

Reprint: Presence Journal, Volume 2, Number 4, Fall 1993, pp. 353-360.

An Integrated Virtual Environment System

James K. Hahn, Larry Gritz, Rudolph Darken, Joseph Geigel, and Jong Won Lee

Department of Electrical Engineering and Computer Science

School of Engineering and Applied Science

The George Washington University

Washington, DC 20052

hahn|gritz|darken|geigellwon@seas.gwu.edu

ENEWS Program

Tactical EW Division (Code 5707)

Naval Research Laboratory

Washington, DC 20375-5000

Abstract

Virtual environment research involves a number of related problems from a variety of domains. A joint research at the George Washington University and the Naval Research Laboratory is bringing together issues from these domains to study the factors that contribute to an integrated virtual environment. The research can be divided into three general categories: human factors, motion control, and sound synchronization. Human factors issues involve the development of new paradigms for movement and navigation, essential for performance of general tasks in virtual spaces. Novel approaches to motion control are being explored to help users of virtual environments interact and control virtual objects. This involves both interactive control as well as automation through evolutionary approaches. The sounds being generated as a result of these motions are modeled with compositional techniques to parameterize and synchronize them to the events in the environment. The research is being approached from both a fundamental point of view typical of an academic environment as well as from an application oriented point of view of interest to the Navy. The cooperative relationship has benefited both the George Washington University and the Naval Research Laboratory.

1. Introduction

Joint research in virtual environments (VE) is being conducted at the George Washington University's Computer Graphics and User Interface Group in conjunction with the Tactical Electronic Warfare Division (TEWD) of the Naval Research Laboratory (NRL). The research involves exploration of a number of areas of general interest to the VE community as well as applications that are of specific interest to the Navy. The research is divided into three general areas: human factors, motion control, and sound. Human factors research involves development of new paradigms for movement and navigation essential for performance of general tasks in virtual spaces. Research in motion control is concentrated on exploring various approaches that help users of VE interact and control inanimate and autonomous objects in the environment. These approaches involve interactive control as well as novel evolutionary programming paradigms. The sounds being generated as a result of these motions are modeled with compositional techniques to parameterize and synchronize them to the events in the environment. The mapping of the parameters from the motion control to the sounds produced contribute greatly to this synchronization. The emphasis is currently on solving some fundamental problems of sound in VE and less on the performance issues. The interdisciplinary nature of VE requires the application of techniques from a variety of domains that need to be integrated. We are studying each of the three areas with special emphasis on this integration problem.

The overall objective of the research is to explore these fundamental problems with an eye toward their application in a specific domain. We feel that this mixture of both pure research and applications of the research is essential in VE. The close cooperation between the George Washington University and the NRL has been very fruitful in this regard.

In section 2, we present a brief introduction to our research in navigation of VE. Motion control research is presented in section 3. Section 4 describes the synchronization of sounds to motions. Section 5 describes one of the important applications of specific interest to NRL. Section 6 concludes with a summary as well as an overall philosophy of the research. Section 7 gives a brief description of the laboratory.

2. Human Factors

At this early stage of development, we find that simple psychomotor tasks in a VE such as changing the viewpoint's position or orientation are difficult to perform at a level comparable to that of real spaces. We have developed a framework for characterizing virtual spaces and the types of psychomotor tasks typically performed in them. Spaces are classified based on size, population density, and action of virtual objects (e.g. static, dynamic). Movement tasks are generally classified as learning (browsing movements) or searching (goal-directed movements). Our initial focus has been on navigation tasks in large, sparsely populated spaces.

The creation and use of cognitive maps for navigation has long been a focus of psychological research (Stevens & Coupe, 1978; Goldberg, 1982). Although mental representations of a space are different for all individual viewers of the space, Lynch (1958) has developed a set of generic components (for urban environments) which are evident in the construction of cognitive maps:

- Paths: linear separators, examples include walkways and passages.
- Edges: linear separators, such as walls or fences.
- Landmarks: objects which are in sharp contrast to their immediate surroundings, such as a church spire.
- Nodes: sections of the environment with similar characteristics. For example, a group of streets with the same type of light posts.
- Districts: logically and physically distinct sections. In Washington, D.C., they might be Foggy Bottom, Capitol Hill, etc.

