikation in den möglichen Lagen benötigt und weist dazu den Roboter an, das auf dem Tisch liegende zunächst unbekannte Teil in verschiedene Lagen zu bugsieren/stoßen/etc. Angestrebt wird, die Möglichkeit zum Begriffs- und Bedeutungslehren in dieser Form für zwei bis drei Objekte vollständig zu demonstrieren, d.h. den Auffüllungsprozeß des Semantik-Teils des Lexikon-Eintrags nur durch Lehranweisungen in unterschiedlichen Modalitäten vollständig durchführen zu können.

Schließlich ist abzuklären, bis zu welchem Abstraktionsgrad Rückfragen auf Anweisungen gestellt werden können, wenn auf den dann erreichten Stand aller anderen Komponenten zurückgegriffen werden kann (Instrukteur: „Bau mir mal ein Flugzeug“ – *SKK*: „Welchen Flugzeugtyp?“ ...). Zur Simulation des vom *SKK* auch mit Hilfe der Antwort auf die Rückfragen erstellten Plans kann ein Simulationssystem verwendet werden, das schritthaltend den jeweils erstellten Plan erhält und diesen dann visualisiert.

## Literatur

- [AAT99] S. Abrams, P. Allen, and K. Tarabanis. Computing camera viewpoints in an active robot work cell. *Int. J. of Robotics Research*, 18(3):267 – 285, 1999.
- [AF88] N. Ayache and O. Faugeras. Building, registering and fusing noisy visual maps. *Int. J. of Robotics Research*, 7(6), 1988.
- [AG92] M. Abidi and R. Gonzalez, editors. *Data Fusion in Robotics and Machine Intelligence*. Academic Press, 1992.
- [Ast65] K. Astrom. Optimal control of Markov decision processes with incomplete state estimation. *J. Math. Analysis and Appl.*, 10, 1965.
- [ATT99] N. Abe, K. Tanaka, and H. Taki. Understanding of mechanical assembly instruction manual by integrating vision and language processing and simulation. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.

<sup>32</sup>Beschränkungen durch fehlende Sicht, durch nicht vorhandene Sensorik, drohende Kollision, ...

<sup>33</sup>Hier bieten sich natürlich die im Umfeld der Aspektgraphen [KvD79] gewachsenen Methoden an.

- [Aya88] N. Ayache. *Construction et fusion de représentations visuelles tridimensionnelles – applications à la robotique mobile*. PhD thesis, INRIA, 1988.
- [BA99] S. Botelho and R. Alami. M+: A scheme for multi-robot cooperation through negotiated task allocation and achievement. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Bak88] A. Baker. Complete manufacturing control using a contract net: A simulation study. In *Proc. Int. Conf. on Comp. Int. Manufacturing*, pages 100 – 109, 1988.
- [BB97] C. Boutilier and R. Brafman. Planning with concurrent interacting actions. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [BBW<sup>+</sup>97] T. Boulton, R. Blum, R. Wallace, G. Zhang, S. Nayar, P. Allen, and J. Kender. Advanced visual sensor systems. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [BCG<sup>+</sup>99] A. Bredendfeld, T. Christaller, W. Göhring, H. Günther, H. Jaeger, H. Kobialka, P. Plöger, and P. Schöll. Behavior engineering with “dual dynamics” models and design tools. In *Proc. 19th Joint Conf. Artif. Intell. IJCAI-99 (to appear)*. Morgan Kaufmann, 1999.
- [BDH99] C. Boutilier, T. Dean, and S. Hanks. Decision-theoretic planning: Structural assumptions and computational leverage. *J. Art. Intell. Res.*, 11:1 – 93, 1999.
- [Ber85] J. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistics. Springer-Verlag, 1985.
- [BFH99] C. Borst, M. Fischer, and G. Hirzinger. A fast and robust grasp planner for arbitrary 3D objects. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [BI97] R. Brooks and S. Iyengar. *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice-Hall, 1997.
- [BKM<sup>+</sup>98] M. Becker, E. Kefalea, E. Makl, C. v. d. Malsburg, M. Pagel, J. Triesch, J. C. Vorbrüggen, and S. Zadel. Gripsee: A robot for visually-guided grasping. In *Proceedings of the Int. Conf. Artif. Neural Netw. 1998*, Sweden, 1998.
