SYMPOSIUM: SUPERCOMPUTER APPLICATIONS IN PSYCHOLOGY

Discussants: Lynne K. Edwards, *University of Minnesota Supercomputer Institute*
Cynthia H. Null, *NASA-Ames Research Center*
Stephen W. Link, *McMaster University*

High-power graphic computers for visual simulation: A real-time-rendering revolution

MARY K. KAISER

NASA-Ames Research Center, Moffett Field, California

Advances in high-end graphics computers in the past decade have made it possible to render visual scenes of incredible complexity and realism in real time. These new capabilities make it possible to manipulate and investigate the interactions of observers with their visual world in ways once only dreamed of. This paper reviews how these developments have affected two preexisting domains of behavioral research (flight simulation and motion perception) and have created a new domain (virtual environment research) which provides tools and challenges for the perceptual psychologist. Finally, the current limitations of these technologies are considered, with an eye toward how perceptual psychologists might shape future developments.

The past decade has witnessed, if not the birth, the infancy and childhood of the field of three-dimensional computer graphics. A milestone marking the field's development is the recent release of the first full-length, completely computer-generated feature film (Disney Studios' "Toy Story"). Equally remarkable are the advances made in real-time computer-generated imagery (CGI). Ten years ago, only a handful of high-fidelity visual simulation systems existed in the world. These systems were honed for the small niche market of flight simulation. The cost of acquiring and maintaining such systems was high; the accessibility of these facilities to the research community was minimal. However, as in all areas of digital computing and storage, the improvement in cost/performance over the past 10-12 years has been remarkable. Table 1 provides a sample of high-performance graphics systems released (or scheduled for release) to the open market from 1983 to 1996.

In this paper, I will examine the impact of these technological advances on two domains of human performance

research: flight simulation and visual motion perception. I will also briefly consider how these developments have engendered a new domain of application, virtual environments (VE). While VE can, in some regards, be viewed as an extension of vehicle simulation technology, new issues and areas of emphasis have emerged that are specific to the domain. VE can thus be viewed both as providing a new set of tools for the study of perceptual processes and as a domain in which the results of such studies can guide and inform technological developments.

Simulating Flight: Creating Visual Worlds

Edwin Link filed his first patent for a flight trainer in 1929. The prototype was jerry-rigged from surplus parts in the basement of the Link Piano and Organ Company in Binghamton, New York. Although early versions were seen as novelties rather than serious training devices (the November 1930 issue of *Science and Invention* suggested that "such devices would make a valuable adjunct to the multitude of miniature golf courses that now dot the country"), Link's company soon developed realistic simulators that became the backbone of U.S. Navy Army Air Corps training. Thousands of Link's "blue boxes" (Figure 1) were used to teach basic instrument flight to World War II pilots and aviators. The key phrase to note, however, is "instrument flight"; no attempt was made, in these early flight simulators, to emulate the external scene (or "contact display," in pilot parlance).

The author would like to thank Steve Ellis and Walter Johnson, for their comments on an earlier draft of this article, and Robin Petrus of Hughes Training, for providing Link archival photos. This work was funded by NASA RTOP 199-06-12. Portions of this paper were presented at the 25th annual SCiP meeting. Correspondence should be addressed to M. K. Kaiser, Mail Stop 262-2, NASA-Ames Research Center, Moffett Field, CA 94035-1000 (e-mail: moose@eos.arc.nasa.gov).

Table 1
Cost, Size, and Performance Parameters for High Powered Graphics Systems

System	Release Date	Cost (at Time of Release)	Number of Textured Polygons per Frame (60-Hz Update Rate)	Physical Size
Singer-Link Image II	1983	> \$2,000,000	256 per channel* (3 channels)	4 rack-mount cabinets
Silicon Graphics Onyx RE2	1992	~ \$400,000	7,500 per pipeline† (2 pipeline)	1 rack-mount cabinet
Silicon Graphics Indigo Max. Impact	1995	~ \$50,000	~ 3,000 ² (single channel)	large desktop (or desk side)
PC Platform w/3-D Graphics Boards	1996	< \$15,000	> 1,000 ² (single channel)	desktop

*Polygons with up to 15 sides. †Three-sided polygons.

