*The 4th International Conference on Field and Service Robotics, July 14–16, 2003*

# Scan Alignment and 3-D Surface Modeling with a Helicopter Platform

**Sebastian Thrun**
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA
thrun@cs.cmu.edy

**Mark Diel**
Whirled Air
Helicopters
Menlo Park, CA
mark@whirledair.com

**Dirk Hähnel**
Computer Science Department
University of Freiburg
Freiburg, Germany
haehnel@informatik.uni-freiburg.de

**Abstract**

*This paper describes initial results for a laser-based aerial mapping system. Our approach applies a real-time laser scan matching algorithm to 2-D range data acquired by a remotely controlled helicopter. Results obtain for urban and natural terrain exhibit an unprecedented level of spatial detail in the resulting 3-D maps.*

## 1 Introduction

In recent years, a number of research teams have developed robotic systems for mapping indoor [12, 20] and outdoor environments [9]. Since sensing is usually confined to the immediate vicinity of the vehicle, active sensors such as sonars and laser range finders have become the technology of choice [19]—albeit some notable exception using passive cameras [16]. For the problem of acquiring accurate maps of outdoor terrain, ground vehicles are limited in two aspects: First, the ground has to be traversable by the vehicle itself. Many environments are cluttered with obstacles that are difficult to negotiate [4]. Second, all important features in the environment have to be perceivable from relatively low vantage points. Moreover, the set of vantage points that can be attained usually lie on an approximate 2-D manifold parallel to the ground surface, since most ground vehicle cannot vary the height of their sensors. This is a severe limitation that is particularly troublesome in complex, natural terrain.

In complimentary research, there exists a huge body of literature on high aerial and satellite-based mapping (see e.g., [2, 6, 11]). At high altitude, it is usually impossible to deploy active range sensors; instead, these techniques are commonly based on passive computer vision systems. While traversability is not an issue for high aerial vehicles, the relatively high vantage points makes it difficult to map vertical structures, and it limits the resolution at which maps can be acquired. Furthermore, clouds can cause obstruction or cast shadows in the imagery. And while air vehicles can change altitude and are therefore not subject to the 2-D manifold constraint characteristic of ground vehicles, such changes have next-to-zero effect on the visual appearance of the surface structure.

Low-flying air vehicles, such as helicopters, promise to overcome these limitations: they are much less constraint than ground vehicles with regards to their navigational capabilities, yet they can fly low enough to acquire data of vertical structures at high resolution. A seminal system by Miller et al. [15, 14] has demonstrated the feasibility of acquiring high-resolution ground models using active laser range sensors on a low-flying helicopter platform.

This paper describes a similar system for acquiring high-resolution 3-D models of urban and ground structures. The system, shown in Figure 1, is based on a Bergen Industrial Twin helicopter, equipped with a 2-D SICK range finder and a suite of other sensors for position estimation. Figure 2 shows some of the computer equipment mounted on the vehicle. The 2-D range finder provides the vehicle with 2-D range slices at a rate of 75Hz, oriented roughly perpendicular to the robot's flight direction. The helicopter is flown under manual control.

Building 3-D maps with centimeter resolution is difficult primarily of two reasons:

1. Using GPS and other proprioceptive sensors, the location and of the sensor platform can only be determined up to several centimeters accuracy. Similar limitations apply to the estimation of its angular orientation. The ground position error induced by those angular errors are often in the order of several meters. This is illustrated by the ground model shown in Figure 3. Shown there is a 3-D model reconstructed using the vehicle's best estimates of its position and orientation. These limitations mandate the use of the range finder as a means for further reducing the errors in the vehicle pose estimate, and to jointly solve the environment mapping the vehicle localization problem.

