



Jahresübersicht 2001

Veröffentlichungen des Jahres 2001 inklusive aller verfügbaren Abstracts

T. Ahlers, G. Albe, R. Frauli, S. Koch, N. Leßmann, C. Oßieck, J. Schütte, B. Jung, J.-T. Milde:
Bart+Rosie: A Simulation Environment and Architecture Testbed for Instructable Robotic Agents

In V. Paelke & S. Volbracht (eds.):
*Proceedings User Guidance in Virtual Environments:
 Workshop on Guiding Users through Interactive Experiences,*
 (C-LAB Publications; Bd. 8), Shaker Verlag, 2001, 137-142.

C. Bauckhage, H. Rieser, B. Jung und I. Voß:
**Struktur, Semiotik und Dynamik -
 ein Überblick über Aspekte und Methoden der Modellierung mechanischer Aggregate.**

SFB 360 Report 01/01, Universität Bielefeld.

Abstract:

Im Konstruktionsszenario des SFB 360 bilden Objektaggregate den Kern vieler Fragestellungen und müssen aus verschiedensten Perspektiven betrachtet werden: Für den Instrukteur sind Aggregate ein Planungsproblem. Er muß eine Sequenz auszuführender Konstruktionsschritte festlegen oder gegebenenfalls eine eigentlich vorgesehene Reihenfolge modifizieren. Der Konstrukteur hingegen nimmt Aggregate vor allem als Zwischenschritte eines Konstruktionsprozesses wahr; für ihn sind sie Resultate von Manipulationen, die der Instrukteur in seinen Montageanweisungen vorgibt. Im kooperativen Dialog zwischen Instrukteur und Konstrukteur erproben Aggregate außerdem die Lernfähigkeit des Konstrukteurs. Versucht dieser Instruktionssequenzen zu antizipieren, kann der Konstruktionsverlauf beschleunigt werden. In dieser Hinsicht kommt Aggregaten darüberhinaus eine wichtige semantische und pragmatische Rolle zu. Durch Einführung neuer Bezeichnungen für komplexe Objekte werden intrinsische Orientierungen etabliert, die die Dialogpartner zur Koordination nutzen, so daß sich folgende Konstruktionsanweisungen weniger umständlich formulieren lassen. Im Rahmen der technischen Realisierung eines künstlichen Kommunikators stellt die Einführung von Aggregatsbenennungen natürlich auch ein Problem der Spracherkennung und -verarbeitung dar. Ferner tauchen Aggregate als zu erkennende Strukturen in Bildsignalen auf, so daß bestehende Zusammenhänge zwischen erkannten verbalen oder gestischen Bezeichnungen und in Bildern gefundene Strukturen ermittelt werden müssen. In der Simulation von Konstruktionsprozessen mit VR-Techniken stellt die Rekonstruktion physikalischer Eigenschaften dynamisch erzeugter Aggregate ein entscheidendes Problem dar. Sowohl in virtuellen als auch in realen Montageumgebungen müssen semantische Aspekte behandelt werden, da sich die Frage stellt, ob und wann ein Aggregat als Modell eines realen Gegenstandes, d.h. als Instanz eines Konzepts, zu sehen ist. Schließlich erweisen sich Aggregate im Hinblick auf die physikalische Manipulation durch einen Roboter auch als ein Problem der Griiffs- und Trajektorienplanung.

Die Realisierung eines kooperativen, künstlichen Kommunikators in einer Konstruktionsdomäne erfordert also, geeignete Formalismen zur Repräsentation mechanischer Aggregate bereit zu stellen, die all den aufgeführten Aspekten situierter, flexibler Montage gerecht werden. Schwierig erscheint dabei vor allem die integrierte Behandlung der verschiedenen Modalitäten, zumal es weltweit wenig andere Projekte gibt, die sich ähnlichen Fragen widmen und somit kaum auf Ergebnisse früherer Arbeiten zurückgegriffen werden kann. Allerdings sind einige der oben genannten Aspekte nicht nur bei der Entwicklung eines künstlichen Kommunikators von Interesse, sondern stellen auch gängige Probleme aus den Bereichen der robotischen Montage und der Automatisierungstechnik dar.