- [BLG97] B. Bonet, G. Loerincs, and H. Geffner. A robust and fast action selection mechanism for planning. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [BM98] I. Bloch and H. Maitre. Fusion of image information under imprecision. In B. Bouchon-Meunier, editor, *Aggregation and Fusion of imperfect information*, Studies in Fuzziness and Soft Computing. Physica-Verlag, 1998.
- [Bog87] P. Bogler. Shafer-Dempster reasoning with applications to multisensor target identification systems. *IEEE Trans. Systems, Man, and Cybernetics*, 17, 1987.
- [BP97] S. Baluja and D. Pomerleau. Dynamic relevance: Vision-based focus of attention using artificial neural networks. *Artificial Intelligence*, pages 381 – 395, 1997.
- [BS95] A. Bell and T. Sejnowski. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7, 1995.
- [BSF88] Y. Bar-Shalom and T. Fortmann. *Tracking and Data Association*. Academic Press, 1988.
- [BU97] I. Bratko and T. Urbančič. Transfer of control skill by machine learning. *Eng. Appl. of Artif. Intell.*, 10(1):63 – 71, 1997.
- [Bur75] J. Burg. *Maximum Entropy Spectral Analysis*. PhD thesis, Stanford University, 1975.
- [Bur96] E. Burdet. *Algorithms of Human Motor Control and their Implementation in Robotics*. PhD thesis, ETH-Zuerich, 1996.
- [BV97] J. Blythe and J. Veloso. Analogical replay for efficient conditional planning. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [CA99] C. Colombo and B. Allotta. Image-based robot task planning and control using a compact visual representation. *IEEE Trans. on Systems, Man, and Cybernetics (A)*, 29(1), 1999.
- [Car98] J. Cardoso. Blind signal separation: Statistical principles. *Proc. IEEE*, 86(10):2009 – 2025, 1998.
- [CGS97] V. Cantoni, V. Di Gesu, and A. Setti, editors. *Human and Machine Perception: Information Fusion*. Plenum Press, 1997.
- [CLZ96] K. Chang, J. Liu, and J. Zhou. Bayesian probabilistic inference for target recognition. In *Signal Processing, Sensor Fusion, and Target Recognition V*, volume 2755 of *SPIE Proceedings*. SPIE – Society of Photo-Optical Instrumentation Engineers, 1996.
- [CM99a] T. Celinski and B. McCarrager. Achieving efficient data fusion through integration of sensory perception control and sensor fusion. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [CM99b] T. Celinski and B. McCarrager. Improving sensory perception through predictive correction of monitoring errors. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Cor93] P. Cork. Visual control of robot manipulators – a review. In K. Hashimoto, editor, *Visual Servoing*, pages 31 – 31. World Scientific Publishing, 1993.
- [CY90] J. Clark and A. Yuille. *Data Fusion for Sensory Information Processing Systems*. Kluwer Academic, 1990.
- [Dah88] K. Dahlgreen. *Naïve semantics for natural language understanding*. Kluwer, Dordrecht, 1988.
- [Das93] B. Dasarthy. *Decision Fusion*. IEEE Press, 1993.
- [DHW93] D. Draper, S. Hanks, and D. Weld. Probabilistic planning with information gathering and contingent execution. Technical Report 93-12-04, University of Washington Department of Computer Science and Engineering, 1993.

- [DL95] T. Dean and S. Lin. Decomposition techniques for planning in stochastic domains. In *Proc. 14th Joint Conf. Artif. Intell. IJCAI-95*, pages 1121 – 1127. Morgan Kaufmann, 1995.
- [DR98] M. Devaney and A. Ram. Needles in a haystack: Plan recognition in large spatial domains involving multiple agents. In *Proc. 15th Nat. Conf. on Artif. Intell. AAAI-98*. AAAI Press / The MIT Press, 1998.
- [DT99] B. Dasarathy and S. Townsend. FUSE – fusion utility sequence estimator. In *Proc. 2nd Int. Conf. on Information Fusion*, Sunnyvale, CA, 1999. Int. Society of Information Fusion.
- [DW87] H. Durrant-Whyte. Sensor models and multi-sensor integration. *Int. J. Robotics Research*, 6(3), 1987.
- [DWJM98] C. Dorai, G. Wang, A. Jain, and C. Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(1), 1998.