2. The rich literature on *simultaneous localization and mapping*, or SLAM [7, 12], has developed extensive techniques for simultaneously estimating the structure of the environment (3-D model, or map) and the pose of the vehicle. Range scan alignment techniques (applied in 2-D in robotics[13] and in 3-D in object modeling [17]) overcome this problem by cross-registering multiple scans. To do so, these techniques rely on multiple sightings of the same environmental feature. This is not the case for

**Figure 1**: Instrumented helicopter platform: The system is based on the Bergen Industrial Twin, with a modified SICK LMS laser range finder, a Crossbow IMU, a Honeywell 3-D compass, a Garmin GPS, and a Nikon D100 digital camera. The system is equipped with onboard data collection and processing capabilities and a wireless digital link to the ground station.

**Figure 2**: Some of the electronics onboard the helicopter: An Intel Stayton board with a 400Mhz XScale processor interfaces to the SICK LMS laser via a high speed RS422 serial link, and to all other devices (compass, GPS, IMU) via RS232. The communication to the ground is established via a 802.11b wireless link.

a 2-D sensor that is moved through the environment in a direction perpendicular to its own perceptive field: Here consecutive measurements always correspond to different things in the world.

We have developed a probabilistic SLAM algorithm that addresses both of these problems. Our approach acquires 3-D models from 2-D scan data, GPS, and compass measurements. The algorithm exploits a local smoothness assumption for the surface that is being modeled, but simultaneously allows for the possibility of large discontinuities of the mapped structure. By doing so, it can utilize range scans for vehicle localization, and thereby simultaneously improve both the pose estimate of the helicopter and the accuracy of the resulting 3-D environment model.

We believe that the maps acquired by our system are significantly more accurate and spatially consistent than previous maps acquired by helicopter systems. A key reason of this increase in accuracy comes from the fact that scans are used for the pose estimate of the vehicle's sensor platform.

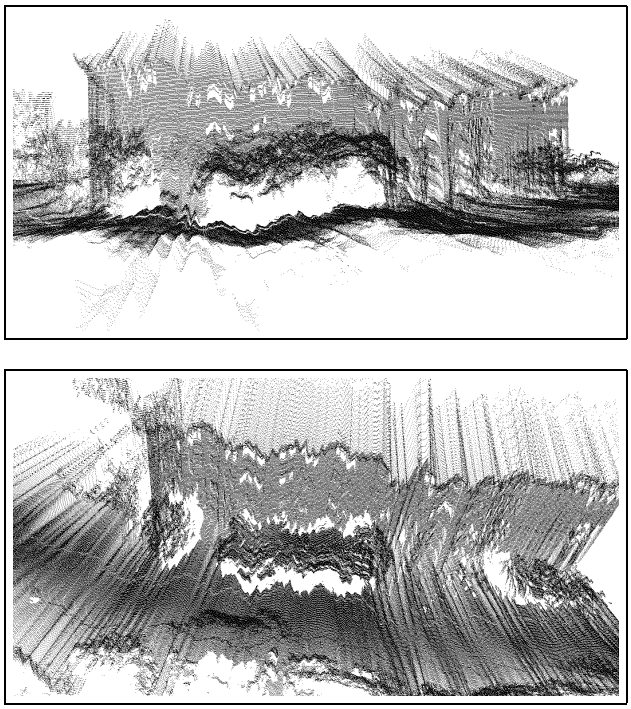

**Figure 3**: Raw data of a multi-storey building

## 2 3-D Modeling Approach

### 2.1 Vehicle Model

Let $x_t$ denote the *pose* of the sensor's local coordinate system at time $t$, relative to a global coordinate system of the 3-D model. This pose is specified by the three Cartesian coordinates and the three Euler angles (roll, pitch, yaw). In irregular intervals, we receive GPS and compass measurements for the pose, denoted by $y_t$. The probability of measuring $y_t$ if the correct pose is $x_t$ is Gaussian and denoted $p(y_t \mid x_t)$:

$$p(y_t \mid x_t) \quad \propto \quad \exp -\tfrac{1}{2}\,(y_t - x_t)^T\, A^{-1}\,(y_t - x_t) \quad (1)$$