Based on these components and our understanding of how humans orient and navigate in the real world, we have built a number of tools founded on real world analogs for use with these tasks in virtual spaces. The toolset includes:

- Flying: based on avian navigation abilities.
- Spatial audio: an acoustic landmark.

- Landmarks: synthetic landmarks placed randomly in the world.
- Breadcrumb markers: markers dropped by the user in the space for uses similar to trailblazing.
- Coordinate feedback: a textual readout of the user's global position either in Cartesian coordinates or polar coordinates.
- Districting: visible lines drawn in the world to subdivide it into sections.
- Mapview: a dynamic map which moves relative to the viewpoint. It represents the user's present position and the positions of objects in the space. Its orientation can be maintained relative to the viewpoint or to the world.

We performed a preliminary study with these tools to determine if there exists a correlation between the cues or tools and the navigational behaviors exhibited by the subjects (Darken & Sibert, 1993). Our initial findings suggest that there is indeed such a correlation and that people will tend to take advantage of environmental cues in predictable ways. Variations in behaviors between different treatments were much larger than those among individual subjects. Finally, when a virtual world exhibits similarities to the real world, principles from the real world can be extended to apply to the virtual world.

From this point, we intend to extend our research efforts in order to provide guidelines as to how cues or tools can apply to virtual worlds which vary greatly in character. The nature of virtual space offers the opportunity for alternative techniques of visualization and sonification. As virtual worlds become more abstract and complex, development of new paradigms for movement and navigation will become essential for performance of general tasks in these spaces.

3. Motion Control

Motion control has been an active area of research in the computer animation domain. The basic idea is to allow the animator to specify the degrees of freedom (DOF) in the scene for the duration of the animation. The main problem is in allowing the animator sufficient control without burdening him/her with a large number of DOF many of which are constrained (e.g. prevention of interpenetration). A number of approaches have been proposed including kinematic (e.g. key-

framing), dynamic (physically-based simulation) (Hahn, 1988), behavioral (e.g. flocking motion based on simple behaviors) (Reynolds, 1987), and constrained optimization (*spacetime constraints*) (Witkin & Kass, 1988).

Controlling the motions of objects in a VE has similar associated problems. Since most of the actions are based on real-time, simulation-based approaches (like physically-based and behavioral motion control techniques) are more important in VE. We are interested in creating general motion control methodologies that will allow the user to control and interact with an environment that moves automatically. The requirement for real-time performance makes it especially difficult for physically-based modeling of rigid and deformable objects. In our earlier work (Hahn, 1988), we demonstrated that realistic physically-based motion control of rigid bodies is possible in near-real time if the problem of spatial reasoning and collision detection are resolved. We used some spatial coherence to optimize collision detection. In the same work, we also explored mixing kinematic control with simulation. This is essential in modelling the interaction of the user (essentially kinematic from the point of view of the simulated world) with the simulated environment.

We are currently investigating the possible use of evolutionary programming and traditional learning models in order to evolve / teach "coordinated group motion" as well as to apply them in a variety of novel dynamic and kinematic motion control techniques. The basic strategy is to allow the user to specify the goals as well as the evaluation metric for the motions and let the system generate the motions.

3.1. Evolutionary Programming

Evolutionary Programming (EP) is a method of solving optimization problems which uses an evolutionary metaphor. The basic procedure is as follows:

1. A "population" of data structures representing potential solutions to a particular problem is initially created at random.
2. Each "individual" in the population is rated for "fitness," or how well they solve the problem at hand.

3. Individuals are stochastically selected for the next generation in a manner proportionate to fitness. That is, individuals which are more fit (i.e. come closer to solving the problem) will have a higher chance of returning for the next generation.

