

Seminar „Autonome Roboter“ im SS '99

# Aktive Objekterkennung und -vermessung zur Steuerung eines Demontageroboters

---

**Axel Guicking**

E-Mail: <guicking@gmx.de>

11. Juli 1999

---

---

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung und Motivation</b>	<b>1</b>
<b>2</b>	<b>Systemarchitektur</b>	<b>2</b>
<b>3</b>	<b>Objektmodellierung und -erkennung</b>	<b>3</b>
3.1	Hybride Objektmodellierung . . . . .	4
3.2	Ganzheitliche Objekterkennung . . . . .	6
3.3	Rekonstruktion der räumlichen Tiefe . . . . .	10
<b>4</b>	<b>Ergebnisse</b>	<b>11</b>

# 1 Einleitung und Motivation

Betrachtet man die heute im Einsatz befindlichen Montage- oder Demontageroboter, so führen die meisten von ihnen Arbeiten mit einer stets gleichbleibenden Aktionsfolge aus; die Anpassung an veränderte Umweltbedingungen muß dabei jeweils von einem Programmierer vorgenommen werden.

Um eine autonome Adaption des Roboters an eine neue Situation zu ermöglichen, muß er in der Lage sein, Daten seiner Umwelt mittels Sensoren wahrzunehmen und auszuwerten.

Bei der Demontage kommen einige Faktoren hinzu, die den automatischen Prozeßablauf erschweren (bei der Montage treten diese Faktoren nicht auf) [1]:

- Die zu demontierenden Objekte sind häufig durch Schmutz oder Beschädigungen verändert,
- die Neigung und relative Position der Objekte zum Roboter sind unterschiedlich, und
- der Hintergrund der ganzen Szene ist selten einheitlich.

Um eine zuverlässige Objekterkennung zu gewährleisten, ist also eine Erkennungsstrategie nötig, die in bezug auf Verschmutzungen und Positionierung des Objekts *tolerant*, sowie *vom Hintergrund unabhängig* ist. Um die Lage und Position eines Objekts zu bestimmen, ist ein gutes 3D-Modell der gesamten Demontageszene unentbehrlich.

Der hier vorgestellte Ansatz zur Objekterkennung basiert auf einer Kombination von neuronalen Netzen zur eigentlichen Bilderkennung und einer Wissensrepräsentation des zu demontierenden Objektes mit einem semantischen Netz. Durch diese *hybride Objektmodellierung* ergibt sich eine Performanz, die deutlich über der der jeweils einzelnen Modulen liegt.

Der eingesetzte Roboter soll selbständig Autoräder demontieren; der Ablauf der Objekterkennung eines solchen Demontageprozesses folgt einer *Grob-Fein-Strategie* [1, 2, 9] und gliedert sich prinzipiell in folgende Punkte:

1. Zunächst wird mit einem aktiven Stereokamerasystem ein niedrigaufgelöstes Übersichtsbild der gesamten Demontageszene (des ganzen Rades) aufgenommen.

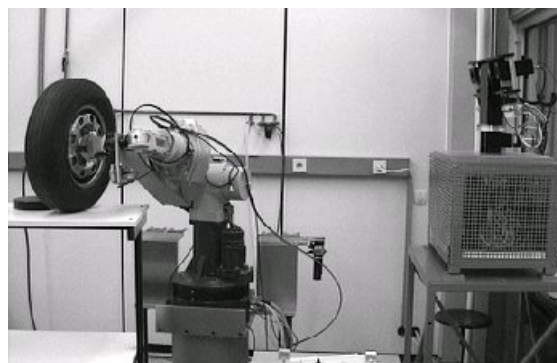
2. Durch Auswertung dieses Bildes mit Hilfe der neuronalen Netze und des semantischen Netzes werden Kameraparameter für detailliertere Aufnahmen berechnet.
3. Aufnahme von Detailaufnahmen von den zuvor berechneten Koordinaten zur genauen Objektklassifizierung und Vermessung der eigentlich zu demontierenden Objekte (der Schrauben).