Here $A$ is the measurement covariance. Since all of these sensors are subject to *systematic error* (e.g., drift), we also employ a differential model $p(\Delta y_t \mid x_t, x_{t-1})$ where $\Delta y_t = y_t - y_{t-1}$ is the differential measurement (angles are truncated in this subtraction).

$$p(\Delta y_t \mid x_t, x_{t-1}) \quad \propto \quad \exp -\tfrac{1}{2}\,(\Delta y_t - \delta(x_t, x_{t-1}))^T \; D^{-1}\,(\Delta y_t - \delta(x_t, x_{t-1})) \quad (2)$$

Here $\delta$ calculates the pose difference. The matrix $D$ is the covariance of the differential measurement noise, whose determinant is smaller than that of $A$. This model is thus implemented by a narrower Gaussian, accounting for the fact that relative information is more accurate than absolute information. However, measurements $y_t$ alone are insufficient for ground mapping as discussed above.

**Figure 4**: Helicopter flying by a building under manual remote control. The image also shows the pilot.

### 2.2 Range Sensor Model

To localize the sensor based on range data, we need a model of the range sensor. Most SLAM algorithms model the probability $p(z_t \mid m, x_t)$ of a measurement $z_t$, given the map $m$ and the pose $x_t$. Such a generative model is the most general approach to robotic mapping; however, it involves as many variables as there are features in the map $m$; thus the resulting likelihood functions would be difficult to optimize in real-time.

For a forward-flying helicopter which never traverses the same location twice, it is sufficient to model a relative probability of subsequent measurements conditioned on the pose: $p(z_t \mid x_t, x_{t-1}, z_{t-1})$. This probability models the spatial consistency of scan $z_t$ relative to the previous scans, $z_{t-1}$, assuming that those scans are taken at the global poses $x_t$ and $x_{t-1}$, respectively.

In our implementation, this model is defined as follows:

$$\begin{aligned} & p(z_t \mid x_t, x_{t-1}, z_{t-1}) \\ & \quad \propto \prod_i \exp -\tfrac{1}{2} \ \min \Big[ \alpha, \min_j \Big\{ (z_t^i - f(z_{t-1}^j, x_{t-1}, x_t))^T \\ & \qquad B^{-1} \ (z_t^i - f(z_{t-1}^j, x_{t-1}, x_t)) \ \Big\} \Big] \end{aligned} \tag{3}$$

The index $i$ iterates over individual beams in the scan $z_t$, and the function $f$ projects measurements in scan $j$ into the local coordinate system of scan $i$. The argument of the exponential involves two minimizations. The inner minimization identifies the index $j$ of the measurement nearest to $z_t^i$, when projected into the local coordinate system of the scan $z_t$. The second minimization upper bounds this distance by $\alpha$. It is best thought of as an outlier detection mechanism.

The measurement covariance $B$ is degenerate: it possesses infinite covariance in the direction of helicopter flight. This degeneracy accounts for the fact that subsequent scans carry no information about the forward motion of the vehicle. The degeneracy implies that the rank of $B^{-1}$ is five, even though the matrix is six-dimensional.

### 2.3 Optimization

The resulting probabilistic model is proportional to the product

$$p(y_t \mid x_t) \ p(\Delta y_t \mid x_t, x_{t-1}) \ p(z_t \mid x_t, x_{t-1}, z_{t-1}) \tag{4}$$

The negative log-likelihood is now given by the following expression:

$$\begin{aligned} & \text{const.} + \tfrac{1}{2} \Big( (y_t - x_t)^T \ A^{-1} \ (y_t - x_t) \\ & \quad + \ (\Delta y_t - \delta(x_t, x_{t-1}))^T D^{-1} \ (\Delta y_t - \delta(x_t, x_{t-1})) \\ & \quad + \ \sum_i \min \Big[ \alpha, \min_j \Big\{ (z_t^i - f(z_{t-1}^j, x_{t-1}, x_t))^T \\ & \qquad B^{-1} \ (z_t^i - f(z_{t-1}^j, x_{t-1}, x_t)) \ \Big\} \Big] \Big) \end{aligned} \tag{5}$$