- [EHW<sup>+</sup>92] O. Etzioni, S. Hanks, D. Weld, D. Draper, N. Lesh, and M. Williamson. An approach to planning with incomplete information. In *Proc. 3rd Int. Conf. on Principles of Knowledge Representation and Reasoning KR-1992*, pages 115 – 125. Morgan Kaufmann, 1992.
- [Erd95] D. Erdmann. Understanding action and sensing by designing action-based sensors. *Int. Journal of Robotics Research*, 14(5), 1995.
- [FGL87] K. Fu, R. Gonzalez, and C. Lee. *Robotics – Control, Sensing, Vision and Intelligence*. McGraw-Hill, 1987.
- [FH86] O. Faugeras and M. Hebert. The representation, recognition, and locating of 3-d objects. *Int. J. of Robotics Research*, 5(3), 1986.
- [FN71] R. Fikes and N. Nilsson. STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2(3/4), 1971.
- [Fox81] M. Fox. An organizational view of distributed systems. *IEEE Trans. on Systems, Man and Cybernetics*, SMC-11, 1981.
- [FP97] J. Fisher and J. Principe. A nonparametric methodology for information theoretic feature extraction. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [FPWK99] J. Fraile, C. Paredis, C. Wang, and P. Khosla. Agent-based planning and control of a multi-manipulator assembly system. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [GI94] K. Gremban and K. Ikeuchi. Planning multiple observations for object recognition. *Int. J. Comp. Vision*, 12(2):137 – 172, 1994.
- [GMC99] X. Gu, M. Marefat, and F. Ciarallo. A robust approach for sensor placement in automated vision dimensional inspection. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [GMKB97] R. Goldman, D. Musliner, K. Krebsbach, and M. Boddy. Dynamic abstraction planning. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [GMN97] I. Goodman, R. Mahler, and H. Nguyen. Mathematics of data fusion. In *Theory and Decision Library. Series B, Mathematical and Statistical Methods*, volume 37. Kluwer Academic, 1997.
- [Hal92] D. Hall. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, 1992.
- [HD98] R. Horaud and F. Dornaika. Visually guided object grasping. *IEEE Trans. on Robotics and Automation*, 14(4):525 – 532, 1998.
- [HK89] S. Hutchinson and A. Kak. Planning sensing strategies in a robot work cell with multi-sensor capabilities. *IEEE Trans. on Robotics and Automation*, 5(6), 1989.
- [HLSvS98] U. Handmann, G. Lorenz, T. Schnitger, and W. v. Seelen. Fusion of different sensors and algorithms for segmentation. In *Proceedings of the IEEE Conference on Intelligent Vehicles*, 1998.
- [HM91] G. Hager and M. Mintz. Task-directed sensor data fusion and sensor planning. *Int. J. of Robotics Research*, 10(4), 1991.
- [HM99] G. Howland and B. McCarragher. Control of sensory perception in a mobile navigation problem. *Int. J. of Robotics Research*, 18(2):201 – 212, 1999.
- [Ho92] Y. Ho, editor. *Discrete Event Dynamic Systems*. IEEE Press, 1992.
- [Hol90] J. Hollerbach. Planning of arm movements. In D. Osherson, S. Kosslyn, and J. Hollerbach, editors, *Visual Cognition and Action*, volume 2. MIT Press, 1990.
- [HSFS99] A. Hauck, M. Sorg, G. Färber, and T. Schenk. What can be learned from human reach-to-grasp movements for the design of robotic hand-eye systems. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [IA97] M. Irani and P. Anandan. Robust multi-sensor image alignment. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [IS94] K. Ikeuchi and T. Suehiro. Toward an assembly plan from observation. Part I: Task recognition with polyhedral objects. *IEEE Trans. on Robotics and Automation*, 10(3), 1994.
- [IV97] K. Ikeuchi and M. Veloso, editors. *Symbolic Visual Learning*. Oxford University Press, 1997.
- [JS95] R. Joshi and A. Sanderson. Multisensor fusion and unknown statistics. In *Proc. IEEE Conference on Robotics and Automation*. IEEE Comp. Soc. Press, 1995.
- [JS99] R. Joshi and A. Sanderson. Minimal representation multisensor fusion using differential evolution. *IEEE Trans. on Systems, Man, and Cybernetics (A)*, 29(1), 1999.

