EVERY PICTURE TELLS A STORY

IMAGE-BASED MODELING AND RENDERING INTRODUCES NEW METHODS FOR CREATING PHOTOREALISTIC IMAGERY

Photographs are a deeply important part of our lives. They store information about where we have been, what we have seen, and we cherish their power to evoke a remembrance of things past. It’s no wonder that family photo albums (or, increasingly, CD-ROMs) are among our most valued possessions.

Photographs are achieving similar importance in the field of computer graphics. In the emerging field of image-based modeling and rendering, researchers and software developers are discovering and using the information contained in photographs to create new models and images, along with new methods for solving visualization problems in diverse fields from architecture to entertainment.

A traditional goal of computer graphics is realistic image synthesis, and most of the work done in the past three decades has followed a standard approach toward achieving that goal: First determine an accurate 3D model and then render that model in physically realistic light. Image-based modeling and rendering follows an inverse path to the same goal: Start with realistic images and interpret them to determine geometric models or synthesize other realistic images.

In general, the term "image-based modeling" is used when a traditional geometric 3D model is derived from one or more photographic images. The term "image-based rendering" is used when a set of reference photographs is processed to produce a new set of synthetic images, usually representing new points of view, without creating a geometric model. For many problems the methods overlap, and the terms are used interchangeably.

The profound advantage of photographs in computer graphics is that they are, of course, photorealistic. Imagery synthesized from photographs retains the visual detail and complexity of the original, resulting in a more convincing look to the created image. Image-based approaches also offer much faster render times. "Image-based methods allow us to achieve a holy grail of..."
computer graphics: completely realistic renderings in real time" says Paul Debevec, a researcher at University of California, Berkeley. "Instead of modeling complicated geometry, materials, and lighting from scratch, we can tap directly into the richness of the real world. We're seeing the result of real things illuminated by real light, with all the complex shading and lighting the world has to offer."

**A Picture Is Worth a Thousand Polygons**

The idea of using a 2D image to represent a 3D scene has been around for many years. Since the early days of cinema, moviemakers have created matte paintings to describe objects—from castles to cities to spaceships—that would otherwise be too expensive to build as sets. These matte paintings are designed to integrate with live-action elements from a predetermined camera point of view, and they can be considered the first examples of image-based modeling and rendering.

Another simple but practical form of image-based rendering has been around for 20 years in the form of texture maps, which were originally designed as a convenient photorealistic solution to the time-consuming problem of modeling structures with complex surfaces. For many scenes, particularly natural scenes that are difficult to model accurately (such as a brick wall, a concrete sidewalk, or stones on a beach), a photograph-based texture map greatly simplifies the geometry of the scenes and significantly reduces the render time, while maintaining a high level of photorealism. Texture maps are also important because of the development of an image-based technique—the projective warp—that transforms the texture image into the 3D perspective geometry of the rendered scene.

The technique of image morphing, which first appeared in mass culture in the 1988 movie Willow and in the Michael Jackson video "Black and White" represents a significant step forward in image-based rendering. Given two photographs as reference images, new images are created from the visual transition between the two reference images. Morphing is important for two reasons: the synthetic transition images are procedurally generated from the reference images, and the transition images themselves are astoundingly photorealistic.

In a 1993 paper, "View Interpolation for Image Synthesis" Eric Chen and Lance Williams, from Apple Computer, proposed an idea for synthesizing intermediate images from computed images at specific viewpoints as a substitute for rendering each scene directly. They proposed their method as a way to avoid the computational expense of continuously rendering complex geometry, especially in flight-simulation and virtual-reality applications.

Many researchers in the past several years have recognized the value of photographic imagery as starting points in the process. "A strength of the image-based modeling and rendering approach is its reliance on images as a basic primitive. Because images are easy to capture, store, manipulate, and transmit on existing platforms, current image-based techniques enable real-time rendering without the need for special-purpose hardware," says Steven Seitz of Carnegie Mellon University.

QuickTime VR, a commercial product developed at Apple Computer in the early 1990s, is an example of image-based software that runs on desktop computers. QuickTime VR uses continuous panoramas of "stitched" photographs, rather than modeled and rendered geometry, to create an immersive virtual-reality (VR) environment. Chert, who was the co-founder and chief technologist of the QuickTime VR team, decided to use panoramas (cylindrical projections of an environment) because they could be created easily with a special camera. He later invented an automatic stitching process that uses photos from regular cameras, to create a continuous panorama.

