

# MIT9904-14

## Image Based Synthetic Aperture Rendering

Leonard McMillan and Julie Dorsey

### Project Overview

The goal of the Image Based Synthetic Aperture Rendering Project is to develop new image-based representations for use in computer graphics. Image-based representations have significant advantages over the traditional computer graphics models. They are easy to acquire, efficient to render, and the resulting images look more realistic than traditional computer graphics renderings. In this research effort, we have focused on developing an end-to-end solution for image-based representations. We have developed new algorithms for rendering novel views from a collection of images. We are also developing devices to acquire this collection of images efficiently and economically, and, finally, we have developed new autostereoscopic display devices for directly viewing these image-based representations. Our ultimate goal is to construct a complete capture-to-display system that will acquire, process, and display three-dimensional real-world scenes in real-time. Such a system could be used in a broad range of applications including remote telepresence, studio video production, and three-dimensional television.

### Progress to Date

In the first year of this effort, we have made significant advances in the areas of capture, rendering, and display. The central component of our image-based rendering system is a collection or database of reference images, called a *light field*. We have designed two capture devices for acquiring light fields. Our first capture device is a low-cost portable device for acquiring static scenes. Our second capture system is a considerably more ambitious design for acquiring dynamic light fields at video-rates. It employs a custom-designed random-access camera module, which we plan to replicate and configure into various two-dimensional arrays. To address the problem of synthesizing novel viewpoints from light fields, we have developed new algorithms and built a real-time renderer. In the area of displays, we have built a static light-field display that allows

multiple viewers to perceive a three-dimensional scene, without any need for special viewing or tracking devices.

Our first capture system addresses the high costs and lack of portability of previous light-field acquisition systems. The key component of our low-cost capture device is an inexpensive flatbed scanner with a custom lens system that mounts onto the glass face of the scanner. The construction of such an acquisition device is relatively straightforward. The primary challenges of the project lie in calibrating and correcting the scanned images. A paper on these methods has been submitted to SIGGRAPH 2000 as a technical sketch.

Our low-cost capture device, as well as the more expensive motion-platform based light-field capturing devices, is only able to acquire static scenes. We have begun work on constructing a second capture system that will be suitable for acquiring dynamic scenes in real time. One way to accomplish this would be to assemble a two-dimensional array of video cameras. A usable system would require approximately 256 digital cameras configured in a 16 by 16 array. In order to process these images, data would have to be transferred into the memory of a host processor at a data rate of approximately 7 GB/s. In an effort to reduce this bandwidth requirement, we have designed a random-access imager that can be mapped into the memory space of a host processor. In this architecture, only those pixels that are need to generate a specified novel view need be accessed. We are implementing our random-access imager using a sequential-access color CMOS imager that is connected to a synchronous DRAM and a field programmable gate array (FPGA). The FPGA is the primary control logic, which 1) drives the imager, 2) manages the transfer of image data between the imager and memory, and 3) provides a dual-port access to the SDRAM memory buffer from the host computer. This memory-buffering scheme provides the illusion of random access that is necessary to support our rendering algorithms. To date, the sensor pod FPGA circuitry's design has been created.

To produce high-quality novel views at interactive rates, we have built a rendering system running on Microsoft Windows systems using DirectX 7. This renderer allows the user to create images with dynamically controllable position, orientation, focus, and aperture size.

We have also developed new rendering algorithms that reorganize light fields allowing them to be directly viewed using a simple display device. Multiple viewers will perceive the display as a three-dimensional scene, without any need for special eyeglasses or tracking devices. Using our reparameterization technique, we can capture a scene using

one of our light-field capture devices, reparameterize it, and display the result as a three-dimensional image, much like a holograph. All that is required for our 3-D viewing system is a \$30 acrylic lens array. The resulting images exhibit considerable depth, both in front and behind the actual display surface.

The dynamic rendering techniques developed through our collaboration with NTT will be presented at SIGGRAPH 2000. "Dynamically Reparameterized Light Fields" will appear in the proceedings and will be presented during the conference in July.

### **Expected Progress through June 2000**

For the real-time camera, we plan to complete the printed circuit board layout for the sensor pod, and then fabricate and test a small number of them. We also plan to build a test fixture, which will simulate the interface between the sensor pod and the camera motherboard. This will allow us to test our random-access imager design.

For rendering, we plan on increasing the texture bandwidth of the real-time renderer so that we can render from significantly larger data sets and improving the accuracy of the renderings by using more accurate hardware-accelerated projective texture mapping. Furthermore, we will start design of a new renderer optimized for accessing our camera array (the texture mapping method will not be sufficient for the real-time camera).

Although we have presented prototypes for the autostereoscopic display, we would like to improve on the quality and resolution of the three-dimensional images. In addition, we will be exploring the possibilities of building a video display using our direct light field display technology.

### **Second Year Research Plan**

Our primary goals in the second year of this research effort are to fabricate and demonstrate the real-time camera array for acquiring dynamic scenes, and to couple this system to our real-time rendering system. A second objective is to build a prototype display device by combining our autostereoscopic lens array to a flat-panel display device. Both of these efforts support dynamic light fields.

To fabricate the real-time camera array we will manufacture sixteen random-access sensor pods and a motherboard with an interface to a host PC. As stated in the

previous section, a prototype random-access sensor pod will be tested and debugged during the summer. The motherboard will be designed and fabricated during the early part of the second year. The motherboard will have control logic to control the multiple sensor pod data busses planned for our system, and will connect to the host PC through a PCI card. When this work is completed, we should have a sixteen pod random access camera suitable for small video-rate scenes.

We will also develop a new rendering platform for use with the camera array. Because the current rendering system must preload all available images into graphics hardware texture memory, a new rendering platform will need to be written without assumptions made for static scenes. Note that our current techniques for dynamic reparameterization will remain valid: we will only need to rewrite the renderer to recall pixels from the camera array instead of from texture memory. Once the traditional algorithm has been implemented on the camera, we can experiment with distributed computation to increase the quality of image synthesis while decreasing the amount of data transfer.

Once this system is completed, it should be very easy to capture static light fields and to play video-rate light fields at run time. After the prototype light-field camera is complete, we would like to extend the array to 256 cameras and design a compatible, higher-resolution camera pod. Extending the array should be easy, as the system has been designed to be modular and to support larger array sizes.