More color, increased model complexity, and the continuing quest for photorealism will challenge computer graphics during the next decade. What changes does this imply?

More Accurate Simulations at Faster Rates

Donald P. Greenberg
Cornell University

When asked by *IEEE Computer Graphics and Applications'* renowned editor to write a brief piece about what has occurred in computer graphics during the past two decades and what will occur during the next one, I thought it would be an easy task. I immediately accepted, flattered by the invitation. After all, hindsight is always 100 percent, and the older one gets, the easier it is to recall events of long ago. The days of IBM punch cards, paper tapes, calligraphic displays, hidden line algorithms, 32-Kbyte memories, 1.2-Mbyte disks, and 24-hour-a-day operating hours (because the multiuser systems served an entire laboratory) are fondly and vividly etched in my mind. At early Siggraph conventions, one knew everybody and could get through the vendor exhibits in one hour. Also, because so few people were involved, anything could be accepted for publication. It's somewhat embarrassing now to read some of the early articles.

Initially, color was frivolity. None of the major computer manufacturers showed any interest in color and raster displays. The cost of memory to store a single very high-resolution image was exorbitant. Our first eight-bit, 512 × 480 display cost more than $80,000 in 1975. Its existence opened up a new world to us. But why bother? Of what use are pretty pictures? No scientists or engineers need color! I have quoted perhaps too many

Illustrations accompanying this article

The two sequences of images accompanying this article indicate the progression of realism, in terms of both the sophistication of light reflection and model complexity. Many of the illustrations have already been overpublicized. Nonetheless, the two sets show not only the chronology but also the fact that, despite the exponential increase in processing power, the demand for graphical computations has increased at an even faster rate.

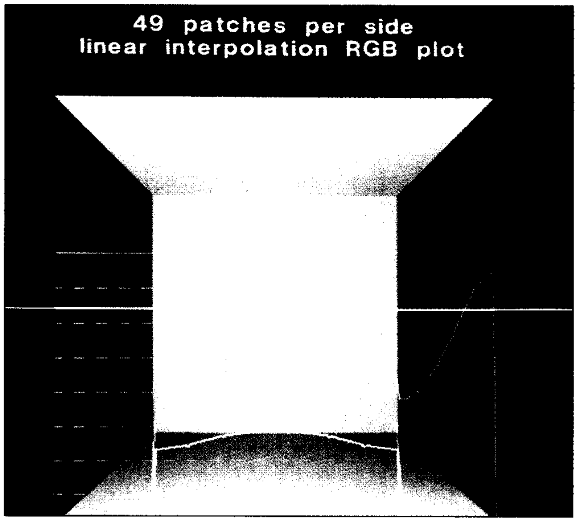
Almost all of the models and images were computed on Cornell's modeling and rendering testbed. The caption for each figure includes an estimate of the approximate times needed for the computation. Note that the model complexity has increased three to four orders of magnitude, while the available processing power at the laboratory increased by two to three orders of magnitude. However, the total time required for each image in the sequence has not been reduced.

This data supports the general observation made by Frank Crow at the Utah Conference on Interactive Graphics (March 1990) that the graphical processing demand exceeds the supply. I might also add that it perhaps implies the threshold of human impatience, or what we might call the "constant time phenomena." Once conditioned, we are willing to wait several hours, or even overnight, for the photorealistic picture to be computed.

The first six images (RGB Cube, Magritte Studio, Simulated Steel Mill, Constructivist Museum, Theory Center Proposed Lobby, and Printing Press) demonstrate radiosity. The last six images (Art Museum, Champagne Glass, Siggraph Watch, Entrance Hallway, Kitchen Counter, and Stair Tower) illustrate direct illumination and ray tracing.

times Faraday's response when asked of what use was his dynamo: "Of what use is a baby?"

Rather than reiterate the happenings of days gone by, it is more fun and interesting to conjecture about the future. (I haven't been kicked out to pasture yet.) I base my comments on the fact that—despite enormous publicity to the contrary—today's graphics simulations do not, in general, produce the same information that any young child observes in his or her environment.