This stitching process and use of regular photographs made the creation of photorealistic panoramas and immersive virtual environments accessible—through QuickTime VR—to a greater number of people. The ability to create a personal image-based immersive environment is now
available to anyone with a camera and a computer. According to Chen, image-based modeling and rendering "gives nonskilled operators the ability to create photorealistic models and images that once could be created only by highly trained professionals. The act of creating 3D models may become as mechanical as taking pictures is today."

**Architectural Models**

Image-based modeling and rendering has achieved impressive results in the field of architecture. In one striking example, Paul Debevec and his colleagues at UC Berkeley used a set of reference photographs--20 images of the Campanile and surrounding campus buildings at UC Berkeley--to create a fly-by animation titled "The Campanile Movie." By collating points and lines of perspective on the reference photographs, they generated a 3D model of the campus. They then applied projectively transformed images from the reference photographs to the geometric model as texture maps. The near-photographic results suggest a new kind of virtual camera that can transform a handful of images into a convincing animation from any imaginable tracking path.

The most commercially active area of image-based modeling and rendering in architecture is in photogrammetric modeling (photogrammetry is the science of making measurements from photographs). Three applications that create 3D models from photographs are currently available. They are used mainly for architectural modeling, but they can also be applied to problems in archeology, anthropology, accident reconstruction, mechanical engineering, and 3D design.

3D Builder Pro, from 3D Construction Company (Elizabethton, TN), is a photogrammetric application for Windows that is designed to convert ordinary photographs into dimensionally accurate 3D models. A user begins the measurement process by importing digital images (either from a digital camera or from scans of traditional photographs) into 3D Builder Pro and placing collated points on reference targets in each of the images. The user then locates lines and surfaces of interest in the images, and the 3D Builder math engine extracts measurements and creates 3D models from this data. The original digital images are projected into proper perspective and attached to the models as texture maps. The models can then be exported into an appropriate rendering application in a variety of formats.

Photomodeler Pro, from Eos Systems (Vancouver, BC), is another Windows application that extracts 3D models, data, and measurements from a set of overlapping photographs of an object or a scene. These photographs can come from a variety of digital-image sources, including video cameras. The user marks features of interest on the photographs, such as corners and edges, and PhotoModeler processes the marked images into points, lines, and surfaces in 3D space. Accurate photogrammetric measurements can then be made from the 3D model, and the model can be exported to a variety of applications, including CAD and VR.

CitiBuilder, from Gentech (Tokyo), is designed primarily for creating models of buildings or other flat-surfaced polyhedral structures such as bridges. The user creates preliminary models by assembling predefined generic blocks into structures that resemble the buildings in the photographs. Then CitiBuilder's reconstruct function accurately estimates the size of the model elements from image data specified in the photographs. Like the other photogrammetric applications, the images are used as texture maps on the model.

Gentech also offers Gen-Trix Studio, which allows users to create image-based 3D VR models. The user takes digital pictures of an object on a special Gen-Trix stage, and Gen-Trix Studio processes the image data to create a realistic VR model that can be displayed interactively on a Web page. Like CitiBuilder, Gen-Trix Studio uses the digital images as texture maps to create greater photorealism. Because the Gen-Trix VR model uses only 3D data and a few images, the size of the model is significantly smaller than current VR models that use sequential movie-like images.
Facial Animation

Animating the subtle and intricate motions of the human face has long been a problem in computer graphics. In spite of many attempts to refine geometric models of facial structure and develop techniques to reproduce delicate facial movement, facial animations usually have a stiff, waxy, artificial look that is distinctly nonphotorealistic. Recent work in image-based techniques, however, has produced some exciting new results.

Researchers at the University of Washington, the Hebrew University, and Microsoft Research have combined the visually accurate technique of image morphing with the transformational flexibility of a 3D model of the human face to create startlingly realistic facial animations. Their technique, presented in a paper "Synthesizing Realistic Facial Expressions from Photographs" at the SIGGRAPH 98 Conference, begins with a series of digital still photographs of a subject. Distinct facial elements are marked with points in each of these photographs, and the points are used to determine 3D positions of the facial elements. The 3D positions are then applied to a generic 3D face mesh to create an accurate geometric model for the particular subject. Texture maps from the original photos are developed into a gallery of photorealistic facial expressions that are morphed and applied to the 3D face model.

In ArtiFace, control points on a 3D face model are linked to similar points on video imagery, The facial movements are then mapped onto the 3D model.

