Knowledge-Based Design of 3D Graphics and Virtual Worlds

Steven K. Feiner

Department of Computer Science
Columbia University
New York, New York 10027
212-939-7083
feiner@cs.columbia.edu

INTRODUCTION

There are many situations in which pictures are needed to fulfill a clearly stated purpose, such as describing the results of an experiment or explaining how to perform a physical task. While computer graphics research has given us fully automated solutions to rendering pictures, the process of designing pictures is typically addressed through the development of editors with which a user can express their design decisions. Designing pictures by hand is time-consuming, requires scarce expertise, and invariably must be done in advance of when the pictures are viewed. Consequently, the results are often overly general and do not meet the information needs of each individual viewer and situation. *Knowledge-based graphics* offers an alternative approach in which AI techniques are applied to automate the design of pictures to communicate specific information to a viewer.

A number of researchers have been investigating knowledge-based graphics across a broad range of application domains, including 2D diagrams [10, 11], 3D pictures [2], 2D [7] and 3D [14] animation, and data visualization [9]. In the following sections, we provide a brief overview of our work on knowledge-based 3D graphics and show how we have applied these ideas to the design of 3D virtual worlds.

KNOWLEDGE-BASED 3D GRAPHICS

What decisions must be made in designing a picture? If we take the standard 3D graphics pipeline as a given, we can couch this problem in terms of choosing a set of objects to depict, a lighting specification, a viewing specification, and a rendering specification. (We use the term *rendering specification* to refer to information about how objects should be rendered, such as the pixel resolution of the image, whether a line drawing or shaded image should be created, or whether shadows should be cast.) This is precisely the data that is required by conventional rendering software.

APEX (Automated Pictorial Explanations) [3] was our first attempt at automating the design of pictures intended to show how to perform simple physical tasks. Its input was a hierarchical polyhedral world model, information about the user and their familiarity with parts of the model, and a description of a task to be depicted. Its design approach built up a picture based on a theory of how to select and render objects as a function of the role that they performed in the picture.

Each of APEX's pictures crystallized around those objects that actually participated in an action being depicted. Additional objects were added to fulfill a variety of criteria for the design of effective pictures. For example, objects were added to provide context for an unfamiliar viewer, to disambiguate objects from others of similar appearance, to serve as "landmarks" to help locate other objects, or even because they provided physical support (e.g., a floor) for some object already included. As these decisions were made, changes were propagated to the lighting, viewing, and rendering specifications. For example, the viewing specification was modified so that an object that participated in an action was wholly within the view volume, whereas an object that merely supported another did not merit enlarging the view volume on its behalf.

Columbia's Computer Graphics and User Interfaces Laboratory has extended this work in three major directions: generalizing and improving the picture-making process, designing animated presentations, and generating pictures in conjunction with other media.

IBIS [13] (Intent-Based Illustration System) supports a much richer model of the picture-making process. It takes as input a set of *communicative goals*, specified in terms of a prioritized set of information about objects, their attributes, and changes in them that are to be depicted. It uses a rule-based approach to create pictures that can be *composite* (composed of multiple subpictures) and *dynamic* (allowing the user to modify aspects of the picture, such as its viewing specification, while the system incrementally redesigns it).

ESPLANADE [8] (Expert System for Planning Animation Design and Editing) designs explanatory animation. Given a full description of the objects and actions to be depicted, it develops a hierarchically structured narrative, using editing effects such as cuts and wipes to join together individual shots.

COMET [4] (Coordinated Multimedia Explanation Testbed) couples IBIS with a text generation component [12] to create a knowledge-based multimedia system whose output includes both text and graphics. An additional stage precedes the text and graphics generators, and determines which parts of the input communicative goals should be expressed in graphics, which in text, and which in both.

KNOWLEDGE-BASED VIRTUAL WORLDS

As described thus far, a picture, be it still or animated, presents a view of some larger world. Through a combination of high-performance graphics workstations and 3D display and interaction devices, we can instead try to present a user with the world itself. Although at any given time the user's view of such a *virtual world* is just a single picture (or a pair of pictures, in the case of stereo), the ability to manipulate the world and our relationship to it interactively, can create the illusion of something more. We have been exploring the knowledge-based design of virtual worlds for both abstract and concrete domains, paying special attention to the active role of the user/participant.

AutoVisual [1] designs virtual worlds for visualizing multivariate functions. The design approach attempts to take into account the interactivity of the resulting world. For example, if a function is displayed as a surface, the number of samples with which it is approximated is adjusted to achieve reasonable interactive performance. Since the world is intended to be explored, rather than passively viewed, part of the design process determines which of a set of tools to make available to allow the user to examine the world. (For example, a "dipstick" lets the user determine the precise value of a height field at a desired location.)

One especially promising kind of virtual world is an *augmented reality* in which the virtual world is intended to augment, rather than replace, the surrounding world. This can be accomplished through the use of a see-through head-mounted display whose graphics is overlaid on the user's view of the world. Systems of this sort could be exceptionally effective for maintenance and repair applications in which users must otherwise cope with the cognitive overhead of using documentation that is separate from, rather than integrated with, the physical domain that it explains.

