Image-Based Rendering: Really New or Déjà Vu?

Organizer
Michael Cohen (Microsoft Research)

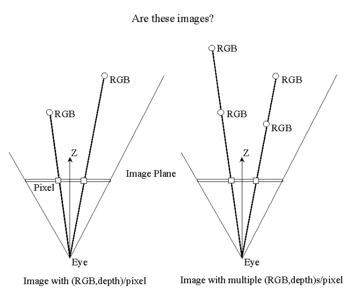
Panelists
Marc Levoy (Stanford University)
Jitendra Malik (University of California, Berkeley)
Leonard McMillan (Massachusetts Institute of Technology)
Eric Chen (Live Picture)

Michael F. Cohen

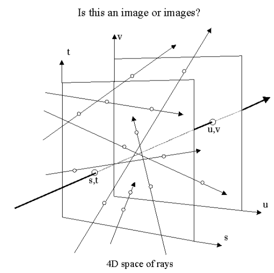
Image Based Rendering (IBR), techniques which generate new images from other images rather than geometric primitives, appeared to burst onto the computer graphics scene in the last few years. IBR seems to hold the promise of leaping right over the difficulties of traditional modeling and rendering or at a minimum hiding the latency between rendered frames. This panel will first define what is meant by image based rendering and place this emerging technology in the context of a continuum of developments. The panel of IBR researchers will then speculate on the long-term impact of IBR on computer graphics by addressing issues such as:

* what is IBR good for?
* will IBR replace polygons?
* what could IBR mean for graphics on the net?
* what is the relationship between IBR and traditional computer vision research?

Perhaps the first task is to redefine our own understanding of the terms “image”, “model”, and “rendering”. Traditional image synthesis rendering has meant simulating the flow of light from a source, reflecting it from a geometric and material description of a model, into a simulated camera and onto a film plane to produce an image. (Of course all this is done backwards if you are a ray tracer.)



In particular, the result of this process, an “image”, has generally been considered a 2D array of RGB valued pixels (or if you prefer, a continuous RGB valued function in 2D). What if a depth value is added at each pixel (or point in 2D)? Is this still an image? Or is this now a model consisting of colored points floating in 3D? What if instead of a single RGB and Z (depth value) per pixel, there is a list of these? Is this still an image, or a sparse volume? What is a 4D light field (or lumigraph) relative to an image?



The computer graphics community is already familiar with concepts such as texture mapping, and environment mapping. These are certainly a form of image based rendering. What is the new excitement all about?

To begin to answer these questions, it is important to look at what has been achieved in the computer vision community as well. Vision researchers have struggled with the problems of producing accurate models from images, while the graphics community has tried to both create models through sophisticated user interfaces and to render these models. Vision technology continues to have a difficult time producing the rich detail in complex geometric models (but does have the visual effects of the detail in the original images). Meanwhile, graphics technology has difficulty in producing the same geometric detail and also in rendering its effects. However, by placing the partial results of computer vision methods end to end with those from graphics, we seem to be able to (in some cases) simply skip both difficulties. To a great extent, this has been the cause of the excitement. However, this paradigm is not yet well understood and its limitations are even less well understood. Hopefully, this panel will help clarify the questions if not the answers in this emerging research endeavor.

Marc Levoy

The study of image-based modeling and rendering techniques is essentially the study of sampled representations of 3D objects. This idea has been around for a long time in the form of textures, sprites, shadow and environment maps, range images, movie maps, and so on. In computer vision, classic examples include reflectance maps, disparity maps, optic flow fields, and epipolar volumes. The new vigor in this area seems to arise from two factors: a marked increase in the dimensionality and size of the representations - 3D volumes, 4D light fields, 5D plenoptic functions, and hierarchical image caches to name a few - and a shift toward fast, robust algorithms that combine techniques from graphics and vision.

Numerous engineering challenges must be overcome to make image-based techniques practical, including acquisition, compression, fast display, and software interface. The research community will undoubtedly respond to these challenges, leading to many SIGGRAPH papers in this area.

In this panel, I would like to address a different question: what are the outer boundaries of image-based modeling and rendering? In other words, how far can we push this paradigm? This question can be approached from several viewpoints (no pun intended):

* What other sampled *representations* of 3D objects are possible?

  Light fields were recently introduced into computer graphics as a 4D scalar field, but Adelson's plenoptic function, which inspired them, is a 5D scalar field, and Gershun's original light field was a 3D vector field. What other definitions are possible? As another example, light fields store radiance. Is there a place for irradiance fields? How about importance fields?

* What *operations* are possible on image-based representations?

  For example, to what extent can objects be edited via images of them? Can one write a paint program or a 3D modeler that operates entirely on light fields? Is knowledge of surface orientation essential in order to compute reflection, or depth in order to compute occlusion? Are there weaker forms of knowledge that suffice for these tasks?