Gegebenenfalls sind hierbei mehrere Detaillierungsebenen nötig, um die Objekte hinreichend genau zu klassifizieren.

Bevor die Verfahren zur Objekterkennung und 3D-Modellierung erläutert werden, folgt ein kurzer Einblick in die technische Ausstattung des Systems [3].

## 2 Systemarchitektur

Das System besteht aus einem Roboterarm mit sechs Freiheitsgraden; die Objekterkennung und -vermessung erfolgt mit einem aktiven Stereokamerasystem sowie einer zusätzlichen Handkamera an der Roboterhand. Die beiden Stereokameras befinden sich schräg hinter dem Roboter, um einen besseren Szenenüberblick zu erhalten und um Kollisionen mit dem Roboterarm zu vermeiden (s. Abb. 1). Die zusätzliche Handkamera dient gegebenenfalls der Aufnahme von Detailansichten, die wegen evtl. vorhandener Hindernisse nicht vom Stereokamerasystem aufgenommen werden können.



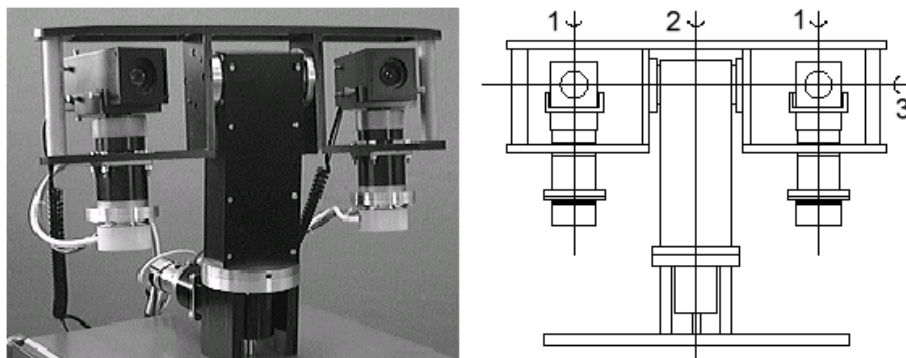
---

**Abbildung 1:** Demontagezene mit Roboterarm und Stereokamerasystem.

---

Das Stereokamerasystem besitzt neben drei optischen Freiheitsgraden (Blende, Fokus und Zoom) vier mechanische Freiheitsgrade, um die Kameras unabhängig voneinander positionieren zu können (s. Abb. 2):

1. je eine vertikale Drehachse,
2. eine gemeinsame vertikale Drehachse und
3. eine gemeinsame horizontale Neigeachse.



**Abbildung 2:** Das Stereokamerasystem.

Dieses Sehsystem ist aktiv, d. h. es ist in der Lage, seine intrinsischen und extrinsischen Parameter der Umwelt anzupassen. Diese Systeme haben gegenüber statischen Systemen den Vorteil, einen größeren Bereich erforschen zu können. Häufig sind solche aktiven Sehsysteme an biologische Systeme angelehnt und nutzen diese als Vorbild.

Die Kameras werden über ein Standard RS232-Protokoll von einer Sun SPARC-Workstation gesteuert; das gelieferte PAL-Signal kann mit Hilfe eines Frame-Grabbers digitalisiert werden [3].

### 3 Objektmodellierung und -erkennung

Gegenüber zweidimensionaler Bildverarbeitung ergeben sich bei der dreidimensionalen Bildverarbeitung folgende Schwierigkeiten [4]:

1. Es existieren mehr Freiheitsgrade bei der Positionierung eines Objekts,
2. die Objekte können sich gegenseitig verdecken, und

3. es ist jeweils nur eine Seite eines Objekts sichtbar – dies reicht häufig zur Klassifikation von (insbesondere ähnlichen) Objekten nicht aus.