- [JWC98] R. Jones, R. Wilson, and T. Calton. On constraints in assembly planning. *IEEE Trans. on Robotics and Automation*, 14(46):849 – 863, 1998.
- [KAF99] F. Kobayashi, F. Arai, and T. Fukuda. Sensor selection by reliability based on possibility measure. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Kai97] M. Kaiser. Transfer of elementary skills via human-robot interaction. *Adaptive Behavior*, 5(3/4):249–280, 1997.
- [KI95] S. Kang and K. Ikeuchi. Toward automatic robot instruction from perception – temporal segmentation of tasks from human hand motion. *IEEE Trans. on Robotics and Automation*, 11(5), 1995.
- [KII92] Y. Kuniyoshi, M. Inaba, and H. Inoue. Seeing, understanding and doing human task. In *Proc. 1992 IEEE Int. Conf. on Robotics and Automation*, Nice, 1992. IEEE-Press.
- [KMM96] G. Kamberova, R. Mandelbaum, and M. Mintz. Statistical decision theory for mobile robotics: theory and applications. In *Proc. 1996 IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems (MFI 96)*, Washinton, D.C., Dec. 1996.
- [KQW97] O. Khatib, S. Quinlan, and D. Williams. Robot planning and control. *Robotics and Autonomous Systems*, 21:249 – 261, 1997.
- [KS96] H. Kautz and B. Selman. Pushing the envelope: Planning, propositional logic and stochastic search. In *Proc. 13th Nat. Conf. on Artif. Intell. AAAI-96*, pages 1194 – 1201. AAAI Press / The MIT Press, 1996.
- [KvD79] J. Koenderink and A. van Doorn. The internal representation of solid shape with respect to vision. *Biological Cybernetics*, 32(4):211 – 216, 1979.
- [Lat91] J. Latombe. *Robot Motion Planning*. Kluwer Academic, 1991.
- [LB91] E. Lebsanft and F. Braun. Ein Expertensystem zur Auftragsreihenfolgeplanung für eine roboterisierte Packstraße in der Lebensmittelindustrie. In H. Bullinger, editor, *Expertensysteme in Produktion und Engineering*. Springer-Verlag, 1991.
- [LC97] D. Leng and M. Chen. Robot trajectory planning using simulation. *Robotics & Computer-Integrated Manufacturing*, 13(2):121 – 129, 1997.
- [LH91] S. Lee and H. Hahn. An optimal sensing strategy for recognition and localization of 3-D natural quadric objects. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(10), 1991.
- [LKW97] D. Leake, A. Kinley, and D. Wilson. Case-based similarity assessment: Estimating adaptability from experience. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [LL98] F. Lin and H. Levesque. What robots can do: robot programs and effective achieveability. *Artificial Intelligence*, 101:201 – 206, 1998.
- [Llo98] J. Lloyd. Removing the singularities of serial manipulators by transforming the workspace. In *Proc. 1998 IEEE Int. Conf. on Robotics and Automation*, Leuven, Belgium, 1998. IEEE-Press.
- [LM99] S. Lee and H. Moradi. Disassembly sequencing and assembly sequence verification using force flow networks. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [LR99] H. Levesque and R. Reiter. High-level robotic control: beyond planning. A position paper. In *AAAI 1998 Spring Symposium: Integrating Robotics Research: Taking the Next Big Leap*, Stanford University, 1999. AAAI-Press.
- [LWG97] M. Leventon, W. Wells, and W. Grimson. Multiple 2D-3D mutual information registration. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [Mae90] P. Maes. Situated agents can have goals. *Robotics and Autonomous Systems*, 6(1):49 – 70, 1990.
- [Man99] G. Mantriota. Optimal grip points for contact stability. *Int. J. of Robotics Research*, 18(5):502 – 513, 1999.
- [MBF98] G. Mauris, E. Benoit, and L. Foulloy. Fuzzy linguistic methods for aggregation of complementary sensor information. In B. Bouchon-Meunier, editor, *Aggregation and Fusion of Imperfect Information*, Studies in Fuzziness and Soft Computing. Physica-Verlag, 1998.
- [MDW95] J. Manyika and H. Durrant-Whyte. *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Ellis Horwood, 1995.
- [MH98] E. Marchand and G. Hager. Dynamic sensor planning in visual servoing. In *Proc. 1998 IEEE Int. Conf. on Robotics and Automation*, Leuven, Belgium, 1998. IEEE-Press.
