

Entwurf verhaltensbasierter visueller Systeme

Gerald Sommer

Institut für Informatik und Prakt. Mathematik

Christian-Albrechts-Universität zu Kiel

Zusammenfassung

Es werden thesenhaft die in der Forschungsgruppe Kognitive Systeme der CAU Kiel angewendeten Paradigmen zum Entwurf verhaltensbasierter visueller Systeme zusammengefaßt. Diese Thesen beschränken sich auf den theoretischen Aspekt der Formulierung allgemeiner Lösungen für ein spezielles Problem.

Der Entwurf verhaltensbasierter visueller Systeme hat zum Ziel, technische visuelle Systeme mit solchen attraktiven Eigenschaften wie Robustheit bezüglich nicht exakt definierbarer Einsatzbedingungen, Adaptivität bezüglich sich ändernder Einsatzbedingungen sowie Echtzeitfähigkeit zur Verfügung zu stellen. Die Konzeption verhaltensbasierter visueller Systeme stellt eine konsequente Erweiterung des Paradigmas "Aktives Sehen" dar [1]. Verhalten als beobachtbare Äußerungsform des Wahrnehmungs-Handlungs-Zyklus (WHZ) bildet das Modularisierungskonzept derartiger Systeme.

Nach Aloimonos [2] besteht das zu lösende Problem in Folgendem:

"Finde eine allgemeine Lösung für ein spezielles Problem".

Hierunter ist zu verstehen, daß "Sehen" nur als von der Körperlichkeit und Situiertheit des Beobachters determiniertes spezielles Problem gelöst werden kann. Dies bedeutet aber nicht, daß nur spezielle Heuristiken zur Problemlösung beitragen. Vielmehr ist zu vermuten, daß bisher nicht hinreichend ausgearbeitete allgemeine Lösungswege genutzt werden können, um spezielle Sehleistungen technischer Systeme zu realisieren. Diesem Thema ist der vorliegende Beitrag gewidmet.

Der Entwurf verhaltensbasierter visueller Systeme stellt in erster Linie eine Herausforderung naturwissenschaftlichen und theoretischen Charakters dar. Ihre Realisierung hat eine ausgeprägte ingenieurwissenschaftliche Komponente, die hier ausgeklammert werden soll.

Im Folgenden werden thesenhaft Positionen des Autors zusammengestellt, die als Grundlage zum Entwurf verhaltensbasierter visueller Systeme genutzt werden.

T1: Im verhaltensbasierten Entwurf visueller Systeme verschmelzen Computer Vision und Robotik zu einer Einheit, dem WHZ. Dies erfordert, daß deren Beschreibungsrahmen vereinheitlicht werden müssen [3].

T2: Signaltheorie und Neuroinformatik haben den Repräsentationsrahmen des WHZ zur Verfügung zu stellen. Dies leisten sie zur Zeit nur unzureichend [4].

T3: Zentrales Problem des WHZ ist die Repräsentation von geometrischen Kategorien der Raum-Zeit. Die verfügbare Sprache der Geometrie ist aber meist deskriptiv und symbolisch.

T4: Die Wahrnehmung von Bewegung als primäre visuelle Aufgabe wird durch die Kopplung mit okulomotorischem Verhalten einfacher. Dies gilt insbesondere in log-polarer Geometrie.

T5: Die Geometrische Algebra (geometrische Interpretation der Clifford Algebra) bietet den geeigneten Repräsentationsrahmen für Computer Vision, Robotik, Neuroinformatik und Signaltheorie mit folgenden Konsequenzen:

1. Die Beschreibung von Prozessen (Strukturen und Bewegungen) in der mehrdimensionalen Raum-Zeit ist intrinsisch gegeben.
2. Strukturkonzepte und Bewegungskonzepte sind äquivalent (gleichberechtigt und vertauschbar).
3. Probleme, die sich in der linearen Algebra als nichtlinear darstellen, werden zu linearen transformiert.
4. Die symbolische und numerische Komplexität wird reduziert und damit werden die Echtzeitfähigkeit und die neuronale Lernbarkeit geometrischer Konzepte (Überwindung des Dimensionalitätsproblems) verbessert.

