Imagination Amplification

Michael F. Cohen
Microsoft Research

It would be easy to gaze into the distant future and speculate what image synthesis technology may be. Since *Star Trek* with its Holodeck has already done this far better than I can, prognosticating will not be my main focus here. Instead, I hope to relate some of my own inspiration for getting involved with computer graphics, an inspiration that is still an unrealized dream, one that I'm uncertain how to fully achieve.

My interest in computer graphics can best be summed up as a desire to create tools for imagination amplification. I suspect I did not invent the term imagination amplification, but I don't remember the source.

We all have thoughts about things that don't exist. These may range from simple musings, to architectural designs, to industrial design problems. We all have drawn on the backs of envelopes or napkins to try to convey a sense of our thought process. Drawing not only communicates, it also helps to clarify thinking. This process of creating images from ideas I call imagination amplification. Computers can help.

The computer can, in the form of an algorithm, be given information about how light reflects from objects, how materials bend and bounce, how ink flows on a piece of paper, how cameras create images from light, and more. Thus, from only partial information, the computer can simulate (to some approximation) the physics of materials, paint, and light to generate images we couldn't create on our own. The computer can show us how things will change when the input changes.

Realistic image synthesis has made tremendous progress over the past three decades. The first major body of work on raster images of synthetic objects came from the University of Utah in the 1970s. To understand just how indebted we are to this time and place in history, see http://www.cs.utah.edu/dept/history/. If the 1970s can be considered the decade of local illumination, the 1980s were the decade of global illumination. Ray tracing and radiosity algorithms provide the ability to capture the subtleties of interreflection and shadows. Figure 1 (left) shows a radiosity image I helped create 10 or so years ago at Cornell University.[1] Recent improvements have resulted in images such as Figure 1 (right), a Florence Courtyard.

Since their inception, ray tracing and radiosity methods have not changed fundamentally in their basic algorithm, although improvements in the 1990's have made them much faster and more widely applicable. Radiosity methods have given way to radiance methods that also capture nondiffuse interreflection. Stochastic and hybrid methods have brought together the best of both families of algorithms. These changes, along with the steady increase in computational power, bring global illumination methods to a point where they could be used on a wide scale. However, they aren't.

The frustration

Why aren't global illumination algorithms used routinely? The most obvious answer is that even with an unbelievably inexpensive modern graphics card, they're still too slow. We still cannot render scenes of interesting complexity at interactive rates.

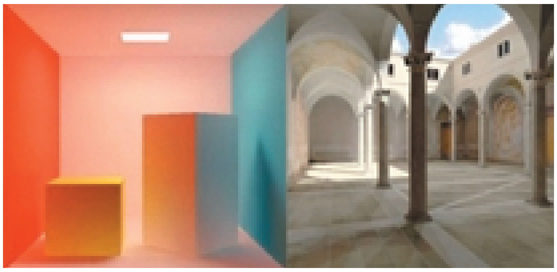
Computational costs aren't the only barriers to wider acceptance of global illumination algorithms. As algorithms become more sophisticated, the requirements imposed on the models to be rendered increase. Surface geometry and surface properties need to be more fully and more carefully defined. Radiosity methods often require subdividing surfaces in particular ways to make the algorithm produce accurate images. Both ray tracing and radiosity methods break when slight errors are included in the geometry, leading to ugly artifacts. Lights represented as point sources are no longer adequate. In practice, those designing virtual sets for games or movies often prefer the simpler control possible when lighting isn't realistic. Shadows may not even be a desirable feature. We cannot expect the majority of computer graphics users to go to the trouble to create accurate models.

A temporary refuge

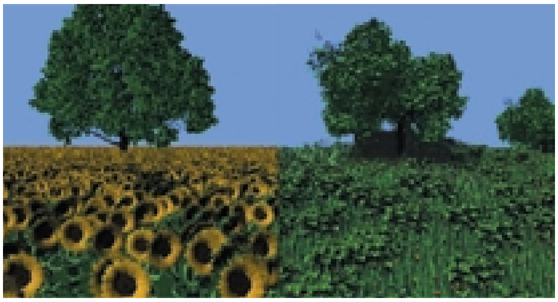
Recent years have seen the rise of a new paradigm for rendering images known as image-based rendering. IBR represents a marriage of sorts between computer vision and computer graphics. While computer vision tries to create a machine that can look at the world and form an accurate mental model of it, computer graphics tries to create a machine capable of taking such mental models and turning them into realistic images.