- [MI98] J. Miura and K. Ikeuchi. Task-oriented generation of visual sensing strategies in assembly tasks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(2), 1998.
- [Mir97] B. Mirtich. *V-Clip: Fast and Robust Polyhedral Collision Detection*. TR 97-05, Mitsubishi Electric, Cambridge, MA, June 1997. <http://www.merl.com/reports/TR97-05/index.html>.
- [MLS94] R. Murray, Z. Li, and S. Sastry. *Robotic Manipulation*. CRC Press, 1994.
- [MR91] D. McAllester and D. Rosenblitt. Systematic nonlinear planning. In *Proc. 9th Nat. Conf. on Artif. Intell. AAAI-91*, pages 634 – 639. AAAI Press, 1991.
- [MS97] J. Miura and Y. Shirai. Vision-motion planning of a mobile robot considering vision uncertainty and planning cost. In *Proc. 15th Joint Conf. Artif. Intell. IJCAI-97*, pages 1194 – 1200. Morgan Kaufmann, 1997.

- [MTYN99] M. Mitsuishi, K. Tanaka, Y. Yokokohji, and T. Nagao. Remote rapid manufacturing with “action media” as an advanced user interface. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Mur98] R. Murphy. Dempster-shafer theory for sensor fusion in autonomous mobile robots. *IEEE Trans. on Robotics and Automation*, 14(2), 1998.
- [Mut98] A. Mutambara. *Decentralized Estimation and Control for Multisensor Systems*. CRC Press, 1998.
- [Mye99] D. Myers. An approach to automated programming of industrial robots. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Nil94] N. Nilsson. Teleo-reactive programs for agent control. *J. Art. Intell. Res.*, 1:139 – 158, 1994.
- [NRK99] J. Nerb, F. Ritter, and J. Krems. Knowledge level learning and the power law: A soar model of skill acquisition in scheduling. *Kognitionswissenschaft*, 8(1):20 – 29, 1999.
- [NYM99] Y. Nakamura, T. Yamazaki, and N. Mizushima. Synthesis, learning and abstraction of skills through parameterized smooth map from sensors and behaviors. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [OHB<sup>+</sup>99] D. Ouelhadj, C. Hanachi, B. Bouzouia, A. Moualek, and A. Farhi. A multi-contract net protocol for dynamic scheduling in flexible manufacturing systems (FMS). In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Par89] H. Parunak. Manufacturing experience with the contract net. In M. Huhns, editor, *Distributed Artificial Intelligence*. Pitman, 1989.
- [PB94] P. Pook and D. Ballard. Deictic teleassistance. In *Proc. IROS '94 – IEEE/RSJ/GI Int. Conf. on Intell. Robots and Systems*. IEEE Press, 1994.
- [PC96] L. Pryor and G. Collins. Planning for contingencies: A decision-based approach. *J. Art. Intell. Res.*, 4:287 – 339, 1996.
- [Pea97] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Academic Press / Morgan Kaufmann, 1997.
- [PL97] A. Pentland and A. Liu. Modeling and prediction of human behavior. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [PO98] M. Parnichkun and S. Ozono. GSGM movement model for cooperative robots system. *Mechatronics*, 8:905 – 925, 1998.
- [Pom93] D. Pomerlau. *Neural network perception for mobile robot guidance*. Kluwer Academic, 1993.
- [Por88] J. Porrill. Optimal combination and constraints for geometrical sensor data. *Int. J. of Robotics Research*, 7(6), 1988.
- [Rev99] S. Reveliotis. Production planning and control in flexibly automated manufacturing systems: Current status and future requirements. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [RW89] P. Ramadge and W. Wonham. The control of discrete event systems. *Proc. IEEE*, 77(1):81 – 98, 1989.
- [Sac74] E. Sacerdoti. Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5(2):115 – 135, 1974.
- [Sad96] F. Sadjadi. Selected papers on sensor and data fusion. In *SPIE Milestone Series*, volume MS 124. SPIE - Society of Photo-Optical Instrumentation Engineers, 1996.
- [Sch98] T. Scherer. *Design and Implementation of a Trajectory Generator for Arbitrarily Moving Targets and On-Line Singularity Robustness*. Diplomarbeit, Universität Bielefeld, Technische Fakultät, Arbeitsgruppe Technische Informatik, 1998.