Die heutzutage bei der Erkennung dreidimensionaler Objekte angewendeten Verfahren zur Objektbeschreibung lassen sich überwiegend in zwei Bereiche einteilen: die *objektbasierten* und die *blickpunktorientierten* Ansätze [5]. Bei den objektbasierten Verfahren wird ein explizites Objektmodell verwendet, bei denen häufig eine Zerlegung der Objekte in Objektprimitive (z. B. generalisierte Zylinder) erfolgt.

Die blickpunktorientierten Methoden bilden sich dagegen anhand mehrerer charakteristischer zweidimensionaler (Detail-)Bilder aus verschiedenen Blickrichtungen ein implizites Objektmodell. Vorteil dieser Vorgehensweise ist, daß das aufgenommene Bild direkt mit den gelernten zweidimensionalen Bildern verglichen werden kann.

Das in diesem Seminarvortrag vorgestellte Verfahren basiert ebenfalls auf der Modellierung durch mehrere Ansichten; dabei wird in einem Einzelbild das Objekt durch Kantenwolken beschrieben, um die gewünschte Erkennungs-Toleranz zu erhalten. In den nächsten beiden Abschnitten wird das hier eingesetzte Modellierungsverfahren genauer erläutert.

### 3.1 Hybride Objektmodellierung

Das hier vorgestellte System modelliert Objekte anhand *ganzheitlich* gelernter und wiedererkennbarer Objekt- und Teilobjektansichten; das Lernen dieser Ansichten erfolgt mit neuronalen Netzen. Eine zusätzliche wissensbasierte Komponente (ein semantisches Netz) dient der Modellierung des Objekts durch Verknüpfung dieser einzelnen Ansichten.

Bei der häufig im 3D-Bereich verwendeten Objektmodellierung anhand von Ecken, Linien und (geraden oder gekrümmten) Kanten erhält man in realen Szenen (mit komplexem Hintergrund) einen riesigen Suchraum.

Für die Modellierung anhand *ganzer Teilansichten* sprechen also folgende Gründe [1, 6]:

- Es genügt eine Beschränkung auf *wenige aussagekräftige Teilansichten*, wodurch der Modellierungsaufwand verringert wird.
- Die Teilansichten haben einen *hohen Informationsgehalt*; da sie jeweils nur Teil weniger Objekte sind, wird der Suchaufwand enorm eingeschränkt.

- Durch Zuordnung von *Positions- und Orientierungsinformationen* zu den Teil- und Objektansichten kann auf die Position und Lage des Gesamtobjekts geschlossen werden. Dadurch wird ebenfalls der Suchaufwand verringert.

### **Aufgabe des semantischen Netzes**

Daß die Ermittlung der notwendigen Kameraparameter für die weiteren Detailaufnahmen nicht allein datengetrieben erfolgen kann (d. h. anhand der Aufnahmen berechnet werden kann), ist klar, da sich bei einem dreidimensionalen Objekt objektspezifische Details häufig erst durch starke Blickwinkeländerung ergeben und daher nicht alle Aufnahmepunkte aus einer ersten Übersichtsaufnahme berechnet werden können.

Demzufolge ist für die Modellierung Vorwissen (in Form eines semantischen Netzes) notwendig. Anschaulich kann man sich die Aufgabe des semantischen Netzes so vorstellen, daß es dem aktiven Sehsystem bei der Entscheidung hilft, welche weiteren Aufnahmen eines Objekts nötig sind, um es zu klassifizieren. Es enthält über *Spezialisierungs-, Aspekt- und Teil-von-Relationen* miteinander verbundene (Teil-)Ansichten eines Objekts (s. Abb. 3).

Die Aspekt-Relationen verknüpfen dabei die charakteristischen Objektansichten, während die Teil-von-Relationen die charakteristischen Detailansichten mit einem Aspekt verbinden [4].

Bei der Erkennung wird das semantische Netz der top-down-Strategie entsprechend zu immer detaillierteren Ansichten „durchlaufen“. Dabei erfolgt auf jeder Ebene eine Modellierung des bisher erkannten Objekts, um die Aufnahmeparameter für die nächste Ebene zu erhalten. Wurde eine modellierte Struktur ganzheitlich erkannt, wird auf die Instantiierung ihrer Subnetze verzichtet, und es wird direkt eine Instanz des bisherigen Konzepts gebildet.