The map and the pose is recovered by minimizing this negative log-likelihood. We note that this problem combines a continuous optimization problem that is quadratic in nonlinear functions of the pose $x_t$, with a discrete one that involves finding the corresponding $z_{t-1}^j$ for each measurement beam $z_t^i$. Following mainstream optimization in the scan matching literature [3], our approach iterates a step in which the minimization is solved for a fix pose $x_t$, which a step in which the optimal pose is determined for a fixed setting of the minimizing indices. Both steps can be carried out highly efficiently [10].

The result is an algorithm that implements the optimization in an incremental fashion, in which the pose at time $t$ is calculated from the pose at time $t-1$ under incorporation of all scan measurements. While such an implementation is subject to local minima, it can be performed in real-time and works well, as long as the helicopter never traverses the same area twice. The 3-D model is also obtained in real-time, by using the corrected pose estimates to project measurements into 3-D space. The model may simply be represented by a collection of scan points in 3-D, or a list of polygons defined through sets of nearby scan points. Both are computed in real-time.

## 3 Results

We have tested our approach in a number of different environments, all of which involved significant vertical structure that cannot easily be mapped by high-aerial vehicles. Figure 4 depicts the helicopter flying by a multi-storey building under manual control; it also depicts the pilot walking behind the vehicle. The raw data acquired in this flight is shown in Figure 3; this plot uses the helicopter's best estimate of its own pose for generating the map. These plots clearly show significant error, caused by a lack of accurate pose estimation.

**Figure 5**: Visualization of the mapping process, carried out in real-time. This figure shows a sequence of snapshots taken from the interactive ground display, which displays the most recent scans with less than 0.5 seconds latency.

Figure 5 depicts a sequence of maps as they are being generated in real-time. The latency between the data acquisition and the display on the ground is less than half a second. Snapshots of the final 3-D map are shown in Figure 6.

Figure 7 shows to maps acquired at a different urban site (left image) and at an ocean cliff (right image). These maps are represented by polygons that are defined over nearby points. The search for such polygons is also carried out in real-time, although rendering the final model requires substantially more time than rendering the multi-point model.

Unfortunately, we do not possess ground truth information for the mapped structures. This makes it impossible to assess the accuracy of the resulting models. However, the models appear to be visually accurate, locally consistent. The spatial resolution of these models in the centimeter range.

## 4 Conclusion

This paper described initial results for an instrumented helicopter platform for 3-D ground modeling. A real-time algorithm was developed that integrates pose estimates from multiple sensors with range data, acquired by a 2-D laser range finders oriented perpendicular to the vehicle's flight direction. The algorithm uses a fast optimization technique to generate maps in real-time. Relative to prior work in [15, 14], our approach utilizes the range data for vehicle localization, which results in maps that are spatially significantly more consistent and—we suspect—more accurate. Experimental results suggest that the maps acquired by our approach are of unprecedented detail and accuracy; however, the exact accuracy is presently not known. Nevertheless, we believe that the findings in this work are highly promising.

It is important to notice that this paper does not address the popular topic of autonomous helicopter control, see [1, 5, 8, 18] for recent work in this area. However, an integration of accurate mapping and autonomous flight would make