The success of this technique depends on an animation interface that allows the artist to create facial images from a set of reference images called the expression gallery. The artist uses this animation interface to mix any of the different reference expressions into seemingly unlimited combinations. In addition, the researchers have developed a sophisticated method of weighted morphs of the texture maps and the reference images to control subtle transitions between expressions. The resulting facial animations demonstrate an unprecedented photorealistic accuracy.
Commercial applications of image-based facial animation have become available in the past year. TechImage Ltd. (Herzelia, Israel), recently announced a facial-animation application called ArtiFace. This application, a plug-in for SoftImage 3D and Alias/Wavefront PowerAnimator, combines live-action facial expressions filmed on video with a user-chosen 3D model of a face. ArtiFace uses proprietary image-analysis algorithms to transfer the 2D motion data from the video onto the 3D model without the need for markers or other user-intensive intervention. The motion data and the 3D model are combined to produce a realistic animated sequence that accurately imitates the original facial gestures and speech. An application such as ArtiFace has obvious advantages in high-volume animation houses, where accurate animations need to be produced quickly and efficiently from prerecorded soundtracks.

Sven Technologies (Palo Alto, CA) has a similar application called SurfaceSuite, which uses an adaptive image-based surface texturing algorithm to map digital image data such as facial photographs onto 3D models. SurfaceSuite, currently available as a 3D StudioMax plug-in, uses strategically placed control points on a digital image and a model to adapt the image to fit the model. The process is interactive, so control points can be moved or added until the resulting texture mapping is accurate. SurfaceSuite can also combine photographs to create seamless blends and wraparounds.

Some research-generated techniques in computer graphics are developed well before significant uses are found for them. Facial animation, however, is an area of research that has many potential uses. Character animation in films and advertising, video teleconferencing, and on-line facial avatars for chat rooms are three application areas that could benefit enormously from improvements in image-based facial animation. Web-based applications could include real-estate visuals, travel destinations, and on-line shopping.

The idea of a virtual actor is a compelling motivation for new work in facial animation. "What about producing a movie of Abraham Lincoln giving the Gettysburg address?" asks MIT's Tomaso Poggio. "We have a system for text-to-visual speech that could be used someday to make anybody you want (provided you have a picture of him or her) say whatever you want."

**Entertainment and Visual Effects**

Extraordinary examples of image-based visual effects have begun to appear in commercials, videos, and movies. Some of the most striking work has been done by Parisand Los Angeles-based BUF Compagnie, which has done commercials for Perrier and Renault, movies such as City of Lost Children, and astonishing morphs in the Rolling Stones video of the classic Bob Dylan song "Like A Rolling Stone."

Another interesting image-based effect was created, primarily for use in commercials, by New York-based Digital Air and its TimeTrack camera. This camera, which has a series of either forty or eighty lenses mounted on a flexible track, can take simultaneous photographs from different camera positions. When transferred to film, this sequence of still images creates the visual effect of a camera moving around a still image, and has been used in commercials for The Gap (a leaping swing dancer), Miller Lite beer (the exploding beer spray), and other products. By attaching standard movie cameras to either end of the TimeTrack camera track, the filmmakers can create the look of a camera that continues to track even though filmed action is frozen in midair.

Another new technique, which draws on the imagebased architectural modeling and rendering methods developed at UC Berkeley, could eventually replace motion-control photography. George Borshukov of the Campanile project describes this technique, which is called virtual cinematography. "From a few still photographs of an original movie set, we can then match or create any camera move the director wants. Virtual cinematography has been pursued by many industry leaders as a way to provide directors with freedom of camera motion and lighting in
postproduction, within movie sets that have already been taken down." Borshukov is applying these techniques to the visual effects in an upcoming movie, The Matrix, starring Keanu Reeves.

An application called SynaFlex, from SynaPix (Lowell, MA), is one of the first high-end 3D applications to make use of a variety of image-based techniques. The application creates what is called the Virtual Theater, where 2D live-action film and video images are transformed into 3D surface representations and combined with special 3D computer graphics effects. The primary advantage of this tool is the creation of a unified interactive workspace for compositing footage and effects, without the often tedious limitations of working in traditional compositing layers. Lights and shadows in the Virtual Theater interactively affect all the elements at once, and digital effects can be attached to real, moving objects quickly and easily.

