

# **High-Resolution Mapping and Modeling of Multi-Floor Architectural Interiors**

**MIT9904-20**

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

**Seth Teller**

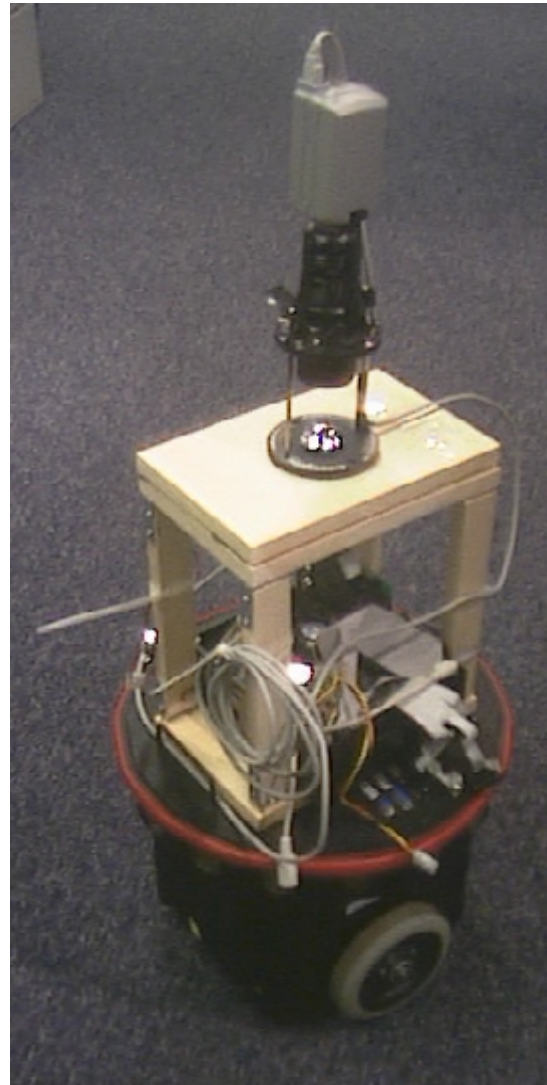
## **Project Overview**

The goal of our project is to achieve an automated mapping capability for architectural interiors. We are deploying a small rolling robot with an omni-directional video camera, an on-board CPU, and wheel encoders for odometry. We intend to have the robot roam about the space to be acquired, gathering video frames, tracking geometric features and constructing a geometric model of its environment on the fly. We also plan to establish a wireless network link between the robot and a base station. The base station will monitor telemetry from the robot, and will serve the results of the robot's mapping efforts through a standard web interface. The base station could also allow semi-autonomous direction of the robot, for example by one or more interested visitors to the web site.

In our initial proposal and previous progress report, we described the equipment we planned to use and our initial mapping efforts. This progress report describes the prototype robot we have constructed, and our initial mapping efforts.

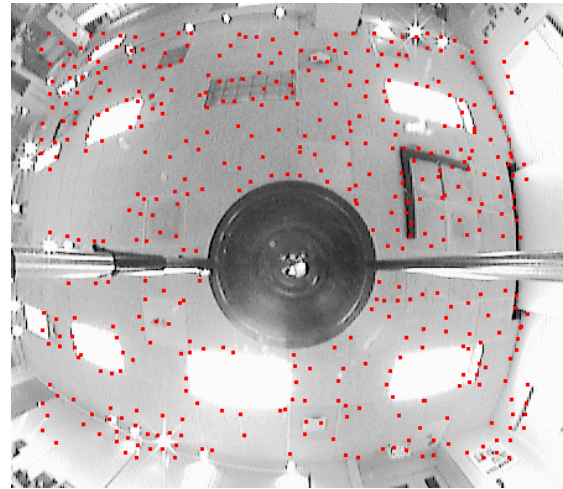
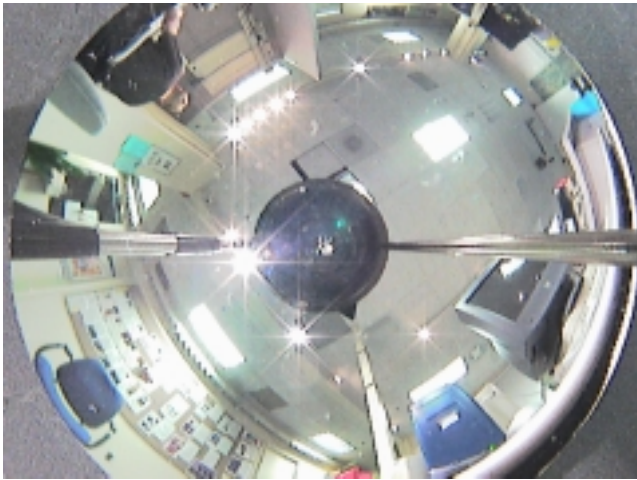
## **Progress Through June 2000**

Our initial prototype was a four-wheeled radio-controlled vehicle. Since the vehicle's wheel mounts did not permit the establishment of reliable odometry, we decided to switch to a custom platform, a two-wheeled rolling robot with integrated wheel encoders (see figures). Prof. John Leonard of the Ocean Engineering Department lent us the Animatics robot.