Kann das Objekt nicht klassifiziert werden (also auch nicht anhand der charakteristischen Detailaufnahmen eines gewählten Aspekts), so werden nacheinander die verschiedenen Aspekte und deren jeweilige Subnetze durchlaufen, bis eine Erkennung erfolgt ist oder das Objekt endgültig nicht erkannt wurde.

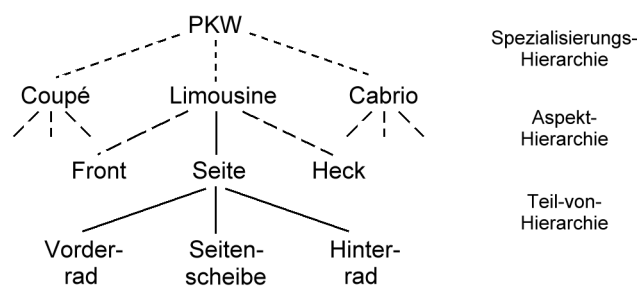
## Ein Beispiel: Erkennung eines PKW

Abb. 3 zeigt ein 3D-Objektmodell eines PKW mit verschiedenen Spezialisierungshierarchien.

Zunächst wird die gesamte Szene mit den Stereokameras als Übersichtsbild aufgenommen (oberste Ebene). Entweder kann durch die versuchte ganzheitliche Erkennung das Objekt bereits komplett identifiziert werden oder es wird eine Hypothese über die *Art* des Objekts generiert.

Parallel dazu wird eine erste grobe räumliche Tiefenkarte erstellt (s. Abschnitt 3.3), die eine Hypothese über die *Lage* des Objekts liefert.

Aufgrund dieser Hypothesen wird im semantischen Netz eine genauere Untersuchung im Subnetz „PKW“ angestoßen, und anhand der erhaltenen Kameraparameter werden nun Bilder von weiteren Aspekten (Front, Seite und Heck des PKW) aufgenommen. Entsprechend wird dann mit den Details der Seitenansicht verfahren.



**Abbildung 3:** Teil des semantischen Netzes zur Modellierung eines PKW.

Durch die Neuberechnung der Kameraparameter in jedem Detaillierungsschritt aus der vorhergehenden Hypothese können durch Fehlschätzungen bzw. aufgrund der Weitwinkelaufnahmen des Sehsystems entstehende Positionierungsfehler minimiert werden.

Die im Laufe des Erkennungsprozesses generierte Instanz des semantischen Netzes enthält schließlich die notwendigen Demontageparameter – bei der Demontage eines Autorades also z. B. Anzahl der Schrauben und Positionen von Felge und Schrauben.

## 3.2 Ganzheitliche Objekterkennung

Die Objekterkennung erfolgt *unabhängig von Hintergrund, Position und Orientierung*. Der Vorgang des Erkennens geschieht ganzheitlich, d. h. ein Ob-

jekt bzw. eine charakteristische Detailansicht wird mit neuronalen Netzen als Ganzes gelernt bzw. erkannt.

Anders als bei herkömmlichen Einsätzen von neuronalen Netzen, bei denen anhand einer relativ langwierigen Trainingsphase das Netz eine generalisierte Darstellung der zu erkennenden Objekte *lernt*, geschieht hier die Generalisierung schon durch die *Repräsentation* der Objekte selber.

