

# Situated Verbal Interaction in Virtual Design and Assembly

Yong Cao, Bemhard Jung, Ipke Wachsmuth  
AI & Computer Graphics Lab  
Faculty of Technology  
University of Bielefeld  
D-33501 Bielefeld, Germany  
{yong, jung, ipke}@techfak.uni-bielefeld.de

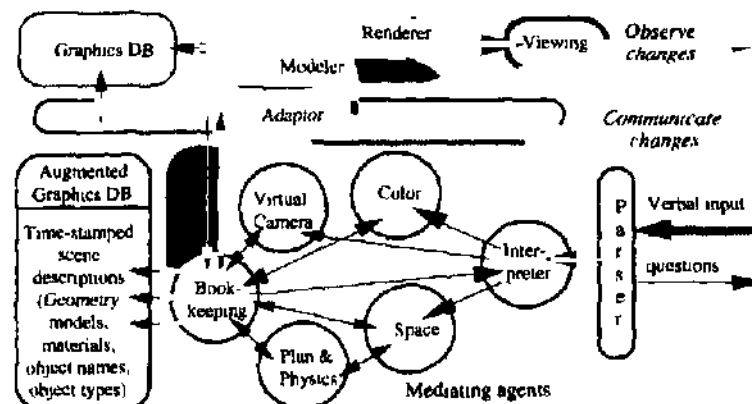
## Abstract

The video demonstrates research accomplishments in interactive design and assembly with 3D computer graphics environments. Agent techniques and dynamic knowledge representation techniques are used to process qualitative verbal instructions to quantitative scene changes. A key idea is to exploit situated 'perceptive' information by inspecting the computer graphics scene models.

Computer-based presentations of synthetic geometry data, transformed to visual surface structures by way of rendering techniques, are of growing importance in the design and construction areas. To make better profit of this new technology, new ways of human-computer interaction are called for. Previous work has concentrated on gesturing and pointing, for instance, by using the data glove. An alternative way we explore at the AI & Computer Graphics Lab is to use verbal interaction to communicate alterations. These are put in effect by a mediating system which changes the arrangement or assemblage of scene objects. Our general goal is to use AI to establish a communication link between humans and multimedia. This is, however, possible only when the image of a design or an assembly is not a meaningless visual presentation, but is coupled to an internal semantic representation of the presented images.

As the user gets immersed in the visual scene, verbal statements likely make reference to what can be seen in the current situation. Thus the semantic processing of verbal instructions needs to draw on what is true in a situation. We do not try to give a full propositional account of the scene, which might also be affected by direct mouse or glove manipulation, because of the known problems of "keeping track" (the frame problem). Rather, we exploit the current situation by inspecting the internal scene description (geometry models, materials, etc.). Since the user perceives the scene by eye inspection, system and user can communicate about scene details from the same "point of view." We call this *situated verbal interaction*.

The video illustrates our approach in two examples, one in virtual design and one in virtual assembly. It is our aim that user and system can communicate about a *dynamic* scene. To this end, we have developed dynamic knowledge representations to account for changing conceptualizations of objects and scene, and we use software interface agents which inspect graphic scene descriptions and keep track of varying situation parameters. Discourse in simple written natural language is used for communication (in the future we want to use voice input). Qualitative communications of the user are evaluated so as to produce appropriate quantitative changes in the scene description which is then visualized for the user.



Virtual Design Environment In the VIENA project ("Virtual Environments & Agents"), we have chosen interior design as an example domain, and we build on the SOFTIMAGE modeling and rendering software installed on a Silicon Graphics Indigo. Instead of using the complex mouse-and-menu commands to manipulate objects we communicate with the system by way of natural language. Our aim is to keep the user (designer) free from technical detail such as geometry planning. We have developed a set of software interface agents (see figure above), realized as autonomous Unix processes, which altogether form an intelligent mediator. Taking special responsibilities in processing verbal instructions,

the VIENA agents cooperate with each other and with the user to offer a goal scene meeting the user's wants. For instance, a space agent translates qualitative relations such as to appropriate scene coordinates. The offer can be changed in further interaction, that is, semantics construction is coupled into the visible scene and can be negotiated in discourse.