4. In order to increase diversity of the new population, new individuals are created by genetic operators such as mutation (which makes a random change to an existing data structure) and crossover (which combines two individuals in a manner analogous to sexual reproduction).

5. Steps 2-4 are repeated until an individual is found which adequately solves the problem at hand.

While the theory behind evolutionary programming and its properties is not yet well understood, empirical results indicate that this metaphor is a surprisingly good method for solving certain optimization problems. EP is especially adept at solving non-intuitive, highly nonlinear problems which are not well suited to other optimization methods.

The type of data structure being evolved, the choice of genetic operators, the "fitness evaluation function," the termination criteria, and other parameters, are problem-dependent. Two well-known special cases of EP are being explored. When the data structures being evolved consist of fixed-length binary strings, the techniques are known collectively as Genetic Algorithms (GA). When the structures undergoing evolution are computer programs, they are known as Genetic Programming (GP) (Koza, 1992).

GP is especially interesting because it can be viewed as a method of program induction. That is, GP can evolve a computer program for solving a particular problem without a human explicitly programming it. The user merely needs to specify a metric for determining how well a candidate program solves the particular problem (i.e. the fitness function).

We have begun exploring the applications of EP to the problem of motion control for simulation and computer animation. Specifically, we are exploring behavioral systems and articulated figure control in dynamics simulations.

3.2. Behavioral Systems

The term "behavioral systems" (Reynolds, 1987) refers to small control programs which determine the behavior of autonomous agents in a simulation. These systems are especially

interesting in cases in which identical control programs are running in multiple autonomous agents. Surprisingly, even trivial control programs often result in complex aggregate behavior such as coordinated group motion (realistic flocking, for example). However, it remains difficult to precisely control the aggregate behavior of a group via traditional programming methodologies.

We believe we can successfully use GA and GP methodologies to develop behavioral systems for use in a variety of simulations. Applications include automatic programming of robotic simulations which have particular non-trivial behaviors such as wall-following, obstacle avoidance, coordinated group motion, coordinated group planning, and task-oriented behavior. Having multiple complex autonomous agents such as these could enhance simulations in which many "background" agents are to perform non-trivial tasks.

3.3. Articulated Figure Control

We are also exploring the use of GP for controlling articulated figures in a dynamics simulation. When simulating dynamics, it is not easy to translate a high level description of a task (for example, grasping an object) into the necessary forces and torques at the joints of an articulated figure. We are applying GP to this inverse dynamics problem. Applications involve more realistic and robust control over articulated figures for both robotics and computer animation.

Another goal for this research is to develop a method for producing expressive motion. In both robotics and computer animation, we are still largely unable to produce smooth, natural-looking, expressive motion. Even for simple tasks (for humans) such as walking, grasping of objects, and facial expression, natural-looking solutions still elude us. For example, though much has been done in simulation of walking, the research has concentrated on objective physical criteria such as minimizing the energy required to walk across the room. It is certainly not true that humans expend minimal energy when walking, yet we are remarkably efficient at it. In robotics and computer animation, not only have walking simulations resulted in motion which is not natural-looking, but in general such motion is not very robust when performed in non-trivial environments such as uneven terrain or in the presence of obstacles. We are using GP techniques to solve some of these motion control problems in ways which can handle a wide variety of situations very robustly and also produce motion which is convincing, similar to human motion, and perhaps even expressive.

4. Sound in VE

The generation of sounds, their synchronization to motion, and modeling of environmental effects are important problems in VE and computer animation. In VE, the research has concentrated mainly on Head Related Transfer Functions (HRTF) and localization of sounds in anechoic environments (Wenzel, 1992) for real-time performance. The issues of generation and parameterization of sounds as well as calculating general environmental effects have received little attention, although we get so much information about our environments from sound (Foster & Wenzel, 1991). In computer animation, soundtracks are traditionally hand synchronized using recorded sounds. Therefore the timing, as well as having the sound reflect the characteristics of the motion, is extremely difficult. This is especially a problem for continuously varying sounds, where the characteristics of the sounds should be mapped to the motion. A notable exception is in (Lytle, 1991) where a MIDI script was used to drive the motions in an animation. There are three major problems: sound synthesis, sound synchronization, and "sound rendering." The approach we are taking in our research is to consider some of these fundamental problems without the limitations of hardware performance considerations.

