Editor: Nevenka Dimitrova
Philips Research

Through the Looking Glass: The Synthesis of Computer Graphics and Computer Vision

Alyn Rockwood
Power Take Off Software

Janet McAndless
Capstone Systems

If we could only get through into Looking-Glass House. I'm sure it's got, oh! such beautiful things. Let's pretend the glass has got all soft like gauze, so we can get through.

—Lewis Carroll, *Through the Looking Glass*

In one abstract view, computer graphics deals with building computer models or representations and then displaying them by some method or theory to produce a high-quality image. A classical example is the Utah teapot benchmark (so-called because the actual teapot belonged to a professor at the University of Utah). The teapot is a typical model representation made of 32 polynomial surface patches. It has been rendered and illuminated by theories of occlusion, lighting, shadowing, and so on, to produce more images than perhaps have been published even of Marilyn Monroe.

Computer vision is a dual to computer graphics. It starts with an image or animation and deduces the model representation for the computer. A classical example here is classifying and sorting objects based on camera input. Taken together, this dual relationship looks like Figure 1. Computer graphics goes down from model to image, and computer vision goes up from image to model. This is, of course, a simplification, but it serves as a good basis to understand recent developments intersecting the two fields.

The image is it

First, there's the room you can see through the glass—that's just the same as our drawing-room, only the things go the other way.

—Lewis Carroll, *Through the Looking Glass*

Computer graphics and computer vision have had a somewhat checkered relationship. As both have matured, however, the differences between the disciplines has softened and thinned. This natural trend has been observable for some years in areas such as image-based modeling and rendering (IBMR).[1] IBMR uses the viewing information plus the image itself to reconstruct the underlying geometry or a re-rendered image. The model, then, is the image plus the information of how it was captured (such as range data and the location of light sources).

To better understand the interrelationship of computer graphics and computer vision, we'll showcase a representative set of papers that meld these two disciplines. We drew from papers accepted at the Siggraph 99 conference in Los Angeles. From the conference we see that a significant number of researchers in computer graphics are drawing inspiration from computer vision.

An outstanding example of such work presented at Siggraph 99 was the research paper "A Morphable Model for the Synthesis of 3D Faces."[2] A technique for modeling textured 3D faces, Morphable Faces begins by morphing a collection of photographed faces to produce a model of "the average face" in three dimensions. Given a new target photograph, the author's technique orients and scales this average face to approximate the photographed image's position. At this stage, an optimization algorithm attempts to perturb the 3D model until it looks like the photograph. The authors accomplish this effectively using their own novel routines. Once the model exists in three dimensions, you can apply the range of computer graphics techniques to shade, texture map, morph, and so on.

The authors presented impressive results including the reconstruction of believable, texture-mapped 3D faces of Tom Hanks and Audrey Hepburn from single frames of film. Using this method, we can reconstruct a probable facial rep-

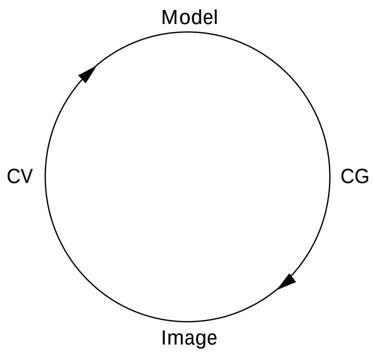
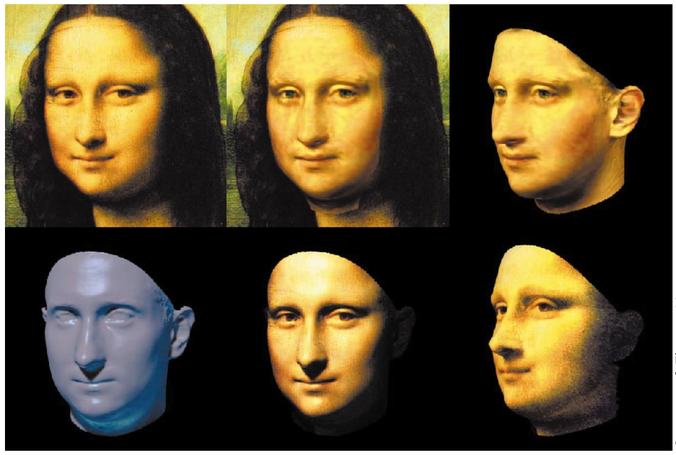


Figure 1. The dual relationship between computer graphics and computer vision.