KARMA (Knowledge-based Augmented Reality for Maintenance Assistance) [6] creates an augmented reality that explains simple end-user maintenance for a laser printer. We attached several 3D trackers to key components of a printer, allowing the system to monitor their position and orientation, so that the physical and virtual worlds can be registered. IBIS is the design component. Rather than designing pictures from scratch, its viewing specification is determined by the position and orientation of the user's head and its graphics must mesh with what the user already sees. For example, if one communicative goal is to show an object to the user, and the object is determined to be within the user's field of view and unobscured by other objects, IBIS need not generate any graphics. If the object were obscured, however, IBIS will draw it, allowing it to be seen through the obscuring objects.

CONCLUSIONS AND FUTURE WORK

We have presented a series of thumbnail sketches of knowledge-based graphics systems that design 3D pictures and virtual worlds. There are a number of directions in which we intend to take this work. One involves exploring real-time generation of animated presentations. For example, ESPLANADE plans an animation entirely in advance of presenting it, and thus has access to the full results of the planning process. Instead, we are interested in planning and displaying a presentation as the actions that it describes occur. This requires imposing hard real-time constraints on the presentation planning process and being willing to accept some predetermined amount of lag time to allow limited lookahead.

Another direction will extend COMET to include animation and both speech and non-speech audio [5]. Temporal constraints will be a key issue here. A presentation must not only fit within the time that is allocated, but the duration of each of its component parts must also be constrained so that desired temporal relationships can be achieved both within a single medium and across media boundaries.

ACKNOWLEDGMENTS

This research was supported in part by the Office of Naval Research under Contract N00014-91-J-1872, the New York State Center for Advanced Technology in Computers and Information Systems under Contract NYSSTF-CAT-92-053, the Columbia Center for Telecommunications Research under NSF Grant ECD-88-11111, NSF Grant CDA-90-24735, and equipment grants from the Hewlett-Packard Company, Ascension Technology, and Digital Image Design. Earlier work on IBIS and COMET was supported in part by the Defense Advanced Research Projects Agency under Contract N00039-84-C-0165.

REFERENCES

1. Beshers, C. and Feiner, S. AutoVisual: Rule-Based Design of Virtual Worlds for Interactive Multivariate Visualization. *IEEE Computer Graphics and Applications* 13(4), July, 1993.

2. Emhardt, J. and Strothotte, T. Hyper-Rendering. In *Proc. Graphics Interface '92*, pages 37–43. Vancouver, Canada, May 11–15, 1992.

3. Feiner, S. APEX: An Experiment in the Automated Creation of Pictorial Explanations. *IEEE Computer Graphics and Applications* 5(11):29–37, November, 1985.

4. Feiner, S. and McKeown, K. Automating the Generation of Coordinated Multimedia Explanations. *IEEE Computer* 24(10):33–41, October, 1991.

5. Feiner, S., Litman, D., McKeown, K., and Passonneau, R. Towards Coordinated Temporal Multimedia Presentations. In Maybury, M. (editor), *Intelligent Multimedia Interfaces*. AAAI Press, Menlo Park, CA, To appear in 1993.

6. Feiner, S., MacIntyre, B., and Seligmann, D. Knowledge-Based Augmented Reality. *Communic. ACM* 36(7), July, 1993.

7. Kahn, K. Creation of Computer Animation from Story Descriptions. PhD thesis, AI Lab, MIT, January, 1979. MIT AI Technical Report 540.

8. Karp, P. and Feiner, S. Issues in the Automated Generation of Animated Presentations. In *Proc. Graphics Interface '90*, pages 39–48. Halifax, Canada, May 14–18, 1990.

9. Kochhar, S. Friedell, M. and LaPolla, M. Cooperative, Computer-Aided Design of Scientific Visualizations. In Nielson, G. and Rosenblum, L. (editors), *Proc. IEEE Visualization '91*, pages 306–313. IEEE Computer Society Press, October, 1991.

10. Mackinlay, J. Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics* 5(2):110–141, April, 1986.

11. Marks, J. A Syntax and Semantics for Network Diagrams. In *Proc. IEEE 1990 Workshop on Visual Languages*, pages 104–110. Skokie, IL, October, 1990.

12. McKeown, K.R., Elhadad, M., Fukumoto, Y., Lim, J., Lombardi, C., Robin, J., and Smadja, F. Language Generation in COMET. In Mellish, C., Dale, R., Zock, M. (editors), *Current Research in Language Generation*. Academic Press, London, 1990.

13. Seligmann, D. and Feiner, S. Automated Generation of Intent-Based 3D Illustrations. In *Proc. ACM SIGGRAPH '91* (*Computer Graphics*, 25:4, July 1991), pages 123–132. Las Vegas, NV, July 28–August 2, 1991.

14. Zeltzer, D. Knowledge-based Animation. In *Motion: Representation and Perception* (*Proc. ACM SIGGRAPH/SIGART Workshop on Motion*), pages 187–192. Toronto, Canada, April 4–6 1983.