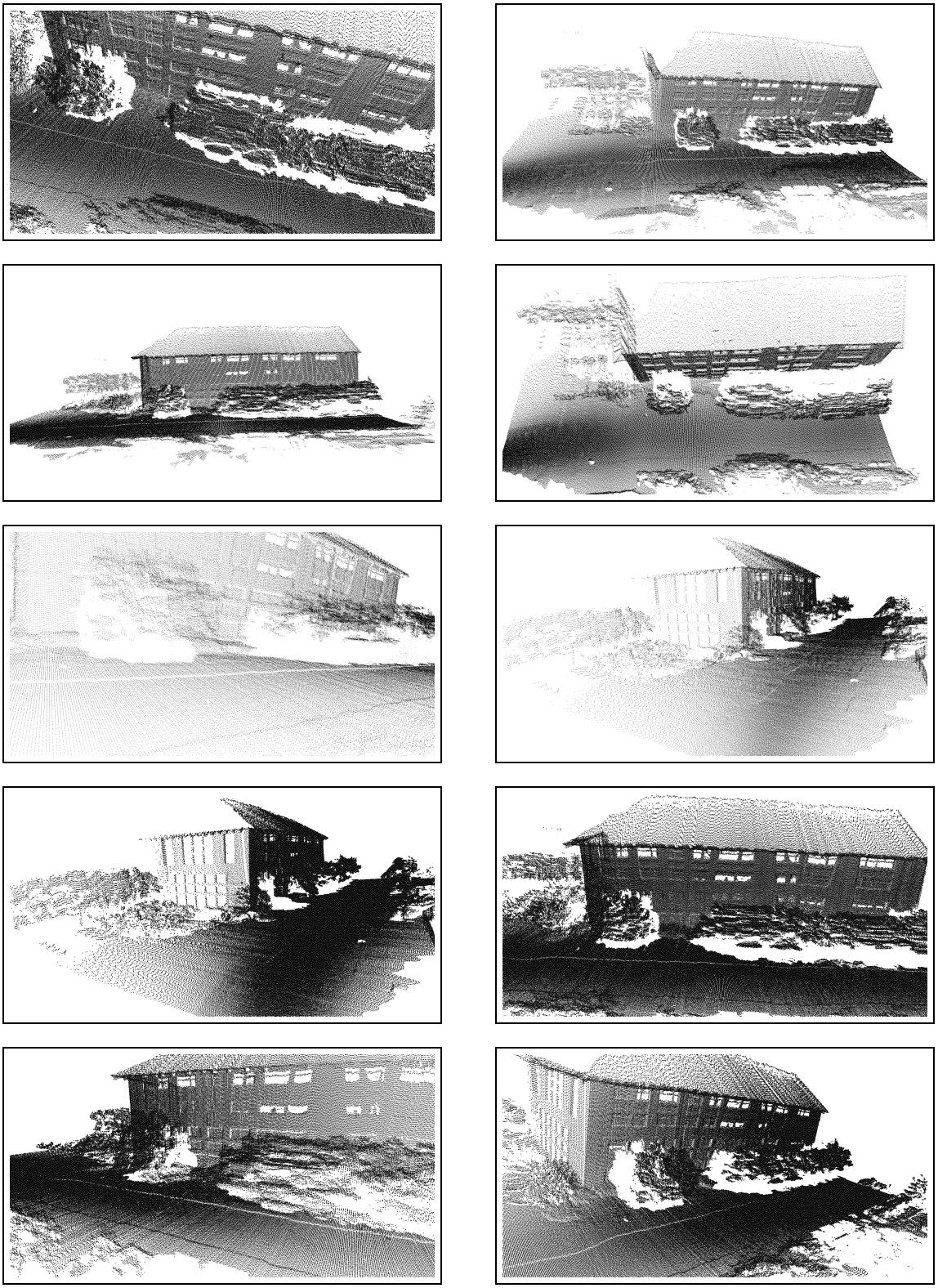

**Figure 6**: Snapshots of the 3-D Map of the building shown in Figure 4. The map is represented as a VRML file and can be displayed using standards visualization tools.