**An Evolving Technology**

Image-based methods can solve many problems in computer graphics, but they are not a cure-all. "Image-based modeling and rendering is limited by difficulties in dealing with a broad range of viewpoints and with control of the illumination" declares Poggio. "It is not a replacement for the more traditional physics-based approach."

Eric Chen has a similar complaint. "On the modeling side, capturing accurate range data, necessary in most image-based techniques, is still the biggest hurdle. On the rendering side, free navigation and interaction with an image-based environment is still a difficult task" he says.

One of the biggest problems is getting usable information--particularly lighting information--from photographs in the first place. "It's always hard to acquire all the data that you would want, especially for surfaces that are shiny in interesting ways," says Berkeley's Debevec. "Sampling issues make it difficult to produce consistently high-resolution images. Most importantly, direct image-based modeling and rendering techniques record only the appearance of a scene or object under given lighting, which makes it difficult to modify, add to, or relight the scene."

Barry Ruff, a rendering specialist with SynaPix, summarizes the state of image-based work: "As with any new technology, everyone is doing things their own way. Techniques in image-based modeling and rendering are ad hoc, and each developer is inventing (or reinventing) what they need as they go."

Despite these limitations, opportunities for creative new uses of image-based modeling and rendering are practically unlimited. Academic researchers continue to explore new possibilities for image-based work, and commercial products using image-based techniques are beginning to appear in the marketplace.
such as these single-image multiple-perspective views created by researchers at University of North Carolina.

One of the areas primed for image-based development is computer games. The use of image-based techniques will not only improve the real-time rendering speed and detail of games but will also bring a heightened level of photorealism. Eric Chen sees the use of image-based techniques as an obvious next step. "The game quality will be drastically improved, and the production time will be cut down through the use of images. The new games will have Myst's visual quality and Doom's playback speed." The use of facial animation in game avatars could bring a greater level of personal immersive realism as well.

Future uses of image-based modeling and rendering will continue to develop in extraordinarily powerful and perhaps unimaginable ways. "Image-based modeling and rendering has already changed our conception of what computer graphics can do," saysDebevec. "What will inevitably happen is that we will want to go beyond what is real--to reinterpret these sampled spaces through our own creativity--and to do this we need conventional computer graphics to become more integrated with what can be done with image-based modeling and rendering. Successfully bridging image-based and traditional graphics, making it possible not only to sample the world but also to bring it to life with our own visions, will bring forth some of the most powerful tools to create and entertain yet developed."

For More Information

PhotoModeler Pro

Eos Systems, Inc.
205-235 West 12th Ave.
Vancouver, BC, Canada v6J2G2
604-732-6658 phone
604-732-6642 fax
www.photomodeler.com
InfoNOW 170

3D Builder Pro

3D Construction Company
122 Creative Station
Elizabethton, TN 37643-5358
423-543-8917 phone
423-543-4011 fax
www.3dconstruction.com
InfoNOW 171

CitiBuilder and Gen-Trix Studio
PHOTOS (COLOR): A photorealistic front view of the Mona Lisa (center) was created from the original image (left) and a mirror image (right) by using an image-based rendering technique called view morphing.

PHOTO (COLOR): Researchers at Berkeley used image-based modeling techniques on this set of still photographs to generate a 3D model of the Campanile and surrounding campus buildings.

PHOTO (COLOR): A photogrammetric application such as 3D Builder Pro uses control points on photographs of an object to create a 3D geometric model.

PHOTOS (COLOR): A 3D model of the Campanile (top) was derived from the reference photographs (opposite) and used to create a synthetic photograph of the structure (directly above) from a new point of view.

PHOTO (COLOR): PhotoModeler Pro has been used on complex 3D modeling projects such as the restoration of a stone bridge in Germany.

PHOTOS (COLOR): Strikingly expressive facial animations can be created by combining geometric 3D face models with morphed imagery from still photographs of a person’s face.

PHOTO (COLOR): Satelite, Inc. of Japan created this anime image by using Houdini from Side Effects Software to combine traditional 2D animation with application-level techniques of particle systems, motion blur, and image-based modeling and rendering.
PHOTO (COLOR): Academic research in image-based modeling and rendering continues to develop new of photographs.

~~~~~~~~

By RANDALL WARNIERS

Randall Warniers is managing editor of the Lincoln Laboratory Journal. He can be reached at warniers@ll.mit.edu.

____________________________________

Copyright of Computer Graphics World is the property of California Offset Printers, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder’s express written permission. However, users may print, download, or email articles for individual use.