

# Architectural Interiors MIT9904-20

## Proposal for 1999-2000 Funding

Seth Teller

### Project Overview:

We propose to extend and complement our successful automated mapping system for exterior urban spaces, to incorporate mapping capability for *interior* spaces. We propose to deploy a rolling sensor equipped with laser range-finder, high-resolution digital camera, and positioning instrumentation (wheel encoders and precise inertial measurement system) to allow the acquisition and registration of dense color-range observations of interior environments. The sensor package will be directed by remote radio control, but could someday move autonomously. In either case, its positioning instrumentation will track both horizontal motion (e.g. along hallways and through rooms), and vertical motion (e.g., up and down wheelchair ramps, and while inside elevators).

Our system will be the first capable of producing extended, multi-floor maps (CAD models) of interior spaces. The CAD models produced will consist of high-resolution geometry (synthesized from registered laser range data) and high-resolution and high-dynamic range texture (synthesized from multiple digital photographs and inverse global illumination computations). Applications are numerous, including: civilian and military planning and rehearsal; architectural visualization; and embedding of information spaces.

In Year 1, we request \$250K, to acquire instrumentation (\$100K for a Nomadics rolling robot, a Canon Optura digital video camera, a K2T scanning laser range-finder, GEC Marconi inertial measurement unit, and Pacific Crest radio modems), for student support (\$95K for one PhD student, one MEng, two full-time UROPs), for travel (both foreign and domestic), and other minor expenses such as materials and services and communications. In Year 2, we request large-capacity storage media for the gathered data (\$75K), an increased level of student support (\$108K for one PhD, one MEng, three UROPs), travel (both foreign and domestic), and other minor expenses.

At the end of 1999 we expect to demonstrate the integrated sensor, and data collected from several floors of Technology Square (including LCS and the AI Lab), with geometry and texture data captured to better than 10cm resolution.

## **The City Scanning Project**

The MIT Computer Graphics Group is pursuing an ambitious project to achieve automatic 3D mapping of urban areas -- that is, to acquire 3D textured CAD models of extended urban exteriors, automatically, from large numbers of geo-registered digital images. This is a difficult problem, and our strategy has been to deploy a variety of complementary sensors and computer vision algorithms to infer 3D structure from imagery. So far we have deployed three kinds of sensors: a ground-based digital frame camera, mounted on a pan-tilt head and rolling tripod, which is manually propelled; a ground-based "omni-camera" (hemispherical field of view) mounted on a remote-controlled model car; and an aerial frame camera which is mounted in a low-flying, low stall speed aircraft. Each of these sensors is augmented with some combination of inertial navigation sensors, GPS receivers, magnetometers, and (where applicable) wheel encoders to track the 3D position and orientation (pointing) of the camera in Earth coordinates for each acquired image, as well as the absolute time of each image acquisition.

Our technical strategy has four components. First, we are developing the sensor, described above, which will produce high-resolution digital imagery, with position and orientation estimates that are "good enough" for use. Second, we are developing automated image registration algorithms that produce refined exterior orientation estimates for each image, based on the initial estimates provided by instrumentation. Third, we are developing automated algorithms that extract geometry and texture (appearance information) from large numbers of controlled (oriented) digital images. Finally, we are developing a collection of visualization and interaction tools that allow responsive interaction with the recovered, textured 3D model. For example, we employ efficient spatial database techniques to organize the data for proximity and visibility queries to allow rapid collision detection and shedding of irrelevant model data (e.g., polygons that will contribute nothing to the rendered image).

The project has so far been quite successful. We have achieved an end-to-end system capable of reconstructing a textured, 3D model of Technology Square nearly completely automatically. (The remaining semi-automatic component performs partial image exterior orientation; we expect to achieve complete automation of this component within three months.) The total CPU time required for the process is about one CPU-day on a 200-MHz processor; we are mapping the system onto a 64-processor Linux network, and plan to achieve a 20-minute end-to-end processing cycle, nearly real-time. Please see the URL <http://graphics.lcs.mit.edu/city/city.html> for data, pictures, and videos.

## **Acquiring Architectural Interiors**

The City Scanning Project's major emphasis to date has been the collection of observations corresponding to urban exteriors: essentially, the outer shells of built structures in the urban environment. It is natural to consider extending our acquisition, mapping and modeling system to handle architectural interiors as well. That is, we wish to construct a powerful sensor, and a set of novel algorithms, to allow interior environments to be "scanned" into CAD systems in much the same way documents are acquired by flatbed scanners, or small 3D objects are scanned by CyberWare systems.

The major research challenges lie in: 1) the development of the sensor, so as to acquire color and range observations, along with robust estimates of sensor position and orientation; 2) the robust registration and aggregation of observations taken at different times, from different positions, and under different lighting conditions; and 3) the efficient derivation of 3D CAD models, including geometry and texture (reflectance, appearance) information to represent the acquired environment.

We have gained considerable experience manipulating large numbers of geo-located digital images from our present research. Much of this experience is transferable to the interior problem domain. There are several new technical issues to face, however. For example, we plan to operate a laser range-finder, capable of acquiring several thousand ranges per second, for minutes or hours from a slow-moving, rolling platform. In so doing we will amass tens of millions of range observations, many of them redundant in the sense that they can be predicted from the positions of their neighbors. Standard registration algorithms (such as Iterated Closest Point, or ICP algorithms) work well for hundreds or thousands of points. But to work well in this new context, such algorithms must be mediated by carefully built spatial data structures, to combat the combinatorial explosion that would overwhelm a naïve algorithm.

Another issue is environmental changes over time. Outdoors, buildings and their appearance change rather slowly, due to construction, changes in sun position, season, etc. Indoors, change is much more rapid, due to for example the tendency of building occupants to move furniture, remodel office and common space, et cetera. Thus, more so than outdoors, we must develop a collection of robust algorithms for change detection, and flag spaces in which rapid change is occurring so that they may be observed with higher-than-normal frequency. In this way, we plan to produce high-resolution static models of spaces that are largely static. However, for spaces which change rapidly, or in which there is much activity, we plan to produce a kind of "3D time-lapse" model – a CAD model which is indexed not only spatially (permitting viewing from any angle), but in time as well (so that the user may view the model at any time as well).

Given our extensive experience with the outdoor modeling project, we look forward to bringing up the indoor sensor rather quickly, probably within the first three to six months of the grant period. During the bring-up, and thereafter, we will begin to acquire data. By the end of 1999, we plan to exhibit a functioning, integrated sensor, capable of acquiring image and range data, and moving about, under remote control. We also plan to exhibit a large dataset incorporating color, range, and navigation data from several floors of Technology Square, including portions of LCS and the AI Lab.