Courtesy of Thomas Vetter

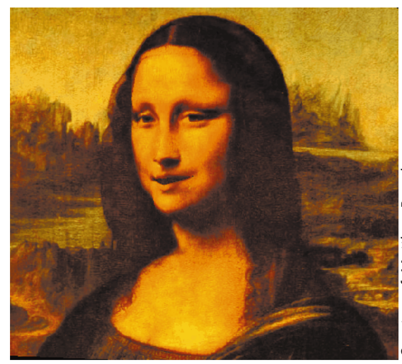
Figure 2. The "Mona Lisa" in 3D.

resentation of the "Mona Lisa," viewable from multiple angles (see Figure 2). This work also demonstrates face manipulations according to complex parameters such as gender, fullness of a face, or distinctiveness.

Adding research from a second paper, "Voice Puppetry,"[3] we can begin to envision myriad applications wherein the "Mona Lisa" instructs visitors on her genesis, history, and creator (see Figure 3). Movie postproduction can edit and replace sound tracks without requiring the presence of actors. Educators can use animated historical figures or fictional characters to teach material (see Figure 4). Such near-horizon applications will breathe new life into the concept of virtual actors. In the Voice Puppetry work, 2D reference video frames and sound are used to train a 3D model to exhibit facial expressiveness including mouth and eyebrow animation. A new voice track can then trigger the model's learned expressive behavior.

> You can just see a little peep of the passage in Looking-Glass House, if you leave the door of our passage as far as you can see, only you know it may be quite different on beyond.
>
> —Lewis Carroll, *Through the Looking Glass*

Researchers commonly apply IBMR techniques to environments or scenery. For example, with "Inverse Global Illumination: Recovering Reflectance Models of Real Scenes from Photographs,"[4] the authors presented a technique that uses image-based rendering to construct estimates of radiance and directional irradiance for small patches of the model. These estimates are then used to create high-resolution illumination maps. The technique develops a lighting-independent model of the scene that can be re-rendered under differing light conditions and includes newly introduced digital elements. The film and television visual effects industry as well as architectural space planners could benefit greatly from such techniques.



Courtesy of Matthew Brand

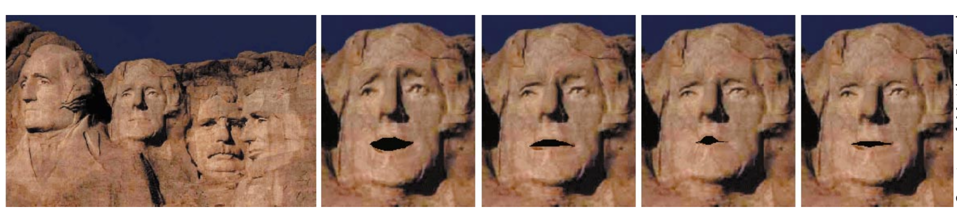
Figure 3. The "Mona Lisa" as Voice Puppet.

Interestingly, technology transfer and the rate of advancement moves so quickly in the computer graphics field that already much of the earlier global illumination work by the authors discussed above was applied in the paper "Reflection Space Image-Based Rendering"[5] to the problem of rendering high-quality, physically accurate scenes at interactive rates. With this work, the bidirectional reflectance distribution function (BRDF) for a lighting environment can be precomputed on a simple cube or sphere. The result drives an environment-mapped lighting model, which is fast, but closely approximates the feel of the physical environment.