"RGB Cube" by Cindy Goral (Program of Computer Graphics, Cornell University, 1984) shows a simulated cube with walls subdivided into smaller patches and rendered with linear interpolation. This simple example was one of the first radiosity images and compared favorably with the physical model. The "color bleeding" caused by the interreflections of the diffuse surfaces is illustrated by the scan-line plots of intensity for each of the red, green, and blue components. The image was computed on a VAX 11/780 and rendered on a Grinnell frame buffer at 512 × 480 resolution.

For starters, the world is not flat (although 90 percent of workstations sold are 2D). Our 3D world consists of many smooth and arbitrarily curved shapes, not just polygons. We see a full range of colors. Most real environments are dynamic, not static. All surfaces have textures. Shadows and shading help in understanding position and geometry. We are totally surrounded by space, and our peripheral vision helps us understand it. Do our simulations provide the same perception?

Commercial offerings, both hardware and software, still fall far short of providing a realistic simulation of the real world. The objective of trying to attain this "artificial reality" (today's buzzword) colors the following thoughts about what trends I think will occur in the next decade.

Increase in model complexity

I anticipate a vast increase in model complexity. With the availability of interactive modeling systems and graphical editing tools, we will move from the design of elements to components to assemblies to systems. The size of typical modeled environments, when combined with automatic meshing and adaptive subdivision techniques, will increase by two to three orders of magnitude in geometric complexity alone. This data explosion will require all available perceptual cues, including motion, just for user comprehension (see below).

"Magritte Studio" by Michael Cohen (Program of Computer Graphics, Cornell University, 1985). The introduction of the hemicube algorithm allowed simulation of environments containing occluded surfaces. The radiosity solution for this environment required approximately four hours for computation and 20 minutes to render in software on a VAX 11/780.

Continued quest for photorealism

With the continued increase in processing power, global illumination algorithms will become the standard. Users will no longer be satisfied with the results of direct lighting models only. Of course, we will obviously want better pictures that more closely represent reality. However the demand and use will not be just for realism. As the environments become more and more complex, the perceptual cues provided by shading, shadows, texture, and interreflections all contribute to our understanding of a scene. They provide the tools for disambiguation of incomplete information.

I hope that the standards of the future change so that our metric of excellence is no longer transformations per second or polygons per frame. Instead, we will evaluate the quality of the simulation in terms of how closely it represents the true environment and how long it takes to render an image that is perceptually indistinct.

Progressive rendering algorithms

Computation times for photorealistic images are excessive. Despite enormous advances in hardware performance, the demand for processing power to compute these images will always outpace the supply. With modeling complexity increasing at a faster rate than machine performance and the required computational times increasing exponentially, will the situation get worse?

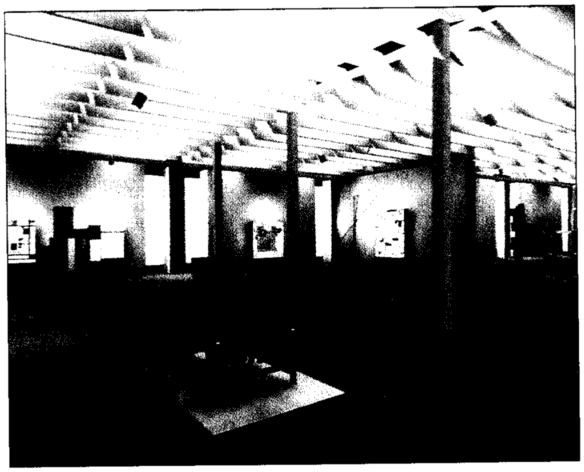
"Simulated Steel Mill" by Stuart Feldman and John Wallace (Program of Computer Graphics, Cornell University, 1987) was created using a modified version of the hemicube radiosity algorithm, computed on a VAX 8700, and displayed on a Hewlett-Packard Renaissance Display. The environment consists of approximately 2,000 patches and 55,000 elements and was one of the more complex environments computed to date.

The answer lies with progressive solutions, which will start with simple models such as the Phong direct lighting model but migrate smoothly and continuously to more refined solutions. Thus, initially people will use the fastest lighting models, since speed and interaction have the highest priority. But at each step, with all available and normally unused cycles, the quality of the image will improve.