Der generelle Erkennungsvorgang besteht hauptsächlich daraus, aus dem aufgenommenen Bild eine Darstellung zu generieren, die es ermöglicht, auch ohne Aufnahme des Objekts von einem *exakten* Blickpunkt aus das Objekt (ggf. mit Hilfe weiterer Detailaufnahmen) eindeutig zu klassifizieren. Diese hier eingesetzte sogenannte „extrafoveale Erkennung“ hat vor allem folgende Vorteile:

- Eine Segmentierung des Objekts kann entfallen – die wesentlich einfacher zu realisierenden fovealen Erkennungsverfahren setzen solch eine Segmentierung des Objekts voraus. Vor einem stark strukturierten Hintergrund bzw. bei Verschmutzungen und Beleuchtungsunterschieden ist dies aber – wenn überhaupt – nur sehr schwer möglich.
- Beim Anfahren der ermittelten Blickpunkte für die Detailaufnahmen des Objekts reicht eine Schätzung der Koordinaten aus, bzw. das System ist robust gegen leichte Abweichungen bei der Kamerakalibrierung.

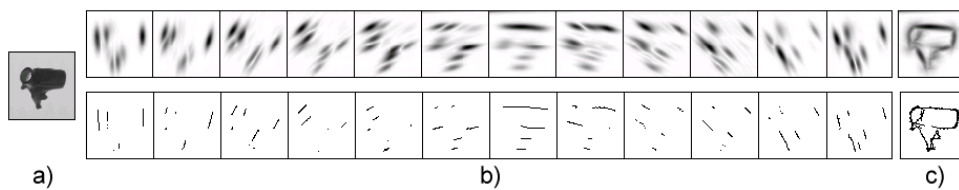
Für die Erkennung werden die zu demontierenden Objekte (die Schrauben) und die zu untersuchenden Szenen (das Rad) im Modell *unterschiedlich repräsentiert*: während die Schrauben mit einem möglichst schmalen Umriß (ihrem *Skelett*) dargestellt werden, werden die zu untersuchende Szene und deren Konturen wolkenförmig verbreitert dargestellt, um die Schrauben darin – auch trotz leichter Lageverschiebungen und verändertem Aussehen – zuverlässig wiederfinden zu können.

Beide Repräsentationen basieren auf der gleichen Vorverarbeitung, in der aus den aufgenommenen Bildern Informationen über die *Konturen* extrahiert werden, dies wird im folgenden Abschnitt näher erläutert.

### **Orientierungsinvarianz**

Zur *orientierungsinvarianten* Objekterkennung werden sog. orientierte Gaborfilter eingesetzt, die nur die Konturen entsprechend ihrer Orientierung „durchlassen“ [7]. Es werden insgesamt zwölf Filter verwendet, die sich

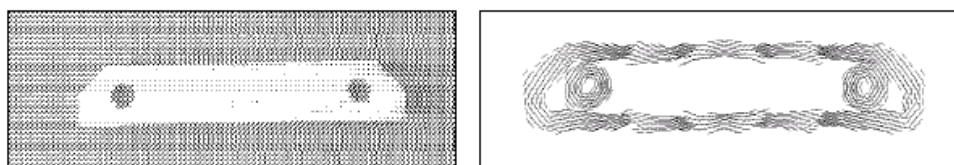
um jeweils  $15^\circ$  in ihrer Orientierung unterscheiden, wodurch sich ein Gesamtwinkel von  $180^\circ$  ergibt, der zur Erkennung jeder Orientierung ausreicht. Abb. 4 a) zeigt als Beispiel ein zu untersuchendes Objekt, in Abb. b) sind die Filterantworten und darunter ihre Skelette dargestellt. Abb. c) zeigt die aus den Filterantworten zusammengesetzten Konturinformationen.



**Abbildung 4:** Originalbild mit Filterantworten und deren Skelette.

Anschaulich kann man sich vorstellen, daß im neuronalen Netz sog. „einfache“ Neuronen mit kleinen, aneinander angrenzenden rezeptiven Feldern diese Gaborfilter realisieren; da benachbarte Neuronen auch benachbarte rezeptive Felder haben, ergibt sich damit eine linienförmige Konturrepräsentation (das Skelett, s. auch Abb. 6 c).

Die tolerante Repräsentation einer Szene erhält man durch sog. „komplexe“ Neuronen mit großen, gegenseitig stark überlappenden rezeptiven Feldern, wodurch sich eine wolkenförmige Konturrepräsentation ergibt (s. Abb. 5 und Abb. 6 b).