The flowering of synthesis

> "Where do the paths lead?" asked Alice.
>
> "One leads to the left and the other to the right," replied the caterpillar.
>
> "Which one should I take then?" she asked, perplexed.

Figure 4. Mount Rushmore characters recite "The Jabberwocky."



Courtesy of Matthew Brand

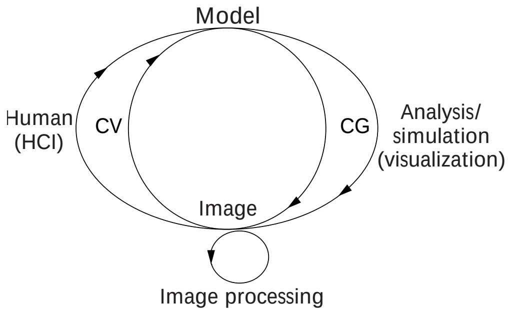


Figure 5. The schema for computer-based imagery.

"That all depends," replied the caterpillar, "on where you want to go."

—As Lewis Carroll might have written

The previous examples illustrate the need for a significant level of synthesis and raw computational power. The Morphable Faces[2] paper draws on refined methods of morphing objects in 3D and on optimization theory, not to mention the underlying assumptions about occlusion, shadowing, texturing, and illumination. These assumptions occupied computer graphics through much of the 1970s and 1980s. Now this expertise is in place—in many cases as a well-defined commodity purchased at the nearest electronics store.

With this firm foundation, we can begin building applications. Although advances made in the basic tools of computer graphics and computer vision will continue, the future will see more and more amalgamating of ideas. Where will the pathways lead us? To begin, let's take another look at Figure 1. You can see that we're missing some obvious elements to round out the world of computer imagery. From the image we can draw another arrow, this time to the human being, and thence on to the model. This represents the area of human-computer interfaces, or HCI. From the model, we draw an arrow to analysis and physical simulation, and from there to the images. This represents the area of visualization, which includes scientific, financial, engineering, medical, and data visualizations to name a few. To finish our diagram we cycle another arrow from the image to itself, creating the area of image processing. (See Figure 5.)

Near-future possibilities

But how curiously it twists! It's more like a corkscrew than a path!...Well then, I'll try it the other way.

—Lewis Carroll, *Through the Looking Glass*

The following presents some of the possible research areas implied by current trends.

Reaching the cognitive centers of our minds

Current graphical user interfaces (GUIs) and other computer interfaces have emerged as a side effect to the advent of desktop computing. However, we now have a powerful new toolbox at our disposal, together with a new generation of users whose experiences with ubiquitous computing will push the interface to the next level. For example, the synthesis of much current work in computer graphics and computer vision brings us closer to realizing the intelligent interfaces seen in popular science fiction—often the portent of future trends.

Involving the other senses

The field of virtual reality has driven much progress in immersive computing experiences, but a whole field of inspiration comes from the area of interface gadgets beyond the monitor, keyboard, and mouse. Science fiction has long dreamed of the ultimate computer interface that will "wet-wire" directly to our brains. What might the new generation of users develop to improve upon head-mounted displays and power gloves? As stated previously, the field of HCI has much to add to computer vision and computer graphics as our interfaces move off the desktop.

Smart environments

The concept of the environment as the interface follows our interest in cognition and sensory perception. Desktops and file systems draw from our experience, but researchers have begun to explore new paradigms that take the computing space off the CRT monitor. An example appears in the work "Emancipated Pixels: Real-World Graphics in the Luminous Room,"[6] which implements smart work spaces using the IO Bulb, a two-way optical transducer (projector and camera) that interprets physical inputs from and projects graphics onto arbitrary surfaces. An especially interesting application of this work examines wind flow around buildings in the tabletop planning of a business park. In this example, a simple architectural figure is placed on a table with a wind indicator, from which the system computes the wind field and displays it around the object. You can easily see how this application ties into current work[7,8] to stabilize and interactively compute physics-based models for fluid-like behavior.