4.1. Sound Representation and Synthesis

There has been a lot of work in computer music for sound synthesis, including functional composition. However, since the sounds generated does not necessarily correspond to physical motions, there is generally no way to map sound parameters to motion parameters. In (Gaver, 1993), auditory icons were synthesized so that they were parameterizable along the attributes of the events responsible for the sounds in user interfaces. However, the approach lacks a general methodology to represent and map sounds to arbitrary motions.

What is needed is a general system for representation and evaluation of sounds that is parameterizable. We have been approaching the problem using a functional composition representation, called *timbre trees*, that is analogous to *shade trees* (Cook, 1984) in image synthesis. Timbre trees are completely general representations of sounds. They are composed of nodes that operate on other timbre trees, sampled sounds, special waveforms, or other parameters. Operations to be performed include standard arithmetic operators, dot product, vector normalization, etc. and the set is expandable. Evaluation of the tree produces a particular sound.

The advantage in using a tree structure is the modularity and simplicity of composing an endless variety of techniques. Since the representation is completely general, any new sound construction methodology can be represented as a timbre tree. One particular tree, with all the parameters fixed, represents an *instance* of a sound (e.g. bee sound), whereas the tree structure itself represents an entire class of sounds (e.g. insect sounds). By leaving the structure of the tree intact and just changing the parameter values, we can easily construct other instances of that class (e.g. wasps, mosquitoes, or even chain saws).

Timbre trees can be constructed using a library of elementary nodes or complete timbre trees, much like an artist learning to use and mix colors. Often, simple heuristics can be used to design an effective timbre tree that describes the process responsible for a sound. They can also be derived based on more rigorous physical principles. Most physical processes that generate sounds are difficult to simulate and an accurate simulation is usually not necessary to generate realistic sounds. However, attaching a few sampled sounds to events is not sufficient since sounds usually vary with parameters of the event. For example, many sounds can be generated by striking an object in different ways. We can approach the problem by constructing and attaching a timbre tree for a class of physical objects and excitation modes (e.g. colliding, sliding, rolling) (Hahn, 1988). Since the parameters within the trees are variable, different sounds can be produced by evaluating the tree with different arguments.