Link and his colleagues were well aware that external scene simulation was desirable for flight trainers, and they initiated the development of a series of increasingly more sophisticated visual world simulations. The earliest attempt, the cyclorama, is shown in Figure 2. Basically a static backdrop, it nonetheless provided the pilots with useful attitude (pitch, yaw, roll) information and feedback to their control inputs. Of course, the information that pilots typically glean from translational flow was missing. In response to requests from the military, the Link Company developed the Celestial Navigation Trainer shown in Figure 3. The primary function of this silo-sized structure was to teach flight crews to navigate by the stars. Thus, the lack of translational flow did not adversely impact the simulation fidelity.

The next real advancement in visual scene simulation utilized model board technology. In these systems, physical models of the gaming area are built on huge boards. A scale factor of 3,000:1 was used for the Mark IV Terrain model shown in Figure 4, meaning that a 20 × 20 ft board was needed to encompass a land mass of 100 square nautical miles. A video camera would then “fly” over the board; the image would be conveyed on a monitor (or projection screen) to the pilot in the simulator. The pilot could control

the path of the camera in real time, although the movement was limited by the dynamics of the camera’s mounting and drive system. The range of motion was likewise limited; mounts had a maximum height, and any errors that resulted in ground strikes resulted in damage to the model board and angry technicians. Simulation fidelity was also constrained by limitations in the performance of the video system. Changes in the visual scene required painstaking work by the modelbuilders. Finally, these systems were physically cumbersome. Although the boards were oriented vertically (to minimize space requirements and to simplify the camera tracking mechanisms), large high-bay areas were usually required to mount these systems.

In the late-1960s, Link introduced a visual scene generator that utilized motion picture technology. The Variable Anamorphic Motion Picture (VAMP) system utilized full-color, 70-mm films of actual, prerecorded flight sequences. The sequences were filmed during nominal approach, landing, and taxi operations. In the simulator, deviations from these nominal paths were possible within a limited window (see Figure 5). The visual scene presented to the pilot was derived from the film sequence by using two processing techniques: (1) subsampling the region of the 70-mm frame to correspond to the simulated aircraft’s pitch and heading; and (2) optically warping the image with anamorphic lenses to simulate vehicle roll.

In many regards, the VAMP technology represented a major breakthrough for visual scene generation. The level of fidelity was unprecedented. However, the logistics of gathering film sequences was daunting: the initial O’Hare set required accommodating the project’s helicopter into a very busy air traffic schedule; the stress of properly mimicking glide-slope and taxi heights of various aircraft (and docking the helicopter on a special platform designed to accommodate the underbelly camera pod) was sufficient to cause one pilot to excuse himself from the project because of “nervous exhaustion” (Kelly, 1970). Also, there were technical drawbacks to the simulation itself. Repeatedly viewing a fixed number of sequences, pilots quickly learned to use extraneous cues. For example, one cited knowing he was on proper glide-slope at proper speed if he crossed a landmark (the outer marker) at the same time a semi-truck crossed a bridge on a nearby roadway. Also, despite the high-quality optics, some distortion occurred if one deviated too far from the nomi-

Figure 1. The Link AN-T-18 Instrument Trainer was used in the training of over half a million pilots in the WWII era. (Photo courtesy of the Link Division of Hughes Training, Inc.)

Figure 2. The Link 1-CA-2 flight trainer used a static backdrop (the "cyclorama") as a simple external scene "generator." (Copyright Robert B. Parke. Reprinted with permission.)

nal path. Thus, pilots learned to fly so as to null optical distortion, a strategy that does not transfer well to the actual task of flying an aircraft.

The limitations and high overhead associated with these non-CGI systems led to the simulator community's becoming highly receptive to computer-based solutions for visual scene generation. In fact, the first-generation CGI visual systems for simulators were of lower fidelity than these alternative methods. But the user community was highly impressed with the CGI systems' flexibility and promise. This early faith has been amply rewarded.

Motion Perception Research: Controlling Visual Worlds

Like flight simulation, the domain of motion perception existed before the advent of CGI systems. Although motion perception is one of the more recent topics of studies by perceptual psychologists, empirical research in the area has been actively pursued since the turn of the century (following a number of intriguing demonstrations in the 19th century); the growth of the field has been quite prodigious since the invigorating work of J. J. Gibson in the 1950s and the subsequent computational modeling efforts. And like the scientists and engineers who developed flight simulators, researchers in this domain required a good deal of ingenuity (and tenacity) to find suitable means to generate visual stimuli for their observers.