View-independent global illumination algorithms, not constrained to static environments or simple reflection functions, will be used. Because of their progressive nature, high-quality rendered images will be generated at interactive speeds. Thus, no longer will modeling and rendering occupy separate stages in a graphics pipeline. In addition to providing better perceptual cues for interaction, the rendering-while-modeling paradigm will allow contextual approaches to design.

Progressive modeling systems

Progressive approaches do not apply to rendering alone, but also to modeling. Easy to use, comprehensive modeling systems remain elusive and difficult to devise. In fact, the modeling problem provides many more difficulties than the rendering problem. People have frequently asked me, "How long did it take to render the picture?" Rarely do questions relating to the duration of the modeling tasks arise.



"Constructivist Museum" by Shenchang Eric Chen, Stuart I. Feldman, and Julie M. O'Brien (Program of Computer Graphics, Cornell University, 1988) includes more than 20,000 elements and 200 discrete area light sources. Computation times were accelerated using the progressive refinement approach, which provides a useful solution almost immediately and progresses gracefully and continuously to the complete radiosity image shown. The complexity of the environment as well as the indirect diffuse skylighting illumination from above would have made it impractical to render with traditional radiosity methods and impossible to simulate using standard ray tracing methods. The computations were performed on Hewlett-Packard 825 SRX workstations and displayed at a resolution of 1,280 × 1,024.

Like rendering, modeling systems will have a physical basis, as geometric descriptions alone do not suffice. The laws of physics, preventing one object from interpenetrating through another, maintaining frictional or gravitational contacts, or obeying the Newtonian laws of motion, should all be part of a modeling system. These constraints must be built into the system. In this way, the user's tasks will become far easier to perform and far more efficient.

Maybe we should also think about modeling space, "voids" but not solids, subtractive rather than additive procedures. This is certainly the way in which we experience our environment.

Perhaps even more important, methods will be derived for displaying incomplete or imprecise information, approximate data. Take, for example, an architectural design. The concepts start with doodles, sketches on the back of an envelope or yellow tracing paper. Dimensions do not exist. As the scheme progresses from preliminary design through design development to working drawings, the geometrical definitions increase in detail and sophistication, from coarse to fine precision.

Progressive modeling systems—ones that can continuously refine themselves through varying phases of abstractions—will evolve.

"Theory Center Proposed Lobby" by Keith Howie and Paul Wanuga (Program of Computer Graphics, Cornell University, 1989) is part of a much larger model of the proposed Engineering Theory Center building. After the progressive radiosity solution was converged, a video walkthrough was created. The solution time for the environment (consisting of 30,000 patches) took approximately five days on a cluster of parallel Hewlett-Packard 835 workstations. Each frame of the video sequence was displayed in two to three seconds.

Physically based light reflection models

Local reflection behavior from a surface with known material properties can be predicted from its bidirectional reflectance distribution function (BRDF). This function predicts the outgoing intensity in any direction as a function of the incoming energy in a specific direction. Thus, for a known finite light source of given shape and emission characteristics, with its light striking a surface at a given angle of incidence, we can describe the outgoing intensities as a function of the wavelength of the light and the absorptive qualities and roughness characteristics of the surface. Correct BRDFs can predict all reflectance types, including diffuse and directional diffuse, as well as specular distributions.

At present, most photorealistic imaging methods fall into one of two categories, ray tracing or radiosity. Ray tracing, although capable of producing images of startling quality, assumes perfect mirror reflections and is not physically based. Therefore, it does not account for diffuse interreflections, and ad hoc procedures must be used to attenuate energy with distance. On the other hand, radiosity, which is thermodynamically correct, is computationally intractable for all but Lambertian environments. Methods for combining both techniques have been tried, but are arbitrary and not energy consistent. Lastly, both approaches produce aliasing artifacts due to the sampling methods used. New comprehensive energy-consistent procedures will emerge to overcome these deficiencies.