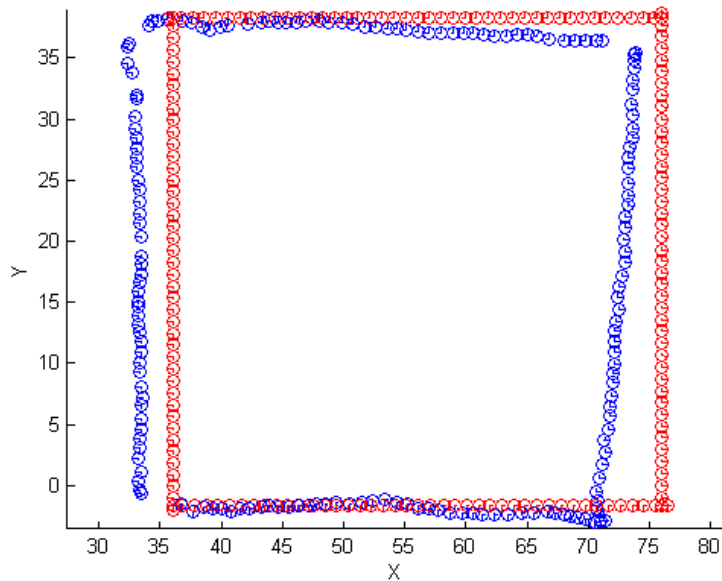


The students Wojciech Matusik, Aaron Isaksen, and Rhys Powell constructed the prototype imaging system and on-board CPU shown here. Two of the students completed the work as a course project in the graduate AI course “Embodied Intelligence”; the third student researcher was funded by NTT. Michael Bosse, another student researcher in the City Scanning project, wrote the modified KLT tracker.

The robot was deployed in a test area measuring about two meters on a side in the Tech Square graphics lab. Using the on-board CPU, the robot was directed to execute a square motion sequence as follows. From the starting point, the robot proceeded counter-clockwise along one edge of the square. After each advance of about two centimeters, the robot stopped and settled, allowing vibratory motions to damp out. An image was then acquired (a typical omni-directional image is shown below at left).



Next, persistent point features were detected using a local gradient operator (point features for one frame are shown above at right). These features were tracked over the sequence using a modified KLT point tracker. The tracker was modified to process omni-directional (rather than rectangular frame) imagery. Finally, a full-path bundle adjustment technique estimated both the 3D positions of the tracked points and the 3D position and orientation of the robot at each frame. The figure below shows the idealized (red) and recovered (blue) paths.



A time-lapse sequence showing the robot's path can be found at:

<http://city.lcs.mit.edu/city/interiors/roverpath.avi>

The acquired omni-directional image sequence can be found at:

<http://city.lcs.mit.edu/city/interiors/roversequence.avi>

## Research Plan for the Next Six Months

We are investigating a number of techniques for the acquisition system over the next 6-12 months. We plan to incorporate robust spherical optical flow algorithms (ego-motion estimation from dense texture), and continue the development of our feature tracking module (ego-motion estimation from corner and edge features). For this purpose we have initiated collaboration with a noted expert in optical flow algorithms, Prof. Michael Black of Brown University.

We plan to investigate the transfer of robust image registration algorithms from the Argus sensor (a high-resolution still image sensor) to the Rover. This involves detection and robust estimation of vanishing points (usually present in architectural scenes) and probabilistic correspondence algorithms for point features.

We plan to scale the acquisition arena up to a full floor of Technology Square. This will require that the robot negotiate long hallways, and over time completely replace any given set of tracked features with a newly visible set. Scaling also requires the abandonment of any batch bundle adjustment technique that assumes availability of the entire image sequence. Instead, our robust ego-motion estimation will use overlapping sets of tracked features throughout the sequence.

Finally, we plan to use model-based knowledge of architectural environments to achieve high-quality, textured geometric models of the acquired spaces. In particular we will enforce assumptions of piecewise planarity, piecewise constant textures or albedos, and generally diffuse environments to achieve robust reconstruction and texturing algorithms.