While researchers in flight simulation strive for visual fidelity, the primary concern in the domain of motion perception is experimental control. It is critical to be able to fully specify the nature of the visual stimulus (at the relevant level of analysis) in order to examine the systematic relationship between phenomenon and percept.¹ Prior to the use of real-time CGI methods, researchers depended on a number of techniques to generate stimuli. These included mechanical apparatus that generated desired motion patterns; "motion strips" that, when viewed through apertures, created motion sequence (see Michotte [1963]

for a description of this technique); sequential illumination of lights that, through properties of long- or short-range apparent motion, created motion sequences; and sampled motion techniques, including film, flip-books, and multichannel T-scopes. Two examples of these methods will be considered.

To facilitate his experiments on the perception of wheel-generated motion, Proffitt and his instrument engineer (Proffitt & Simmons, 1980) designed and built the apparatus shown in Figure 6. Observers were able to move targets on the disk by pressing buttons, which were switches to the electric motors that moved the target mounts along the initial x - and y -axes while the disk rotated. Observers viewed the apparatus on a closed-circuit television system, whose lighting and contrast settings were controlled in such a way that only the targets (and none of the mechanisms) were visible. Thus, the system permitted real-time method-of-adjustment experiments with the motion stimuli. This apparatus is just one exemplar of creative, innovative mechanical devices devised by researchers (and their most clever lab technicians) to study motion phenomena.

The apparatus shown in Figure 7 is a fascinating hybrid of sampled-motion techniques and the computer-generated methodologies yet to come. The photo is from Braunstein's (1976) book *Depth Perception Through Motion*. At that time, the PDP-11 computer in Braunstein's lab was capable of calculating 2-D projections of a 3-D

Figure 3. The Link celestial navigation trainer. (Drawing courtesy of the Link Division of Hughes Training, Inc.)

Figure 4. Model builders put finishing touches on the terrain board of a Link Mark IV system. (Photo courtesy of NASA.)

point-light volume as it underwent rotations and displayed the resulting points on a Tektronix 4010 graphics terminal. However, the program required 5 sec or more in order to generate each frame. To create motion sequences, Braunstein used a motion picture camera to capture each frame (the computer emitted a tone which tripped a sound-actuated relay unit that pulsed the camera's motor drive). The film segments then served as experimental stimuli. Thus, Braunstein was able to overcome the computational and storage limitations of his system and create computer-generated 3-D motion stimuli for psychophysical research. Innovative methods such as his paved the way for more general use of CGI stimuli as graphics systems became more capable and affordable.

At the 1985 Symposium on the Use of Supercomputers in Psychology, the potential for CGI for perception research was considered (although the levels of performance and the affordability that would be achieved in the following decade were not foreseen). Similarly, the promise of CGI for visual scene generation in simulators was recognized in the early 1970s. The IMAGE Society has held an annual conference and exhibition devoted entirely to visual simulation technology since 1977. However, neither forum adequately anticipated the emergence of a new field engendered by the advances in graphical computer technology: the domain of VE.

Virtual Environments: Immersion in Visual Worlds

Generally, the field of VE (also known as virtual reality, telepresence, cyberspace, and a number of other ill-defined buzzwords) is regarded as beginning in the mid-1980s, building on the work at NASA-Ames Research Center, MIT Media Lab, a number of small companies in California's Silicon Valley, and several military laboratories. But more than a few simulator researchers view the field as simply a natural descendent of their domain; indeed, a large impetus for the development of helmet-mounted display systems was the desire to find an eco-

nomical replacement for large dome projection systems (Barrette et al., 1990). The distinction between VE systems and simulator systems can be subtle (Durlach & Mavor, 1995), but generally, VE systems create new and often more stringent demands on visual simulation hardware.

First, since a VE system strives to create an immersive environment, displays tend to be head-mounted, requiring that systems be light-weight, high-resolution, and binocular. Second, because the VE system simulates changes in the visual scene that correspond to body motions (rather than changes in vehicle position resulting from control inputs), lags in VE systems are more noticeable (and annoying) because the operator cannot attribute the delays to natural vehicle dynamics. Finally, cues from other modalities (auditory, tactile, even olfactory) which are effective aids in vehicle simulation are often absent, diminished, or distorted in VE systems. Thus, more dependence is placed on visual simulation fidelity in order to achieve realism.