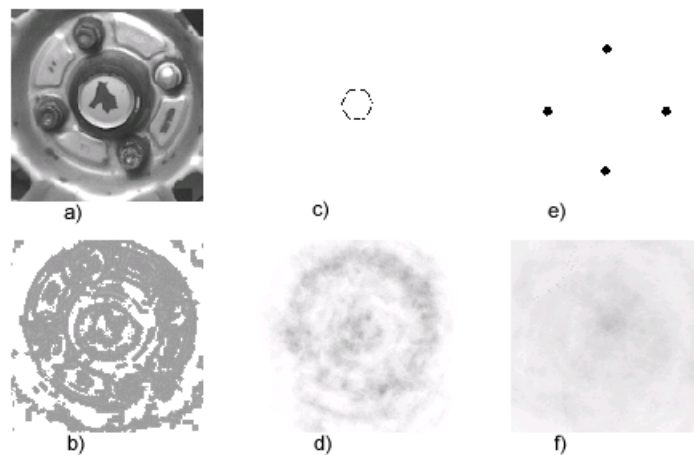


**Abbildung 5:** Ein Werkstück und zugehörige Konturwolke.

Diese Konturrepräsentation ist aufgrund ihrer Wolkenform tolerant in Bezug auf leichte Verschiebungen, Blickwinkeländerungen, Veränderungen der Beleuchtung etc., da auch dann noch die präsentierte Ansicht mit der gelernten weitgehend übereinstimmt [8]. Ein weiterer Vorteil dieser Darstellung ist, daß durch diese Art der Generalisierung das zeitaufwendige Trainieren der neuronalen Netze wesentlich verkürzt wird – es ist quasi ein „Lernen auf den ersten Blick“ möglich.

## Positionsinvarianz

Um ein Objekt im aufgenommenen Bild zu erkennen, wird dieses in überlappende Bildbereiche eingeteilt, für die jeweils ein Ähnlichkeitsmaß mit dem gesuchten Objekt berechnet wird. Der berechnete Wert wird in eine 2D-Matrix eingetragen (s. Abb. 6 d). Wird ein intern festgelegter Schwellwert überschritten, wurde das Objekt an dieser Stelle erkannt.



**Abbildung 6:** Zu untersuchende Szene; a) Grauwertdarstellung – b) Tolerante Repräsentation – c) Gelerntes Objekt in schmaler Repräsentation – d) 2D-Ergebnisspeicher für Teilobjektdetektion – e) Zusammensetzung des Gesamtobjektes aus Teilobjekten – f) 2D-Ergebnisspeicher für Gesamtobjektdetektion.

Zur Steigerung der Erkennungssicherheit und -geschwindigkeit wird zusätzlich eine modellbasierte Gruppierung von Konturelementen zu höheren Merkmalen eingesetzt (s. Abb. 6 e); ein Objekt wird damit anhand seiner Teilobjekte erkannt.

Außerdem werden Skalierungen und Rotationen der Einzelobjekte sowie des Gesamtobjekts zugelassen, womit die Erkennung gegenüber perspektivischen Verzerrungen und stark strukturierten Szenen robuster wird (s. Abb. 6 f).

Abb. 7 zeigt, daß mit dieser Technik die einzelnen Schrauben zu einem Schraubenkranz zusammengesetzt werden können.

Da die Schrauben jedoch aufgrund von Verschmutzungen kaum Konturen aufweisen, ist eine *genaue* Lagebestimmung, wie sie zum Greifen mit dem Roboterarm erforderlich ist, nur mit dem weiter oben in diesem Abschnitt beschriebenen konturbasierten Verfahren nicht ausreichend. Zur



---

**Abbildung 7:** Mit Konturen eingezeichnete Teilobjekte kennzeichnen die ermittelte Lage des Gesamtobjekts und die bestbewertete Orientierung und Skalierung der zugehörigen Teilobjekte.