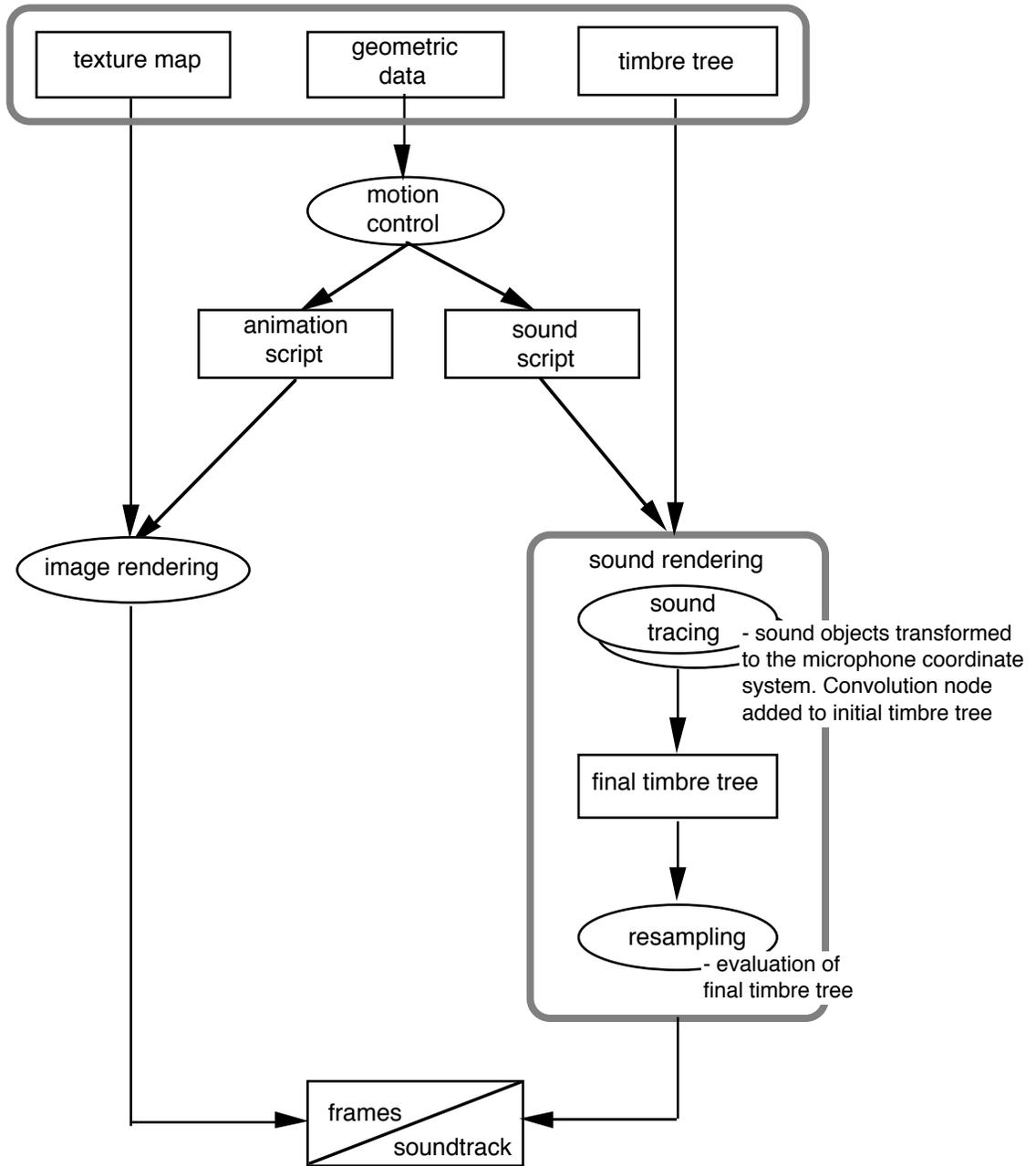
4.2. Sound Synchronization

Sound synchronized to motion provides additional cues that are helpful in defining spatial and temporal relationships. Synchronization does not only refer to timing sound events to motion events but also mapping between the parameters of the motion control system (position, orientation, location of contact between objects, force of contact, etc.) and parameters of the timbre tree. This mapping allows the generation of synchronized sounds that vary with time, based on the motions.

4.3. Sound Rendering

In our previous work (Takala & Hahn, 1992), we first introduced the concept of "sound rendering." The process of sound rendering is analogous to image rendering. Sounds can be seen as one-dimensional (in time) data objects associated with geometric objects in much the same way as texture maps can be seen as two or three dimensional (in space) data objects associated with

geometric objects. These sound objects are "rendered" in a pipelined process paralleling the image rendering process (the figure shows the general sound rendering process). This involves their instantiation based on a sound script derived from the motion control system. Once instantiated, the sounds are traced in an acoustic environment similar to the way that illumination energy is determined by ray-tracing or radiosity. This process also involves the transformation and resampling of sound from the object space to the microphone space. This is analogous to what is done to a texture while mapping it from the texture space to the image space. These environmental effects can be simulated by attaching additional nodes to the timbre tree. The resultant soundtrack is generated by evaluating the final timbre tree.



Parallel pipelines for image and sound rendering

5. Application Directions

One of the major application focuses of our VE research at NRL has been in the visualization of naval simulations for tactical and strategic evaluation as well as possible shipboard employment (Bergen, Darken, & Duckworth, 1993). In traditional system design, a machine is placed between the operator and the world. Information is gathered from the three-dimensional world and is displayed in two dimensions requiring conscious interpretation by the operator. This can be a primary bottleneck in the system. Humans are three-dimensional spatial beings and operate most effectively in three-dimensional space. Therefore the use of virtual environment technology is offered as the means by which operators will visualize information and interact with shipboard systems.