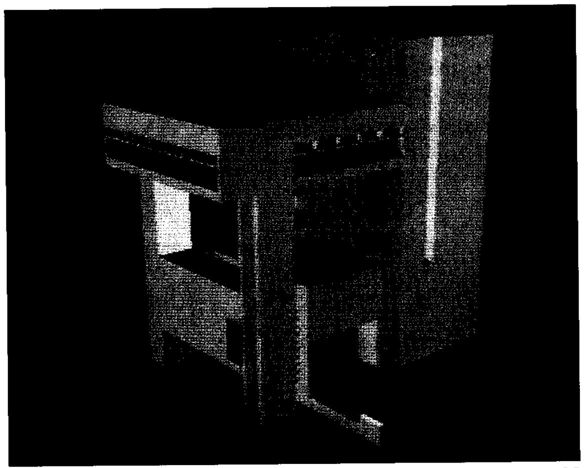
"Printing Press" by Keith Howie, Paul Boudreau, John Wallace, and Eric Haines (3D/Eye, Ithaca, N.Y., 1990) was modeled using Hewlett-Packard's mechanical engineering software (ME30). It consists of 320,000 unmeshed polygons before subdivision. To create this image required using 19 steps of a progressive radiosity solution at 45 minutes per step on a Hewlett-Packard 375 workstation.

These future methods will rely on the BRDFs. Knowledge of the true reflection behavior will allow rational assumptions for the selection of appropriate models for diffuse or specular environments, or combinations of both, and will also enable the rational selection of algorithms dependent upon speed or quality. But all of these approaches will have a physical basis. Furthermore, libraries of BRDFs or, even better, models simulating the BRDFs, will be freely available and distributed by manufacturers on optical disks, so that users only need access the database by name.

Elimination of display lists

The idea of structured and hierarchical display lists came about for several reasons. With limited processing power available, it was necessary to offload the CPU so that its precious computational cycles weren't wasted on display computations.

Furthermore, since processors were expensive, their high cost was justified through multiuser operating systems; no single user should be allowed to gain complete control of the precious resource. Lastly, to allow the display of data in real time at remote sites, it was necessary to place intelligence in the display head. In this way, following construction and transmission of the display list, only a small amount of information needed to be sent over the long-distance, cheap, and limited bandwidth networks, and the traversal and picture generation operations could occur locally.



"Art Museum" by Marc Levoy, Robert Hastings, and Donald P. Greenberg (Program of Computer Graphics, Cornell University, 1971). This direct illumination simulation of Cornell's Johnson Art Museum was rendered and displayed at General Electric's Visual Simulation Laboratory in 1971. We used a list-priority algorithm for the hidden surface removal, which required a substantial amount of preprocessing. Once computed, display time was fairly rapid, with two frames rendered per minute. The background consists of approximately 30,000 triangles. An early film of a walkthrough campus was created at that time.

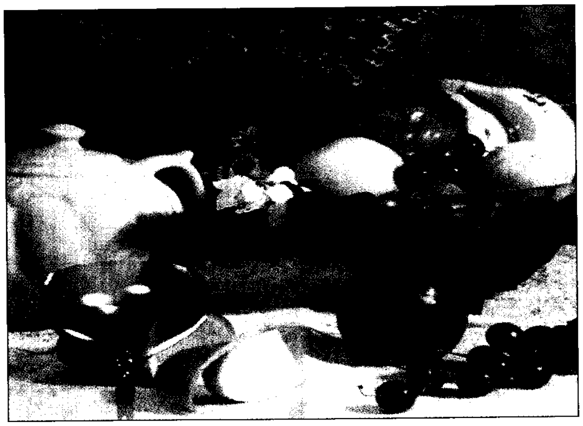
These concepts are now vestigial artifacts of a bygone technology. They are efficient for models that do not change and yield impressive dynamic motion displays that consist primarily of changes in the view position. The greatest benefit might lie in the exaggerated performance numbers that manufacturers can legitimately claim.

As we move into the next decade, many of our simulations will migrate, not from two dimensions to three dimensions, but from three dimensions to four or more with the added dimension of time and the display of multidimensional parameters. Kinematics, dynamics and earthquake analyses, fracture propagation, fluid flow, turbulence, weather modeling, and protein folding, just to name a few, are all simulations that change with time, both geometrically and topologically, and thus are models not well-suited for display lists. The time and cost for rebuilding the display lists are too excessive.