---

pixelgenauen Lokalisierung wird daher zusätzlich eine grauwertbasierte Erkennung der zu demontierenden Objekte angewendet.

Diese Erkennung basiert auf dem Vergleich der aufgenommenen Schraube mit Bildern von zuvor aufgenommenen und generalisierten Schraubenköpfen – die Generalisierung ermöglicht dabei auch die Positionsbestimmung nicht in der Fallsammlung enthaltener Schrauben. Dieses Verfahren allein genügt aufgrund der Generalisierung nicht zur Schraubenerkennung, sondern nur zur genauen Positionsermittlung. Damit ergänzen sich die konturbasierte und grauwertbasierte Methode gegenseitig.

### 3.3 Rekonstruktion der räumlichen Tiefe

Um ein zuverlässiges Erkennen und Greifen der zu demontierenden Objekte zu ermöglichen, ist eine genaue Bestimmung der Lage und Position der Objekte notwendig; dies erfordert eine *realistische dreidimensionale Darstellung* der ganzen Demontageszene.

Die Rekonstruktion der Tiefeninformation anhand der Stereobildpaare wird parallel zur Erkennung (s. Abschnitt 3.2) durchgeführt [1, 9].

Dafür ist es notwendig, die Verschiebung korrespondierender Bildpunkte in den Bildern zu ermitteln – diese ist allerdings in der Regel sowohl horizontal als auch vertikal ungleich Null. Um den Rechenaufwand zu verringern, werden in einem ersten Schritt die Bilder dahingehend transformiert, daß die vertikale Verschiebung gleich Null ist.

Die anschließende Auswertung erfolgt (analog zur Erkennung) konturbasiert – auch hier wird zur Merkmalsextraktion ein Satz orientierter Gaborfilter eingesetzt.

Durch Bildung eines Ähnlichkeitsmaßes wird die Ähnlichkeit der beiden Bilder in lokalen Bereichen bestimmt. Anhand nichtlinearer Differentialgleichungen kann mit Hilfe dieses Maßes die Disparität feiner als das Pixelraster der Bilder berechnet werden, wodurch der mittlere quadratische Fehler um bis zu 60 % reduziert werden kann.

Bevor der Roboter „in Aktion treten“ kann, ist eine Fusion der dreidimensionalen Tiefeninformationen mit den zweidimensionalen Daten der Objekterkennung notwendig. Die extrahierten dreidimensionalen Koordinaten können als grobe Abtastung der Objektoberfläche aufgefaßt werden; durch Bilden von Ausgleichsebenen durch diese Punktmenge können fehlerhafte Punkte eliminiert werden.

Alle ermittelten Daten beziehen sich auf ein Koordinatensystem, dessen Ursprung in der Mitte zwischen den Projektionszentren der Kameras liegt; es ist also eine abschließende Transformation dieser Koordinaten in ein Roboterkoordinatensystem notwendig.

## 4 Ergebnisse

Bisher wurde der beschriebene Demontageprozeß erst zu Demonstrationszwecken implementiert – die Ergebnisse ergaben aber bereits in mehreren Testläufen, daß die mit den beschriebenen Verfahren erreichte Genauigkeit ausreicht, um ein Autorad zu demontieren.

In der Bildebene liegt die Streuung des Fehlers unter 1 mm, in der Tiefe bei ca. 2 mm (der Abstand zwischen Kamerakopf und Roboter beträgt ca. 1,7 m). Bei Tests mit verkippten Autorädern im Bereich von  $\pm 15^\circ$  konnte der Kippwinkel auf weniger als  $1^\circ$  genau bestimmt werden.

Bei den Tests konnten die Radmuttern sicher gegriffen werden, so daß dieses Verfahren zur Steuerung eines autonomen Roboters geeignet ist.