Tatsächlich blickt die Erforschung und Entwicklung intelligenter Montageroboter mittlerweile auf eine lange Tradition zurück. Bereits 1975 stellten Ambler et al. [2] ein situatives System vor, das - mit kognitiven Fähigkeiten ausgestattet - in der Lage war, einfache Spielzeugobjekte in einer nicht normierten

Montageumgebung zu komplexeren Aggregaten zusammensetzen. Nur wenig später haben Nevins und Whitney [70] erstmals systematisch untersucht, welche Kriterien die automatische Planung von industriellen Konstruktionsprozessen zu erfüllen hat und welche Anforderungen an entsprechende Repräsentationstechniken gestellt werden müssen. Seit diesen frühen Arbeiten sind die computergestützte Modellierung und Planung der Aggregation einzelner Objekte Gegenstand intensiver Forschungsbemühungen, die hauptsächlich durch wirtschaftliche Fragestellungen motiviert sind. Dementsprechend finden die bisher entwickelten Verfahren ihren Einsatz vor allem in industriellen CAD- und CAM/CIM-Systemen.

Auch wenn die Planung und Montage von Aggregaten nur ein kleines Teilgebiet der Forschung zur künstlichen Intelligenz abdeckt, ist die Vielzahl der in den letzten Jahren hierzu publizierten Arbeiten nur noch schwerlich überschaubar. Der folgende Überblick erhebt deswegen keinesfalls den Anspruch auf Vollständigkeit, vielmehr sollen nur einige mittlerweile als Standard geltende Techniken sowie Ansätze mit einschlägigem Bezug zu den im SFB 360 behandelten Problemen kurz skizziert werden. Da diese Übersicht aus Literaturstudien, die sich an Fragen des B1 Projekts orientieren, resultiert, werden Ergebnisse aus den Gebieten der Griiffs- und Trajektorienplanung nicht mit aufgeführt, obwohl diese für einige Teilprojekte des SFB von großem Interesse sind. Daher sei auf weitere zusammenfassende Darstellungen über Forschungen zur Thematik der Aggregatmodellierung, Griiffs- und Trajektorienplanung hingewiesen, die sich in [35, 29, 39, 98] und in den Tagungsbänden der IEEE Conference on Robotics and Automation sowie des IEEE Symposiums on Assembly and Task Planning finden.

Im folgenden wird zunächst definiert, was in der Literatur gemeinhin unter "mechanischen Aggregaten" verstanden wird und welche Untereinheiten komplexer Objekte unterschieden werden. Im Abschnitt 3 werden dann verschiedene gängige Methoden zur Repräsentation mechanischer Aggregate beschrieben und Ergebnisse zur mathematischen Komplexität dieser Methoden aufgeführt. Techniken zur automatischen Beobachtung bzw. zur Simulation von Konstruktionsprozessen sind Gegenstand des vierten Abschnitts. Der Bezug der dort vorgestellten Arbeiten zur Thematik des SFB 360 wird anschließend ersichtlich, denn in Abschnitt 5 werden Arbeiten, die sich im Rahmen des SFB mit Fragen der Aggregation auseinandersetzen, dargestellt. Dies sind insbesondere Arbeiten aus den Teilprojekten B1, B3 und C1, so daß vor allem Methoden zur visuellen Aggregaterkennung, zur Beschreibung von Metonymien für komplexe Objekte, zur dynamischen Konzeptualisierung im Verlauf von Konstruktionsprozessen sowie zur Aggregaterkennung anhand prototypischer Formcharakteristika vorgestellt werden. Ein Vergleich der in den einzelnen Teilprojekten entwickelten Ansätze und Techniken und eine Zusammenfassung beschließen diesen Report.

M.E. Latoschik: A General Framework for Multimodal Interaction in Virtual Reality Systems: PROSA.

In: Broll, W & Schäfer, L. (Editors):
The Future of VR and AR Interfaces - Multimodal, Humanoid, Adaptive and Intelligent.
Proceedings of the workshop at IEEE Virtual Reality 2001, Yokohama, Japan.
GMD Report No. 138, March 2001.

Abstract:

This article presents a modular approach to incorporate multimodal gesture and speech driven interaction into virtual reality systems. Based on existing techniques for modelling VR-applications, the overall task is separated into different problem categories: from sensor synchronisation to a high-level description of crossmodal temporal and semantic coherences, a set of solution concepts is presented that seamlessly fit into both the static (scenograph-based) representation and into the dynamic (renderloop and immersion) aspects of a realtime application. The developed framework establishes a connecting layer between raw sensor data and a general functional description of multimodal and scenecontext related evaluation procedures for VR-setups. As an example for the concepts, their implementation in a system for virtual construction is described.

M. E. Latoschik: A Gesture Processing Framework for Multimodal Interaction in Virtual Reality.

Afrigraph 2001, 1st International Conference on Computer Graphics, Virtual Reality and Visualization in Africa, 5 - 7 November 2001.