- [SD81] R. Smith and R. Davis. Frameworks for cooperation in a distributed problem solver. *IEEE Trans. on Systems, Man and Cybernetics*, SMC-11, 1981.
- [SDS97] T. Stahovich, R. Davis, and H. Shrobe. Qualitative rigid body mechanics. In *Proc. 14th Nat. Conf. on Artif. Intell. AAAI-97*. AAAI Press / The MIT Press, 1997.
- [She92] T. Sheridan. *Telerobotics, Automation, and Human Supervisory Control*. MIT-Press, Boston, 1992.
- [SK97] H. Sawhney and R. Kumar. Multi-image alignment. In T. Strat, editor, *Proc. of the 1997 Image Understanding Workshop*, New Orleans, 1997. Morgan Kaufmann.
- [SLG<sup>+</sup>99] G. Smith, E. Lee, K. Goldberg, K. Böhringer, and J. Craig. Computing parallel-jaw grips. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Smi92] M. Smith. An environment for more easily programming a robot. In *Proc. 1992 IEEE Int. Conf. on Robotics and Automation*, Nice, 1992. IEEE-Press.
- [SNI<sup>+</sup>94] T. Sato, Y. Nishida, Y. Ichikawa, Y. Hatamura, and H. Mizoguchi. Active understanding of human intention by a robot through monitoring of human behavior. In *Proc. IROS '94 – IEEE/RSJ/GI Int. Conf. on Intell. Robots and Systems*. IEEE Press, 1994.
- [ŠO98] P. Švestka and M. Overmars. Coordinated path planning for multiple robots. *Robotics and Autonomous Systems*, 23:125 – 152, 1998.
- [Ste81] M. Stefik. Planning with constraints. *Artificial Intelligence*, 16(2):111 – 169, 1981.
- [SWR99] D. Schwammkrug, J. Walter, and H. Ritter. Rapid learning of robot grasping position. In *Proc. Int. Symposium on Intelligent Robotics Systems (SIRS-99)*, 1999.

- [TAT95] K. Tarabanis, P. Allen, and R. Tsai. A survey of sensor planning in computer vision. *IEEE Trans. on Robotics and Automation*, 11(1), 1995.
- [Tod86] D. Todd. *Fundamentals of Robot Technology: An Introduction to Industrial Robots, Teleoperators and Robot Vehicles*. Wiley, 1986.
- [TTA95] K. Tarabanis, R. Tsai, and P. Allen. The MVP sensor planning system for robotic vision tasks. *IEEE Trans. on Robotics and Automation*, 11(1), 1995.
- [TUWR97] E. Trucco, M. Umasuthan, A. Wallace, and V. Roberto. Model-based planning of optimal sensor placements for inspection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(2), 1997.
- [VB97] P. Varshney and C. Burrus, editors. *Distributed Detection and Data Fusion*. Springer-Verlag, 1997.
- [VK99] R. Voyles and P. Khosla. Gesture-based programming: A preliminary demonstration. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Wel94] D. Weld. An introduction to least commitment planning. *AI Magazine*, 15(4):27 – 61, 1994.
- [WHKH97] R. Westermann, F. Hesse, G. Kauer, and A. Hiemisch. Zur Repräsentation systematischer Ähnlichkeiten zwischen analogen Problemen nach zielgerichteten Elaborationen. *Kognitionswissenschaft*, 6(3):101–114, 1997.
- [WI95] M. Wheeler and K. Ikeuchi. Sensor modeling, probabilistic hypothesis generation, and robust localization for object recognition. *IEEE Trans on Patt. Analysis and Mach. Intell.*, 17, No. 3, March 1995.
- [Wil88] D. Wilkins. *Practical Planning: Extending the Classical AI Planning Problem*. Morgan Kaufmann, 1988.
- [WL90] E. Waltz and J. Llinas. *Multisensor Data Fusion*. Artech House, 1990.
- [Wu99] H. Wu. Methodology of generating recovery procedures in a robotic cell. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, May 1999. IEEE-Press.
- [XY95] Y. Xu and J. Yang. Towards human-robot coordination: Skill-modelling and transferring via hidden- markov-model. In *Proc. IEEE Conference on Robotics and Automation*. IEEE Comp. Soc. Press, 1995.