Das vorgestellte System zeigt eine interessante Möglichkeit, die Vorteile von neuronalen Netzen (robuste und tolerante Erkennung sowie Lernfähigkeit) mit den expliziten Objektbeschreibungen von wissensbasierten Systemen zu kombinieren. Die bereits in den ersten Testphasen erzielten Genauigkeiten machen deutlich, daß für einen Demontageroboter als einziger Sensor-Input ein aktives Stereosehsystem ausreicht, um die gewünschten Aufgaben autonom durchführen zu können.

Es bleibt allerdings abzuwarten, inwiefern sich ein industrieller Einsatz tatsächlich als sinnvoll (kostensparender und effizienter) erweist, da die Erprobung erst auf einige wenige Beispielszenarien angewendet wurde.

## Literatur

- [1] BÜKER, U.; DRÜE, S.; GÖTZE, N.; HARTMANN, G.; STEMMER, R.; TRAPP, R.: Aktive Objekterkennung und -vermessung zur Steuerung eines Demontageroboters. *Künstliche Intelligenz 1/99*, 1999, 25–30.
- [2] GÖTZE, N.; STEMMER, R.; TRAPP, R.; DRÜE, S.; HARTMANN, G.: Ein aktives Stereokamerasystem zur Steuerung eines Demontageroboters. In: *Posch, S.; Ritter, H. (Hg.): Dynamische Perzeption. Proceedings in Artificial Intelligence, Vol. 8.* infix-Verlag, Sankt Augustin, 1998, 1–8.
- [3] TRAPP, R.; DRÜE, S.: Ein flexibles binokulares Sehsystem: Konstruktion und Kalibrierung. In: *Mertsching, B. (Hg.): Aktives Sehen in technischen und biologischen Systemen. Proceedings in Artificial Intelligence, Vol. 4.* infix-Verlag, Sankt Augustin, 1996, 32–39.
- [4] BÜKER, U.: Hybrid Object Models: Combining Symbolic and Subsymbolic Object Recognition Strategies. In: *Callaos, N.; Omolayole, O.; Wang, L. (Hg.): Proceedings of 4th International Conference on Information Systems, Analysis and Synthesis, Vol. 1 (ISAS'98).* Orlando, 1998, 444–451.
- [5] BÜKER, U.; HARTMANN, G.: 3-D Objekterkennung durch aktives Sehen und Mehrfachansichten. In: *Paulus, E.; Wahl, F.M. (Hg.): Mustererkennung 1997.* Informatik aktuell, Springer-Verlag, Berlin, 1997, 475–482.
- [6] BÜKER, U.; KALKREUTER, B.: Automatische Generierung hybrider Objektmodelle. In: *Levi, P.; Ahlers, R.-J.; May, F.; Schanz, M. (Hg.): Mustererkennung 1998.* Informatik aktuell, Springer-Verlag, Berlin, 1998, 572–579.
- [7] ZIEGLER, U.; GÖTZE, N.; DRÜE, S.; HARTMANN, G.; STEMMER, R.; BAHRO, J.: Extrafoveale Erkennung von Industrieobjekten vor beliebig strukturiertem Hintergrund durch ein aktives Sehsystem. In: *Groß, H.M. (Hg.): Selbstorganisation von Adaptivem Verhalten.* VDI-Verlag, 1997, 186–193.
- [8] BÜKER, U.; DUNKER, J.; HARTMANN, G.; SEIDENBERG, E.: Aktives Sehen für die 3-D Objekterkennung in hybrider Architektur. In: *Mertsching, B. (Hg.): Aktives Sehen in technischen und biologischen Systemen. Proceedings in Artificial Intelligence, Vol. 4.* infix-Verlag, Sankt Augustin, 1996, 174–181.
- [9] STEMMER, R.; TRAPP, R.; DRÜE, S.; HARTMANN, G.: Dreidimensionale Lagebestimmung zu demontierender Objekte mit einem aktiven Stereokamerasystem. In: *Levi, P.; Ahlers, R.-J.; May, F.; Schanz, M. (Hg.): Mustererkennung 1998.* Informatik aktuell, Springer-Verlag, Berlin, 1998, 253–260.
- [10] Im WWW unter <http://get.uni-paderborn.de/>.