Abstract:

This article presents a gesture detection and analysis framework for modelling multimodal interactions. It is particularly designed for its use in Virtual Reality (VR) applications and contains an abstraction layer for different sensor hardware. Using the framework, gestures are described by their characteristic spatio-temporal features which are on the lowest level calculated by simple predefined detector modules or nodes. These

nodes can be connected by a data routing mechanism to perform more elaborate evaluation functions, therewith establishing complex detector nets. Typical problems that arise from the time-dependent invalidation of multimodal utterances under immersive conditions lead to the development of pre-evaluation concepts that as well support their integration into scene graph based systems to support traversal-type access. Examples of realized interactions illustrate applications which make use of the described concepts.

T. Sowa, S. Kopp & M. E. Latoschik: A Communicative Mediator in a Virtual Environment: Processing of Multimodal Input and Output.

International Workshop on Information Presentation and Natural Multimodal Dialogue, Verona, Italy, 2001.

Abstract:

This paper presents work on multimodal communication with an anthropomorphic agent. It focuses on processing of multimodal input and output employing natural language and gestures in virtual environments. On the input side, we describe our approach to recognize and interpret co-verbal gestures used for pointing, object manipulation, and object description. On the output side, we present the utterance generation module of the agent which is able to produce coordinated speech and gestures.

T. Sowa, I. Wachsmuth: Coverbal Iconic Gestures for Object Descriptions in Virtual Environments: An Empirical Study.

SFB 360 Report 01/03, Universität Bielefeld.

Abstract:

This paper describes an empirical study aimed at investigating object references in Virtual Environments using iconic gestures. Observations are focused on spatial concepts conveyed gestually and their relation to features of the gesture shape. A set of important features and spatial concepts useful for automated gesture recognition is identified. Based on these findings we propose a model of an iconic reference recognizer.

I. Voss: Anticipation in construction dialogues.

In J. Vanderdonckt, A. Blandford & A. Derycke (eds.):
*"Interaction without frontiers",
Proceedings of Joint AFHMC-BCS Conference on Human-Computer Interaction IHM-HCI'2001
(Lille, France, Sept. 2001), Vol. II, pp. 189-190, Cépaduès--Editions, Toulouse, 2001.*

I. Wachsmuth: Timing and Rhythm in Multimodal Communication for Conversational Agents.

In: *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*
(Edinburgh, Scotland, August 1-4, 2001), member abstracts, 2001.

I. Wachsmuth, I. Voss, T. Sowa, M.E. Latoschik, S. Kopp, B. Jung: Multimodale Interaktion in der Virtuellen Realität

In: H. Oberquelle, R. Oppermann, J. Krause (Hrsg.), *Mensch & Computer 2001* (pp. 265-274),
Stuttgart: Teubner, 2001.

Abstract:

Virtuelle Realität oder Virtual Reality (VR) bezeichnet ein neuartiges Kommunikationsmedium, das die unmittelbare Wechselwirkung des Menschen mit räumlich organisierten rechnergenerierten Darstellungen erlaubt. Verbunden mit körperlich verankerter Interaktion finden insbesondere gestische Eingaben starkes Interesse. Dieser Beitrag gibt einen Überblick über Forschungsarbeiten im Labor für Künstliche Intelligenz und Virtuelle Realität an der Universität Bielefeld, mit denen Grundlagen für den Einsatz gestischer und sprachlicher Interaktionstechniken entwickelt werden; als Erprobungsdomäne dient ein Szenario des virtuellen Konstruierens. Für die schnelle Erfassung komplexer Hand-Armgesten werden derzeit Datenhandschuhe und

Körper-Tracker eingesetzt. Die Auswertung erfolgt mit wissensbasierten Ansätzen, die atomare Formelemente der Gestik symbolisch beschreiben und zu größeren Einheiten zusammensetzen. Ein zweites Thema ist die multimodale Interaktion durch sprachlich-gestische Eingaben, z.B. wenn auf einen Gegenstand gezeigt ("dieses Rohr") oder eine Drehrichtung ("so herum") signalisiert wird. Schließlich wird dargestellt, wie die Ansätze zur Formbeschreibung von Gesten für die Synthese natürlich wirkender gestischer Ausgaben mit einer artikulierten, anthropomorphen Figur übertragen werden können, die in laufenden Arbeiten mit Sprachausgaben koordiniert werden.

A. Kranstedt, 28.07.2003