Of course, this discussion assumes that the VE system is being used as a personal simulator, to create a highly veridical sensory emulation of the simulated environment. In fact, VE can be characterized best as a communication medium that can be used not only for environmental simulation, but for more abstract, convention-defined forms of communications: a multidimensional extension of the desktop interface metaphor, so to speak (Ellis, 1995). In such cases, the requirements for visual scene generation will not be defined by notions of fidelity, but rather by task- and context-specific requirements of the user. Still, just as the birth of VE was due, to a large extent, to the emergence of capable, affordable, 3-D computer graphics systems, the growth of the domain seems intimately tied to the improvements in the cost/performance curve of these systems. As is the case with vehicle simulation and motion perception research, VE development profits greatly from advances in computer graphics architecture.

I now consider how perceptual psychologists working in these three application domains can, in turn, provide

Figure 5. "Translational Freedom" envelope for the Singer-Link VAMP system. The envelope is scaled by the height (or above ground altitude) of the aircraft, h . So long as the aircraft remained within this region (and on proper heading $\pm 20^\circ$), the VAMP could generate an appropriate visual scene.

Figure 6. The revolving X-Y plotter. From "A Revolving X-Y Plotter for Studying the Perception of Rotational Motions," by D. R. Proffitt and R. W. Simmons, 1980, *Behavior Research Methods & Instrumentation*, 12, p. 602. Copyright 1980 by Psychonomic Society, Inc. Reprinted with permission.

insight into the evaluation, enhancement, and future development of computer graphics systems.

Human Factors of Visual Simulation: Perceptually Tuning Visual Worlds

Ten years ago, Proffitt and Kaiser (1986) considered the advantages and disadvantages of using computer graphics animation in motion perception research. Many of the advantages, and most of the concerns, apply to vehicle simulation and VE systems. Interestingly, although price and performance are much better than they were a decade ago, the factors that Proffitt and Kaiser then cited are still germane. The advantages of these systems, discussed above, can be summarized in terms of two primary factors: flexibility and control.

The disadvantages, however, also need to be considered. Proffitt and Kaiser (1986) listed six principal limitations: (1) dual awareness—the observer's awareness of both a transforming 2-D pattern on the display screen and the 3-D event being simulated; (2) scaling perspective, depth, and size—most display situations create an ambiguity (or indeterminacy) in these parameters. [These first two limitations can be grouped into a single issue of conflicting/absent depth cues.] (3) limited field of view/resolution—computationally, these are offsetting parameters, and even the best graphics systems still cannot generate a full visual field at eye-limited resolution; (4) motion quality—achieving an acceptable update rate trades off against field of view and resolution, but even at an

acceptable frame rate there are artifacts associated with using sampled motion (Watson, Ahumada, & Farrell, 1986); (5) object realism—despite tremendous advances in this domain in the past decade, object complexity is still a burden on computational resources; (6) adequacy of simulated dynamics—this is more often a limitation of software than hardware; adequate mathematical motion models are available for only a subclass of events users wish to simulate (however, the size and breadth of this catalog has increased enormously in the past decade).

Systems have improved on almost all of these factors in the past 10 years, largely because of increases in raw computational power. But psychophysical insight has led to

Figure 7. Braunstein's computer animation laboratory, using a motion picture camera to capture frames rendered on a graphic display terminal. From *Depth Perception Through Motion* (p. 179), by M. L. Braunstein, 1976, New York: Academic Press. Copyright 1976 by Academic Press, Inc. Reprinted with permission.

Figure 8. A research program demonstrating where perceptual psychologists can impact graphics systems design.

several advances, notably motion-blur algorithms, reduction of critical cue conflicts, and variable resolution/ level of detail rendering. Our lab is currently collaborating with other psychologists and computer scientists to develop what we term "perceptually tuned" rendering techniques: methods to reduce the computational complexity of objects with little or no impact on perceived object realism.