Even more important, scientists want to directly interact with their models, not pass through yet another layer of translations and abstractions before getting to their original data descriptions.

Polygons to pixels

The increase in model complexity will influence a radical change in the design of graphics hardware and the graphics pipeline. Current systems are based on polygons and use incremental algorithms for rapid scan-conversion and depth sorting (z-buffering) operations. These concepts, based on the simple observation that it is cheaper to make an incremental calculation based on previous computations than start from scratch, have served us well.



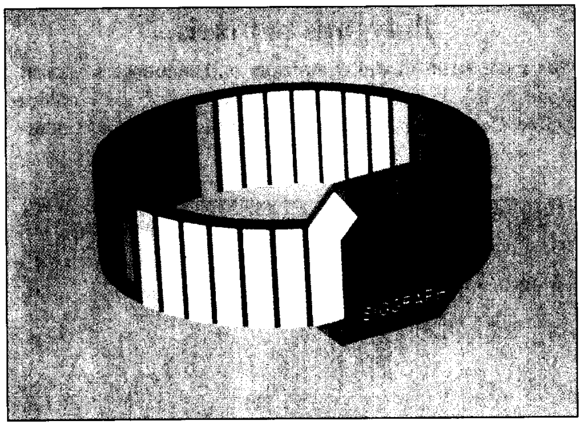
"Champagne Glass" by Doug Kay (Program of Computer Graphics, Cornell University, 1979) shows transparency and refraction effects. It was one of the early examples of ray tracing. Using a back-to-front order, sequentially mapping the current background image onto the next surface closest to the observer, the correct distortions due to refraction were modeled. The image took approximately eight hours to compute on a DEC PDP 11/50.

What happens when the number of elements abstracted from the geometry definition increases by two orders of magnitude? The availability of interactive modeling systems, plus meshing and adaptive subdivision techniques, will cause this to happen.

Today's graphics systems are "balanced" for polygons (usually triangles or convex quadrilaterals) that cover from 100 pixels/polygon (10 × 10) to 900 pixels/polygon (30 × 30). At small sizes, the transformations, geometric and intensity gradient calculations, and edge-sorting operations dominate; at large sizes, the scan-conversion routines take the most time.

Assume that workstations of the next decade increase to a resolution of two million pixels (1,920 × 1,084) to be compatible with the eventual high-definition television (HDTV) standards.

If model complexity remained constant, this doubling of resolution would imply a greater number of pixels per polygon. But with a hundred-fold increase in model elements, the average number of pixels covered per polygon reduces to only one or two. On the other hand, consider the case where 10 pixels are covered by an equilateral triangle. Nine of these represent the edges, which require the major computations. They will have to be antialiased and thus treated differently from the one interior pixel, which can benefit from incremental calculations. With 90 percent of the pixels requiring special operations, we lose all advantages of scan-conversion.



"Siggraph Watch" by Rob Cook, Ken Torrance, and Stuart Sechrest (Program of Computer Graphics, Cornell University, 1981) was directly illuminated by a single light source. This simulated watch was one of the early images created by using actual material properties. Stainless steel and gold of two-surface roughness was simulated using the Cook-Torrance reflection model. Using a polygon visible surface algorithm, the creators generated the image on a VAX 11/780 in approximately five hours.

The logical conclusion is that hardware must migrate from polygon algorithms to pixel algorithms based on the true geometries. A beneficial, and perhaps paradoxical, byproduct of this evolution will be a reduction in database storage requirements and a more accurate geometrical representation.

Sampling separate from discretization

Current methods for high-quality picture generation have intertwined the sampling procedures with the geometrical discretization of the environment. (Note that in this discussion I ignore the view-dependent approaches, such as ray tracing, where sampling occurs on the image plane.) All of the meshing or tessellation algorithms I know are inherently based on geometric subdivision techniques. Even surfaces parametrically defined are geometrically subdivided. This same hypothesis generally holds true for finite element and finite difference simulations.

The rendering problem is really a sampling problem. We want an accurate illumination function over the surface. For high-quality pictures, we must densely sample the surface at regions of high intensity gradients, for example, edge contours or shadow boundaries. In these regions, sampling should occur at high spatial frequencies, whereas in regions of low intensity gradients, low spatial sampling frequencies would suffice.