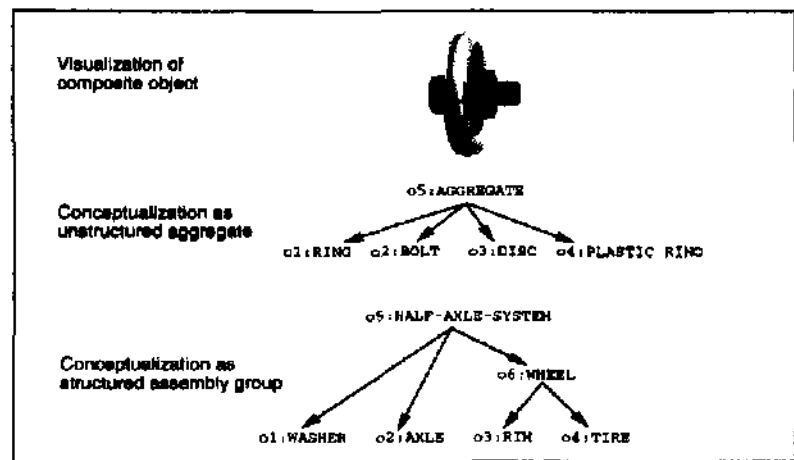
The agent framework developed in VIENA combines different cooperation models such as contract-net, master-slave, and blackboard. Each agent is equipped with a bidder-specific frame which triggers the activation of the agent's local functionality. Out of this, contractor-specific primitives are inserted depending on the situation and the task.

The VIENA system is tested in a prototype scenario with various items of furniture as well as color and light impressions of a virtual office room which can be changed interactively. Sample interactions are shown in the video.

**Virtual Assembly Workbench** In the CODY project ("Concept Dynamics"), we develop a knowledge representation scheme for dynamic object conceptualization in assemblage. In our testbed scenario, we assemble a toy airplane and similar constructs from building blocks on a virtual assembly bench. As shown in the video, the user can inspect the graphics scene from different perspectives and change the building blocks<sup>1</sup> configuration by way of natural verbal communication. Instructions can refer to visual object properties, such as their location and color, but also to conceptual structures that are superimposed dynamically on the geometry models.

Conceptual knowledge is defined in two knowledge bases, one describing the building blocks and their connection ports, the other one containing specific knowledge about the target aggregate's structured assembly groups. This background knowledge enables the user to assemble any complex aggregate from the building blocks with special support for predefined constructs. Concepts for assembly groups are defined by their parts and the relations between them, which are expressed through *logical* and *geometric* constraints. When instances of assembly groups are to be recognized, logical constraints require logical relations, e.g. , to be asserted between part representations, whereas geometric constraints, e.g. , trigger tests in the geometry models on an if-needed basis.

Every time objects are composed on the virtual assembly bench, logical relations such as are inferred from the geometric scene description. In an internal loop, these relations are propagated along the structure in the logical data base. Unstructured aggregate representations are created dynamically and matched against the model knowledge base. The logical description is restructured when instances of assembly groups are recognized. Furthermore, functional names and properties are ascribed to object representations de-



pending on their aggregate context. For example, a bolt functions as an axle when it is part of a half-axle-system (see figure above). Assembly operations, typically, are subject to side effects, i.e., they may affect objects not mentioned in instructions. The system detects these side effects and updates object representations accordingly.

**Conclusion** Natural communication with responsive virtual environments is a decisive factor in making this cutting-edge technology widely applicable. Complementing direct manipulation, situated verbal interaction seems a promising way to make better profit of the ways humans communicate. Our research achievements depend on the integration of various AI techniques, and they lead into various application areas beyond our testbed scenarios, for example, in virtual prototyping and in car assembly planning. At the same time, we found that our scenarios provide a springboard for goal-oriented research on dynamic knowledge representations and on software interface agents.

**Acknowledgment** Credit to further project personnel is given in the video. Research in the VIENA Project is partly supported by the Ministry of Science and Research of the Federal State North-Rhine-Westphalia. CODY is a project in the collaborative research center "Situated Artificial Communicators" at the University of Bielefeld and is in part supported by the German National Science Foundation (DFG).

## References

- [Jung et al., 1995] Bernhard Jung, Britta Lenzmann, and Ipke Wachsmuth. Interaktive Montagesimulation mit wissensbasierter Grafik. Technical Report SFB 360 95/6, University of Bielefeld, 1995. To appear in G. Lux and U. Schroeder, eds., *Proc. 8. ASIM Workshop Simulation und Kunstliche Intelligenz (Darmstadt, 5.-6. April 1995)*.
- [Wachsmuth and Cao, 1995] Ipke Wachsmuth and Yong Cao. Interactive graphics design with situated agents. In W. StraBer and F. Wahl, eds., *Graphics and Robotics*, pages 73-85. Springer, 1995.