Our program of research, as depicted in Figure 8, can be characterized as spanning a spectrum from reactive (i.e., primarily validating and evaluating existing rendering techniques) to proactive (i.e., participating in the design of hardware/software architecture to optimize for properties of human vision). Even evaluation of existing methods can lead to new insights, such as the realization that representing distant objects as a 2-D "billboard" kept normal to the line of sight works for a larger class of objects than database designers had supposed (Kaiser, Montegut, & Proffitt, 1995). However, even more significant contributions can be made when improved methods (such as continuous level-of-detail modulation: Montegut, 1995) or new methods (such as mixed-resolution stereo displays: Proffitt & Kaiser, 1995) are pursued. Finally, if an understanding of the human observer guides the architectural design of the system, significant functional improvements can be gleaned from current hardware capabilities. This is the case with the Address Recalculation Pipeline for VR Displays, which separates geometric transformations that must be performed to accommodate observer head rotation from those resulting from body translation (Regan & Pose, 1994).

Clearly, then, perceptual psychologists can play two roles with regard to high-performance graphics computers. First, as informed end users, we enjoy the capability and flexibility these systems provide as we pursue our research on the underlying mechanisms of visual perception and human performance. Second, we have the opportunity to serve as domain experts and potential collaborators for graphic system engineers who seek to maximize the performance of their rendering engines by optimizing them for the perceptual capabilities and limitations of the operators. Enriched by such collaborations, the advances in visual simulation during the next decade could prove to be even more amazing than the last.

REFERENCES

- BARRETTE, R., DUNKLEY, R., KRUK, R., KURTZ, D., MARSHALL, S., WILLIAMS, T., WEISSMAN, P., & ANTOS, S. (1990). *Flight simulation advanced wide FOV helmet mounted infinity display* (Tech. Rep. No. AFHRL-TR-89-36). Dayton, OH: Wright-Patterson Air Force Base, Air Force Human Resources Laboratory.
- BRAUNSTEIN, M. L. (1976). *Depth perception through motion*. New York: Academic Press.
- DURLACH, N. I., & MAVOR, A. S. (EDS.) (1995). *Virtual reality: Scientific and technological challenges*. Washington, DC: National Academy Press.
- ELLIS, S. R. (1995). Origins and elements of virtual environments. In W. Barfield & T. A. Furness (Eds.), *Virtual environments and advanced interface design* (pp. 14-57). New York: Oxford University Press.
- KAISER, M. K., MONTEGUT, M. J., & PROFFITT, D. R. (1995). Rotational and translational components of motion parallax: Observers' sensitivity and implications for 3-D computer graphics. *Journal of Experimental Psychology: Applied*, 1, 320-330.
- KELLY, L. L. (1970). *The pilot maker* (As told to Robert B. Parke). New York: Grosset & Dunlap.
- MICHOTTE, A. (1963). *The perception of causality*. New York: Basic Books.
- MONTEGUT, M. J. (1995). *The emergence of visual detail*. Unpublished doctoral dissertation, University of California, Santa Cruz.
- PROFFITT, D. R., & KAISER, M. K. (1986). The use of computer graphics animation in motion perception research. *Behavior Research Methods, Instruments, & Computers*, 18, 487-492.
- PROFFITT, D. R., & KAISER, M. K. (1995). *Spatial-temporal resolution process for computationally efficient displays*. U.S. Patent Application (NASA Case No. ARC 12080-1).
- PROFFITT, D. R., & SIMMONS, R. W. (1980). A revolving X-Y plotter for studying the perception of rotational motions. *Behavior Research Methods & Instrumentation*, 12, 601-604.
- REGAN, M., & POSE, R. (1994). Priority rendering with a virtual reality address recalculation pipeline. In *Computer Graphics* (Annual Conference Series 1994: SIGGRAPH94, pp. 155-161). ACM Press.
- WATSON, A. B., AHUMADA, A., JR., & FARRELL, J. (1986). Window of visibility: Psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, 3, 300-307.

NOTE

1. In fact, an ongoing debate pits the need for well-described (and controlled) visual stimuli against the need for stimuli to properly convey the complexity of natural motion phenomena. The latter concern, voiced primarily by researchers from the ecological school, often leads to the retention of non-CGI methods, such as filming natural events for observers to view.

(Manuscript received December 29, 1995;
accepted for publication January 5, 1996.)