For rendering, this problem proves extremely difficult because it is unclear how to predict where the high density sampling should occur. In contrast to most engineering problems, where the effects are predominantly local, illumination of a surface is a global phenomenon affected equally by nearby or distant elements. We have not yet solved this problem, but its solution will have two important beneficial results. First, the calculation of surface illumination functions will provide a view-independent solution, then ultimately allow fast display routines. Second, the existence of the illumination functions will eliminate the database explosion that occurs with geometrical subdivision techniques. Furthermore, the adoption of surface sampling strategies as opposed to discretization methods will accelerate the change in display hardware to a pixel basis (see above).

"Entrance Hallway" by Roy Hall and Hang Weghorst (Program of Computer Graphics, Cornell University, 1983) is an early ray tracing image that used hierarchical bounding volumes and adaptive tree-depth control to reduce computation times. Light source descriptions that accounted for the spatial distribution of the emitted energy were incorporated using goniometric diagrams. The image took approximately five hours on a VAX 11/780. (Printed on the cover of *IEEE CG&A,* Vol. 3, No. 8, Nov. 1983.)

Improved display technology

For years, display technology far exceeded our ability to generate real-time images or render high-quality scenes. Now, we face severe constraints in both the size and dynamic range of display devices.

When we focus on an object, a 60-degree frustum of vision might be acceptable, but to simulate space, or just being inside of an object, requires peripheral vision.

Nineteen-inch monitors, viewed from usual distances, do not suffice. Since the cost for larger screen sizes with CRT technology becomes prohibitive, two types of display systems will become popular. Flat panel displays, either vertical or used as transparent table tops, or drafting boards will allow the presentation of substantially greater amounts of information. Multiple projection systems can do the same, and have the added advantage of surrounding the observer with the simulated space.

"Kitchen Counter" by Eric Haines (Program of Computer Graphics, Cornell University, 1986) shows a ray traced environment containing 224 objects (1,298 polygons, 4 spheres, 76 cylinders, and 35 quadrics) and five light sources. Using the light-buffer approach reduced shadow testing time substantially. The total computation time on a VAX 11/780 was approximately five hours. (Printed on the cover of *IEEE CG&A*, Vol. 6, No. 9, Sept. 1986.)

"Stair Tower" by Keith Howie and Ben Trumbore (Program of Computer Graphics, Cornell University, 1990) shows a ray traced simulation of the stair tower of the new Engineering Theory Center building at Cornell (designed by Gwathmey & Siegel). The image was modeled and generated using the image synthesis testbed at the Program of Computer Graphics. The entire image is synthetic with the exception of the texture-mapped foliage, taken from a real photograph of the site and mapped to a large transparent background cylinder. Computations were executed in parallel on a workstation cluster consisting of Hewlett-Packard HP835s and Digital Equipment DECstation 3100s and 5000s and took approximately five hours.

For increased quality, the dynamic range of the displays must increase substantially. Our visual system can roughly distinguish up to 12 bits of accuracy in luminance, yet the current standard color monitor technology is limited to eight bits.

What do we do now?

I submit that these ideas might not be forward-looking enough. I have trouble conjecturing what will happen when we really have supercomputer power on our desktops—a situation that will inevitably occur within a few years. As with most people, my imagination is constrained by my own past experiences.

I should leave these daring predictions to the newest generation of computer graphicists, just graduating now. Perhaps the IEEE Computer Society should publish a comparison issue of *CG&A* including the opinions of the newest students, ones who have not yet been to a Siggraph convention.

You see, the biggest obstacle ahead is not what we can do technologically in computer graphics, but how can we "untrain" ourselves, to stop continuing from past approaches and to use the new technology to ask the right questions. Hardly a technical problem, it demands adjustments in education and training. Despite the difficulty of teaching an old dog new tricks, we must shed our mental shackles to fully realize the graphics capabilities of tomorrow. This, appropriately, is the challenge for the next generation. I hope I (we) can become part of the solution. It offers exciting possibilities for all of us in computer graphics. ❑