* Recent work in image-based rendering has focused on walkthroughs; what are other *applications* exist for these techniques?

  The vision problem - determining shape from images - is known to be hard. If we are willing to represent shape as images, perhaps as a set of inconsistent range maps rather than as a geometric model, does this simplify the vision problem? Another hard problem is computing global illumination in complex environments. The inverse global illumination problem - determining surface reflectance from measured radiance in the presence of interreflections - is even harder; it has never been solved except on very simple scenes. Can image-based representations help solve these problems?

* Finally, any change in paradigm engenders a crisis in *scholarly methodology*; how shall we judge and compare the results of the new paradigm?

  For geometry-based representations, the graphics and computational geometry communities have developed analytical tools (e.g. visibility aspect graphs), measures of rendering cost (e.g. numbers of polygons, depth complexity, pixel count), and measures of image quality (e.g. aliasing, variance, discrepancy). Image-based representations will require new analytical tools and new metrics.

Leonard McMillan

At this early stage in the development of image-based rendering (IBR) it is worthwhile to ask several pertinent questions:

What good is it?
How will it be delivered?
Where is it going?

IBR can be utilized in many of the same applications where conventional geometry-based computer graphics is employed today. The advantages of IBR methods include the easy acquisition of models from images, computational requirements that are independent of scene complexity, and the delivery of photorealism while avoiding the costs of physical simulation. IBR also holds promise in many new applications such as latency reduction in the network transmission of three-dimensional environments, and in virtual and augmented reality applications (with an emphasis on reality).

In order for IBR methods to be widely adopted in the computer-graphics industry their advantages over geometry-based systems must be realized. This will undoubtedly require the development of special-purpose hardware. It is interesting to speculate on the form that such hardware systems might take. Luckily, we can draw upon our more than 25 years of experience building geometry-based computer-graphics systems. There are striking similarities between IBR methods and the traditional rendering pipeline and the texture mapping approaches used today. In the short term, I expect to see evolutionary modifications to existing geometry-based systems that will render them "IBR capable." However, in the longer term we can expect to see far more revolutionary systems that capitalize on the strengths of the IBR method in representing scenes with apparently high geometric complexity.

Finally, it is interesting to speculate on future directions in IBR research. For instance; how do the image-warping and viewing interpolation approaches compare to the light field and lumigraph methods? Can they be viewed as limiting cases within a common framework? Instead, might image warping be considered as a compression method for the database of rays represented in light-fields and lumigraphs? What are the computational implications of these various approaches? Another class of issues revolves around the dependence of IBR methods on difficult computer vision problems such as image-correspondence and camera calibration. We must first ask if this dependence necessary, and how might we successfully avoid these problems? This will lead to the fundamental issue of whether the problem of visualization is fundamentally a metric (and thus geometric) problem.

Jitendra Malik

Image based modeling and rendering has been presented in its purest form in the Light Field and Lumigraph work. In some ways this has very much the flavor of 'What you see is what you get'. The flip side is, of course, 'What you see is all you get'. In order to go further, one needs to recover geometric and reflectance structure from the collection of images. Barrow and Tennenbaum had proposed producing such a factorization—they called the result 'intrinsic images'—as an agenda for computer vision research way back in 1978. To the extent one is able to perform such a factorization, the ability to produce renderings from novel viewpoints and in novel lighting conditions follows directly. Unfortunately, twenty years of computer vision research in this framework has shown the problem to be much harder than originally suspected and fully automated, general purpose solutions are not yet available. I shall argue that the way forward is with hybrid approaches based on partial factorization into geometric and reflectance structure and representation of the remaining information in the form of unfactored image maps. Practical solutions are likely to be domain dependent. Requirements for geometric fidelity and/or photorealism, as well as what is possible given current computer vision technology, will vary according to the application.

Eric Chen

As the Web gradually moves from pure text-based pages toward multimedia enriched sites, the demand for using virtual reality to enhance the Web browsing experience is increasing. This is evident from the popularity VRML has received so far. However, real-time 3D rendering running on a PC typically does not offer very high image quality. The difficulty in creating good 3D content also presents a hurdle for the wide adoption of traditional 3D environment descriptions.

Image based rendering offers an attractive alternative. The use of images and photographs to “model” virtual environment is easier than 3D modeling in most cases. The rendering speed of IBR has weak correlation with scene complexity and is usually fast enough even on low-end PCs. IBR also allows a virtual environment to be transmitted with fairly constant bandwidth and can use standard image compression methods to reduce the data size. IBR is thus more likely to be accepted as the virtual reality method for the Web and the consumer markets.

IBR already has wide applications on the Web. Travel and real estate sites are using panoramic image rendering to create photorealistic location-based browsing. Image based objects (objects represented with images shot from different directions) are used to create interactive product catalogs. The merging of IBR and 3D rendering allows the creation of a virtual environment that is both photorealistic and dynamic.