

# **Variable Viewpoint Reality**

## **9807-28**

**Progress Report: January 1, 2000 – June 30, 2000**

**Paul Viola and Eric Grimson**

### **Project Overview**

In the foreseeable future, sporting events will be recorded in super high fidelity from hundreds or even thousands of cameras. Currently the nature of television broadcasting demands that only a single viewpoint be shown, at any particular time. This viewpoint is necessarily a compromise and is typically designed to displease the fewest number of viewers.

In this project we are creating a new viewing paradigm that will take advantage of recent and emerging methods in computer vision, virtual reality and computer graphics technology, together with the computational capabilities likely to be available on next generation machines and networks. This new paradigm will allow each viewer the ability to view the field from any arbitrary viewpoint -- from the point of view of the ball headed toward the soccer goal; or from that of the goalie defending the goal; as the quarterback dropping back to pass; or as a hitter waiting for a pitch. In this way, the viewer can observe exactly those portions of the game which most interest him, and from the viewpoint that most interests him (e.g. some fans may want to have the best view of Michael Jordan as he sails toward the basket; others may want to see the world from his point of view).

### **Progress Through June 2000**

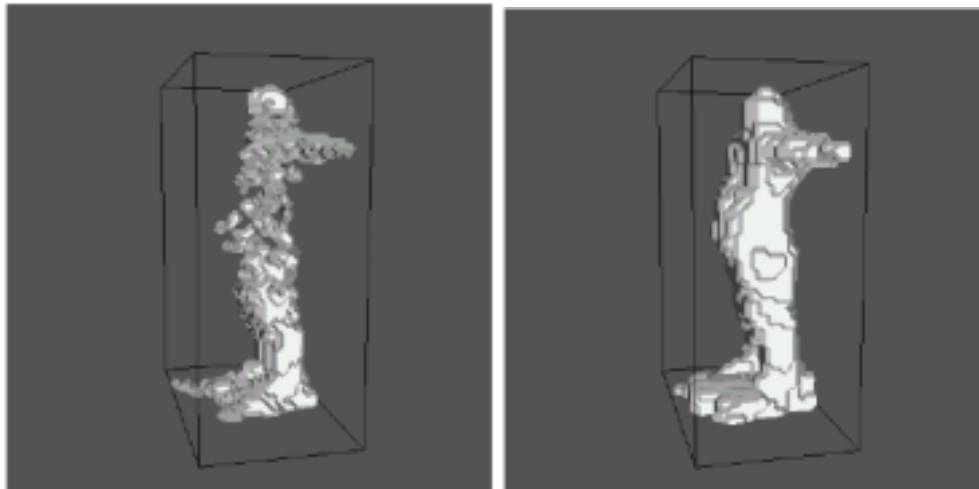
We have made rapid progress on a number of problems related to the goals of the Variable Viewpoint Reality project:

- We have developed a number of basic algorithms for 3D reconstruction. One approach is designed to work in real time on many cameras. Another is a bit slower, but is designed to yield higher quality results. A third attempts to find the arm, leg and body positions of a human being from one or multiple camera views. Each of these algorithms is being tested and improved continually.
- We have designed and setup a multiple camera systems for acquiring data in real-time. This system was designed to be flexible and to work indoors. Right now we have 16 cameras working in synchrony. We would like to setup more.
- We have acquired a great deal of multi-camera data. This is allowing us to test our algorithms and to develop new ideas.

- In collaboration with students working on another project we have been observing outdoor activities. This system provides coarse tracking information of multiple people and cars. The system can also recognize simple activities.
- We have demonstrated the system performing real-time 3D reconstruction using 16 cameras. This system combines many of the results mentioned above.
- We have developed a new algorithm for the reconstruction of 3D shapes. This algorithm addresses one of the key problems we have encountered to date... noise in the camera observations. In previous reconstruction algorithm, each camera attempts to segment the object from the background. These segments are then intersected to form a 3D shape. In some cases noise in the cameras leads to incorrect segmentation. This in turn leads to poor reconstruction. Our new algorithm explicitly models this noise and introduces a prior over shapes. The result is the Bayes optimal reconstruction that is very insensitive to noise. These results are described in a paper at CVPR 2000 which is available from the MIT/NTT web page: <http://www.ai.mit.edu/projects/ntt>.

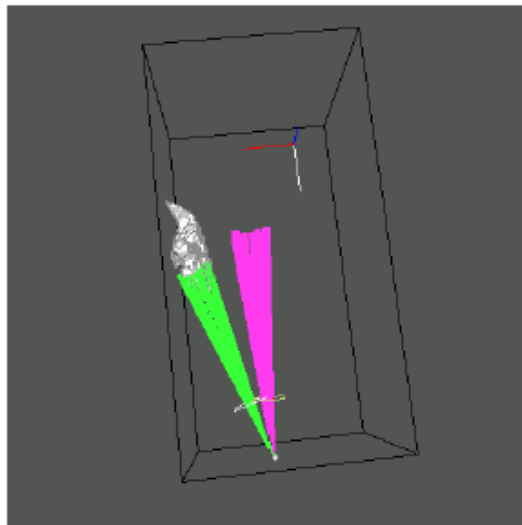


On the left is a typical input scenario. Notice that the imaging conditions are quite difficult, lighting is poorly controlled and the subject is wearing clothing that closely matches the background. On the right is an attempt to segment the subject from the background. These silhouettes are then intersected to in order to reconstruct the 3D shape. The quality of these silhouettes is poor but it is the best that can be done in real-time.



The 3D reconstruction on the left is the intersection of the silhouettes shown above. The holes or gaps result from the gaps in the segmentations. On the right is the output of our new algorithm. With little additional computation the results are significantly better.

- We have developed new algorithms for the automatic calibration of the camera array. Typically, the calibration of 16 cameras is a very difficult and time consuming task. Our approach requires little human intervention and can be used to dynamically update the calibration over time. The algorithm proceeds by a random search over camera poses in an effort to maximize the reconstruction volume.



The significance of correct calibration is shown in this figure. On the left is an overhead view of a 3D object and a single camera. When the camera is correctly calibrated (shown in green) the cone intersects the volume, as is required for correct reconstruction. If the camera were rotated 17 degrees (shown in pink) the cone no longer intersects the volume, and reconstruction is impossible.

## **Research Plan for the Next Six Months**

- Integration of new reconstruction algorithms with real-time system.
- Incorporate the calibration system and the reconstruction system.
- Explore tracking of articulated human body models.
- Incorporating more cameras with real-time system.
- Improving spotting and tracking of people.