- [YG99] X. Yuan and Y. Gu. An integration of robot programming and sequence planning. In *Proc. of the 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [YMC98] C. Yang, M. Marefat, and F. Ciarallo. Error analysis and planning accuracy for dimensional measurement in active vision inspection. *IEEE Trans. on Robotics and Automation*, 14(3), 1998.
- [YUDS99] Y. Yamada, Y. Umetani, H. Daitoh, and T. Sakai. Construction of a human/robot coexistence system based on a model of human will – intention and desire. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [YX96] W. Yun and Y. Xi. Optimum motion planning in joint space for robots using genetic algorithms. *Robotics and Autonomous Systems*, 18:373 – 393, 1996.
- [YXC94] J. Yang, Y. Xu, and C. Chen. Hidden markov model approach to skill learning and its application to telerobotics. *IEEE Trans. on Robotics and Automation*, 10, No. 5, Oct. 1994.
- [YXC97] J. Yang, Y. Xu, and C. Chen. Human action learning via hidden markov model. *IEEE Trans. on Systems, Man, and Cybernetics (A)*, 27(1):34 – 44, 1997.
- [ZBK99] J. Zhang, W. Baqai, and A. Knoll. A Comparative Study of RBFN and B-Spline Fuzzy Controllers. In *Proc. Fourth European Workshop on Fuzzy Decision Analysis and Recognition Technology for Management, Planning and Optimisation*, pages 173–181, Dortmund, 1999.
- [ZFK99] J. Zhang, M. Ferch, and A. Knoll. Extraction and transfer of fuzzy control rules for sensor-based robotic operations. In *International Fuzzy Systems Association World Congress*, Taipei, 1999.
- [Zha95] J. Zhang. *Ein Integriertes Verfahren zur effizienten Planung und Ausführung von Roboterbewegungen in unscharfen Umgebungen*. Infix-Verlag, 1995.
- [Zha99a] X. Zha. Discrete event control with active events. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [Zha99b] X. Zha. An intelligent approach and system for rapid robotic assembly prototyping, planning and control. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [ZK98a] J. Zhang and A. Knoll. Constructing fuzzy controllers for multivariate problems using statistical indices. In *Proceedings of the IEEE International Conference on Fuzzy Systems, Alaska*, 1998.
- [ZK98b] J. Zhang and A. Knoll. Constructing fuzzy controllers with B-spline models – principles and applications. *International Journal of Intelligent Systems*, 13(2/3):257–285, 1998.
- [ZK99a] J. Zhang and A. Knoll. Designing fuzzy controllers by rapid learning. *Fuzzy Sets and Systems*, 101/2, January 1999.
- [ZK99b] J. Zhang and A. Knoll. Integrating deliberative and reactive strategies via fuzzy modular control. In A. Saffiotti and D. Driankov, editors, *Fuzzy logic techniques for autonomous vehicle navigation*. Springer, 1999.
- [ZK99c] J. Zhang and A. Knoll. Towards transparent control of large and complex systems. In *Proceedings of The Third International Conference on Computing Anticipatory Systems*, Liège, Belgium, 1999.

- [ZKR99] J. Zhang, A. Knoll, and I. Renners. Efficient learning of non-uniform B-splines for modelling and control. In *International Conference on Computational Intelligence for Modelling, Control and Automation*, pages 282–287, Viena, 1999.
- [ZKS99] J. Zhang, A. Knoll, and V. Schwert. Situated neuro-fuzzy control for vision-based robot localisation. *Robotics and Autonomous Systems*, 1999. (to appear).
- [ZSK98] J. Zhang, V. Schwert, and A. Knoll. Visuelles Lernen zur Lokalisierung eines mobilen Roboters. In *Tagungsband Autonome Mobile Robotersysteme*, Karlsruhe, 1998.
- [ZSK99] J. Zhang, R. Schmidt, and A. Knoll. Appearance-based visual learning in a neuro-fuzzy model for fine-positioning of manipulators. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [ZTS99] G. Zhao, S. Tatsumi, and R. Sun. A heuristic Q-learning architecture for fully exploring a world and deriving an optimal policy by model-based planning. In *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, Detroit, Michigan, 1999. IEEE-Press.
- [ZvCK99] J. Zhang, Y. von Collani, and A. Knoll. Interactive assembly by a two-arm robot agent. In *Robotics and Autonomous Systems*, 1999.