For naval applications, an immersive interface is superior to two-dimensional displays or flat screen three-dimensional displays for a number of reasons:

- The environments being visualized will encompass very large virtual spaces.
- There are a large number of parameters which will be manipulated by the operator.
- Because of the large number of parameters present, tasks will typically be of a hands-busy nature.
- Perspective is important and must be frequently altered by the operator.

The problem is not that present displays do not offer enough information but rather that they present it in a way that overloads the operator. The goal is to gather more information rather than more data. There are two fundamental obstacles associated with this problem: (1) the lack of a suitable output mechanism for presenting information to the three-dimensional world of the operator; (2) an inadequate means for accepting input from the operator by facilitating natural human psychomotor control.

Our approach is to create an interactive control and display system which presents a panoramic display of all Navy-related information organized spatially and temporally in three-dimensional virtual space. The objective is to communicate to the operator an overall awareness of the combat situation and provide control of the system state relative to the world, threats, and

targets. Situational awareness involves knowing where all the players are, what they are doing, what their intentions are, and knowing available options.

The long term goal for this application is to span field operations and laboratory studies. The role of the operator in mission planning and evaluation in the laboratory can either be observatory or participatory depending on whether or not the operator has the ability to affect the outcome of the simulation. This is also true in the real world case with the operator in the field. Observations can be made with or without having the ability to actively participate in the exercise. Furthermore, the role of the operator can take place either internal or external to the platform. Internally, a simulation of actual naval operations (traditional operator displays) will be available while externally, visualization of tactical and strategic information as well as weather information and electromagnetic spectrum information will provide similar information to the interior model within a multi-sensory three-dimensional framework.

6. Conclusion

There are two important characteristics of VE research. First, it is interdisciplinary in nature. Many components that come from different fields (computer graphics, computer animation, sound, human factors, etc.) must be made to work together coherently. The interactions between the components themselves constitute a large part of the investigation. Second, the field is essentially about the study of tools. It must be studied with an eye towards the applications but with the proper emphasis on fundamental research.

In order to address the first issue, we have chosen to approach the problem of generating an integrated VE from three domains: human factors, motion control, and sound. The human factors problem involves navigating in a virtual environment. The general approach we advocate is to look at the ways in which navigation tasks are performed in the real world. We have implemented some navigation paradigms and have made some evaluations. We will be looking at extensions to these paradigms and continue to apply them to the particular problem of navigation of sparsely populated space typical of naval simulations. Manipulating virtual objects in a VE has not been given enough attention. We have been applying motion control techniques from computer animation to control autonomous objects in a VE. The approach we have been using is evolutionary programming to evolve behaviors for single and groups of interacting objects.

Preliminary results indicate the approach is very robust and capable of adapting to changing environments. Lastly, we have been approaching the problem of integrating sound in the VE from a more fundamental point of view. The key to our approach is to view the motion and sound as an integral process. The parameters generated from the motion are mapped to the parameters associated with the corresponding timbre trees. The resultant sounds are rendered in a process paralleling image rendering.

Academic laboratories are primarily interested in pure research. Government and industrial laboratories often have applications that drive their research efforts. We feel that a balance between the two is especially critical for VE research. The cooperation between the George Washington University and NRL has been very fruitful in this regard. The research is being conducted at both locations and feed on each other. Ideas and personnel are free to migrate from one location to the other. The students who participate in the projects have the opportunity to experience both environments. We foresee this cooperation continuing to benefit both the George Washington University and the NRL.

7. Laboratory Description

There are two laboratories associated with this VE project. The Computer Graphics and User Interface Laboratory is located at the George Washington University. The VE Laboratory at NRL is located at the TEWD. Currently, there are approximately eight graduate students working on the project at both laboratories.