T6: Die Theorie der Lie-Gruppen ist geeignet, die grundlegenden Prozesse der frühen visuellen Wahrnehmung und der Kinematik nach einheitlichen Gesichtspunkten in einer skalierbaren Raum-Zeit zu beschreiben:

1. Wahrgenommene und durch motorische Aktivität erzeugte Muster in der Raum-Zeit stellen Orbits der erzeugenden Operatoren dar.
2. Die erzeugenden Operatoren definieren das Alphabet der irreduziblen Invarianten von Wahrnehmung und Bewegung.
3. Ereignisse oder Mustergrenzen (allgemein Symmetriebrüche) induzieren das Umschalten des Systems auf andere Orbits.
4. Mustererkennung und Robotik stellen sich als Problem der Realisierung solcher Orbit-Sequenzen dar.

T7: Interpolation und Approximation von Mannigfaltigkeiten sind geeignete generische Zugänge für implizite und effiziente Erkennung und Realisierung geometrischer Konzepte in Signaltheorie und Neuroinformatik.

T8: Die Theorie steuerbarer Filter [5] stellt einen leistungsfähigen Rahmen für die Prozesse der frühen visuellen Wahrnehmung dar:

1. Sie ist im Rahmen der Theorie der Lie-Gruppen realisierbar.
2. Sie gestattet, die Freiheitsgrade eines aktiven visuellen Beobachters zu erfassen.
3. Sie gestattet die zielorientierte Einbettung der Anwendung von visuellen Operatoren und verbindet damit die Signaltheorie mit dem Paradigma des aktiven Sehens.
4. Sie erlaubt, die numerische Komplexität der Operationen stark zu reduzieren (durch Interpolation und Approximation).

T9: Dynamische Zellstrukturen sind geeignete neuronale Strukturen für Hybridrealisierungen des WHZ:

1. Sie nutzen die guten Approximationseigenschaften von RBF-Netzen.
2. Sie passen ihre Topologie an diejenige der zu repräsentierenden Mannigfaltigkeit an.
3. RBF-Knoten lassen sich fehlergetrieben adaptieren.
4. Gleichzeitiges Training von verdeckter und Outputsicht vermeidet Suboptimalität.
5. Die Abbildungen haben die Potenz der Echtzeitfähigkeit.

T10: Prozesse der frühen visuellen Wahrnehmung können als neuronale Schichtstruktur progressiv zielorientiert entworfen werden. Dies erlaubt

1. die Unterstützung von attentivem Sehen (und Blickverfolgung),
2. top-down Modulation von bottom-up getriebener Erkennung,
3. schrittweise Aggregation Gibson'scher Invarianten aus einem Satz irreduzibler Invarianten,
4. kontinuierlich anpaßbare progressive Maskenanpassungsfilter nach dem Steuerbarkeitsprinzip,
5. effiziente progressive Rekonstruktion als Wavelet-Netz über irregulären Abtastpunkten.

Die zu diesen Arbeitsthesen in den letzten beiden Jahren erreichten Ergebnisse sind in [6] als Forschungsbericht zusammengefaßt. Dort findet man auch die entsprechenden Referenzen.

Literatur

- [1] Sommer, G.: Verhaltensbasierter Entwurf technischer visueller Systeme, Künstliche Intelligenz, Heft 3 (1995) 42-45
- [2] Aloimonos, Y.: Active vision revisited. In: Active Perception (Ed. Y. Aloimonos), Lawrence Erlbaum, Hillsdale, 1993, 1-18
- [3] Sommer, G.: Geometric Algebra – a framework for the perception-action cycle. Workshop Theoret. Foundations of Computer Vision. Schloß Dagstuhl, März, 1996.
- [4] Sommer, G.: Neural approaches to approximation and recognition. ALCATECH Workshop, Sjaellands Odde, Juli, 1996.
- [5] Michaelis, M., G. Sommer: Steerable filters in finite dimensional function spaces eingereicht J. of Visual Commun. and Image Representation, 1996.
- [6] Sommer, G. et al.: Report of the Research Group Cognitive Systems, Christian-Albrechts-Universität zu Kiel, 1996. <http://www.informatik.uni-kiel.de/inf/Sommer/bericht96/bericht96.html>