Unfortunately, if a state-of-the-art computer vision system is shown a complex scene, it will at best deliver a very coarse approximation (for example, it would not

0272-1716/00/$10.00 © 2000 IEEE



1 The left image known as the Cornell Box[1] demonstrates the color bleeding effects of interreflection faithfully simulated by the radiosity method. Andrzej Zarzycki modeled the right image of the Florence Courtyard (Piazza ss. Annunziata) rendered using Lightscape radiosity software.



2 Sunflowers and stream. These images are snapshots from layered depth images.[2] The viewer can, in real time, render this and nearby views. The LDIs were created and rendered by Jonathan Shade. The data was supplied courtesy of Oliver Deussen.

capture hairs on a stuffed animal). Even if it did, image synthesis algorithms wouldn't be able to deal with such rich detail in anything close to real time.

IBR combines these computer vision and graphics technologies by asking the question, what can I do to render new images directly from a set of images plus limited geometric information that computer vision algorithms can supply? In this way, we are freed from having to create detailed mathematical representations of the world, yet we can potentially create very realistic images. Lumigraphs[2] are one example, constructed from multiple images. The Lumigraph, much like a digital hologram, is a unified representation of what an object looks like from all directions.

The sunflower and stream images shown in Figure 2 were rendered in real time from another data structure referred to as a layered depth image.[3] An LDI falls somewhere between a Lumigraph and an image. It differs from a simple image in that it contains, for each pixel, a list of color values plus depth. The depth value provides the means to induce parallax as the viewpoint changes. As you move your head from side to side, things far away will stay still, while nearby objects will move back and forth in your visual field. Multiple, per-pixel, color-depth pairs fill in the gaps created when an occluded surface becomes visible.

Given all the interesting recent work in IBR, why title this section "a temporary refuge"? IBR provides a tantalizing means to capture very complex real-world geometry plus lighting and a way to leverage slow global illumination methods. However, IBR provides only limited help with imagination amplification. IBR by its nature doesn't help us ask what-if questions. Having to start from images to create new images puts us in a chicken-and-egg bind.

The hope

Despite the frustration just expressed, I'm still both amazed about the progress computer graphics has made to date and very hopeful for its future. Recent movies, such as *Jurassic Park, Titanic,* and *Star Wars Episode One*, have relied heavily on computer graphics rendering with great success. I would challenge anyone to spot when the dinosaurs are pure rendering and not filmed physical models.

The accelerating progression of ever-higher computational and rendering speeds at ever-lower costs will, in the not too distant future, bring real-time ray tracing within reach of an average PC user. This will fundamentally change the look we expect from computer graphics.

I also hope this will inspire a newly energized effort at developing tools for physical simulation. The power of procedural modeling and physically based animation and rendering will help expand our ability to explore new ideas. However, none of these advances will put the power of imagination amplification in the hands of nonspecialists. Better, easier, more intuitive tools need to be built to help us express our imagination. One of my favorite Siggraph papers of recent years is Igarashi et al.'s, "Teddy."[4] This paper presents the type of intuitive interface to modeling that, if extended more broadly, will bring the ability to create imaginary worlds much closer to everyone.

In addition to new technologies, people need to be empowered to participate in the use of new tools. Educational curricula should provide more focus on the kinds of skills needed to leverage this new computational power. The focus on traditional reading and writing skills should be expanded to include formats that are interactive and make increasing use of multimedia. We don't even have verbs to fully describe either the process of authoring or viewing such rich documents. Whatever the name we choose for these processes, we need to teach these skills. ■

References

1. M. Cohen and D. Greenberg, "The Hemicube," *Proc. Siggraph*, ACM Press, New York, 1985, pp. 31-40.
2. S. Gortler et al., "The Lumigraph," *Proc. Siggraph*, ACM Press, New York, 1996, pp. 43-54.
3. J. Shade et al., "Layered Depth Images," *Proc. Siggraph*, ACM Press, 1998, pp. 231-242.
4. T. Igarashi et al., "Teddy: A Sketching Interface for 3D Freeform Design," *Proc. Siggraph*, ACM Press, New York, 1999, pp. 409-416.

Contact Cohen at mcohen@microsoft.com.