A partial description of the computing equipment includes: two SGI Reality Engine Onyx Workstations, two SGI VGX Power Series Workstations, SGI Server 4D/420S, and a number of SGI Indigo and Crimson workstations. Peripherals include: Fake Space Labs BOOM3C, LEEP Systems Cyberface2 Head Mounted Display, Ascension Flock of Birds Tracker with Extended Range Transmitter, Virtual Technologies Cyberglove, BioControl Biofeedback Device, and Visual Synthesis Audio Cube.

8. Acknowledgements

We would like to thank the Naval Research Laboratory (NAVY N00014-91-K-203) and NASA Goddard Space Flight Center (NCC 5-43) for their support. We would also like to extend warm appreciations to the students in the Computer Graphics and User Interface Group at the George Washington University. Daria Bergen, Jean Favre, Sanjeev Sharma, and Suneil Mishra have been especially helpful. We thank the reviewers for their encouragement and many useful suggestions.

9. References

- Bergen, D. E., Darken, R., & Duckworth, A.N. (1993). Applying Virtual Environment Technology to Large-Scale Simulation Systems, (Tech. Rep. NRL/FR/5707-93-9560), Washington, D.C.: U.S. Naval Research Laboratory.
- Darken, R., & Sibert, J.L. (1993). A Toolset for Navigation in Virtual Environments, *Proc. ACM UIST '93*, 157-165.
- Cook, R. (1994). Shade Trees, *Proc. SIGGRAPH'84, ACM Computer Graphics*, 18(3), 195-206.
- Foster, S & Wenzel, E. (1991). Virtual Acoustic Environments: The Convolvotron, [multimedia installation], *SIGGRAPH'91 Virtual reality and hypermedia show*.
- Gaver, W. (1993). Synthesizing Auditory Icons, *Proceedings of INTERCHI*.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.
- Goldin, S.E., & Thorndyke, P.W. (1982). Simulating Navigation for Spatial Knowledge Acquisition, *Human Factors*, 24(4), 457-471.
- Gritz, L, Bergen, D., & Darken, R. (1992). Graphic Violence, (Computer Animation) *ACM SIGGRAPH Video Review Issue 85*.
- The George Washington University (1992). Sound Rendering, (Computer Animation) *SIGGRAPH'92 Film and video show*.

- The George Washington University (1993). Timbre Trees, (Computer Animation) *Video Screening Room SIGGRAPH'93*.
- Hahn, J. (1987). Rigid Body Dynamics Simulations, *ACM SIGGRAPH Video Review #28, Visualization in Scientific Computing, Computer Graphics*, Vol. 21, No. 6.
- Hahn, J. (1988). Realistic Animation of Rigid Bodies, *Proc. SIGGRAPH'88, ACM Computer Graphics*, 22(3), 299-308.
- Koza, J. (1992). *Genetic Programming*, MIT Press, Cambridge, MA.
- Lynch, K., & Rodwin, L. (1958). A Theory of Urban Form, *Journal of the American Institute of Planners*, 24, 201-214.
- Lytle, W. (1991). More Bells and Whistles, (video) in *SIGGRAPH'90 Film Show*. Also described in *Computer*, 24(7), p.4 and cover.
- Reynolds, C.W. (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model, *Proc. SIGGRAPH'87, ACM Computer Graphics*, 21(4), 25-34.
- Stevens, A. & Coupe, P. (1978). Distortions in Judged Spatial Relations, *Cognitive Psychology*, 10, 422-437.
- Takala, T., & Hahn, J. (1992). Sound Rendering, *Proc. SIGGRAPH'92, ACM Computer Graphics*, 26(2), 211-220.
- Wenzel, E.M. (1992). Localization in Virtual Acoustic Displays, *Presence: Teleoperators and Virtual Environments*, 1, 80-107.
- Witkin, A. & Kass, M. (1988). Spacetime Constraints, *Proc. SIGGRAPH'88, ACM Computer Graphics*, 22(2), 159-168.