Networked graphics

The ubiquity of access to the Internet and the World Wide Web has opened new doors of communication and exchange, even as it has demonstrated the need to reexamine and evolve our current methods. From new designs for transfer protocols to real-time algorithms and data compression, researchers are responding to the challenges presented by user expectations for instant access. New areas of application will need to overcome latency issues and enable the transfer of simple data models from which complexity can be reconstructed at the receiving end. For example, the receiver's system might construct a holographic 3D human figure texture-mapped with the sender's image. A camera at the sender's end might capture motion and expression, cull the data, and transmit updates in real time.

Better algorithms and simulations

With the enormous computing capacity available and its increasing rate of advancement, the challenge lies in discovering the right algorithms to deliver more accurate results, bringing us closer to physically accurate computer modeling. Much of this work will involve arriving at results nondeterministically and genetically.[9] Armed with an accurate and true understanding of the physical world, we'll be better equipped to violate its rules and engender breakthroughs that continue to advance numerical analysis techniques at a rate outpacing Moore's Law by the order of magnitude seen in current research.

Conclusion

> "Fan her head!" the Red Queen anxiously interrupted. "She'll be feverish after so much thinking."
>
> —Lewis Carroll, *Through the Looking Glass*

Computer graphics and computer vision are truly complementary disciplines quickly approaching convergence. As noted, the broad study of computer-based imagery extends beyond these two fields to include the areas of HCI, visualization, and image processing. Where does it go from here? As stated so eloquently by Sir Isaac Newton, "Standing on the shoulders of others, we see farther." Ongoing research and development will continue to forge this bond, and we'll begin to see real-world products emerge from these efforts. Then we will see the fruits of this convergence. MM

Acknowledgments

We'd like to thank the Siggraph 99 Papers Committee and referees for their time and diligence in selecting for this year's proceedings an outstanding collection of leading-edge research within the field of computer graphics and interactive techniques. We also thank Visions and Views department editor Nevenka Dimitrova for her excellent suggestions, guidance, and most importantly, patience. Finally, we thank Dino Schweitzer, Siggraph Chief Staff Executive, for providing valuable feedback on our ideas.

References

1. S. Chen and L. Williams, "View Interpolation for Image Synthesis," *Proc. Siggraph 93*, ACM Press, New York, 1993, pp. 279-288.
2. V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 187-194.
3. M. Brand, "Voice Puppetry," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 21-28.
4. Y. Yu et al., "Inverse Global Illumination: Recovering Reflectance Models of Real Scenes from Photographs," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 215-224.
5. B. Cabral, M. Olano, and P. Nemec, "Reflection Space Image-Based Rendering," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 165-170.
6. J. Underkoffler, B. Ullmer, and H. Ishii, "Emancipated Pixels: Real-World Graphics in the Luminous Room," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 385-392.
7. J. Stam, "Stable Fluids," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 121-128.
8. P. Whitting, "Computational Fluid Dynamics in a Traditional Animation Environment," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 129-136.
9. J. Funge, X. Tu, and D. Terzopoulos, "Cognitive Modeling: Knowledge, Reasoning, and Planning for Intelligent Characters," *Proc. Siggraph 99*, ACM Press, New York, 1999, pp. 29-38.

Rockwood was the Siggraph 99 Papers Chair. Contact him at Power Take Off Software, e-mail rockwood@enuxsa.eas.asu.edu.

McAndless was the Siggraph 98 Emerging Technologies Chair. Contact her at Capstone Systems, e-mail janet@siggraph.org.

Contact Visions and Views editor Nevenka Dimitrova, Philips Research, 345 Scarborough Rd., Briarcliff Manor, NY 10510, e-mail nevenka.dimitrova@philips.com.