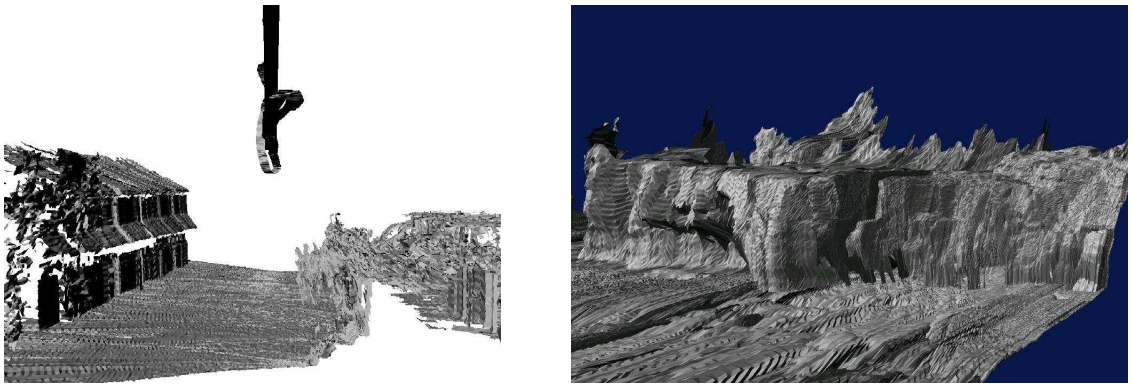

**Figure 7**: Multi-polygonal 3-D models of a different urban site with a smaller building, and a cliff at the pacific coast near Davenport, CA. The left diagram also shows the vehicle's estimated path.

it possible to operate autonomous helicopters in rugged terrain, such as mountainous areas or caves. It would also open the door to the important problem of selecting safe landing pads in uneven terrain.

A second extension concerns the limitation that surfaces are only sensed once. To enable the vehicle to integrate data from multiple fly-overs would require a mechanism for establishing correspondence to previously mapped terrain. While a number of techniques exist that offer this capability [19], it is unclear whether they can be executed in real-time on a helicopter platform.

### Acknowledgments

The authors gratefully acknowledge the donation of two Stayton boards by Intel.

This research is sponsored by by DARPA's MARS Program (contracts N66001-01-C-6018 and NBCH1020014), which is also gratefully acknowledged. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing official policies or endorsements, either expressed or implied, of the United States Government or any of the sponsoring institutions.

## References

[1] J. Bagnell and J. Schneider. Autonomous helicopter control using reinforcement learning policy search methods. In *Proceedings of the International Conference on Robotics and Automation 2001*. IEEE, May 2001.

[2] S. Becker and M. Bove. Semiautomatic 3-D model extraction from uncalibrated 2-D camera views. In *Proc. of the SPIE Symposium on Electronic Imaging, San Jose*, 1995.

[3] P. Besl and N. McKay. A method for registration of 3d shapes. *Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[4] J. Casper. Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center. Master's thesis, Computer Science and Engineering, University of South Florida, Tampa, FL, 2002.

[5] D. Drinka D. Coleman, D. Creel. Implementation of an autonomous aerial reconnaissance system. GeorgaTech Entry into the 2002 International Aerial Robotics Competition, 2002.

[6] P.E. Debevec, C.J. Taylor, and J. Malik. Modeling and rendering architecture from photographs. In *Proc. of the 23rd International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 1996.

[7] G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localisation and map building (SLAM) problem. *IEEE Transactions of Robotics and Automation*, 2001. In Press.

[8] J. Dittrich and E.N. Johnson. Multi-sensor navigation system for an autonomous helicopter. In *Proceedings of the 21st Digital Avionics Systems Conference*, 2002.

[9] J. Guivant and E. Nebot. Optimization of the simultaneous localization and map building algorithm for real time implementation. *IEEE Transactions of Robotics and Automation*, May 2001. In press.

[10] D. Hähnel, D. Schulz, and W. Burgard. Map building with mobile robots in populated environments. In *Proceedings of the Conference on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland, 2002.

[11] G. Konecny. *Geoinformation: Remote Sensing, Photogrammetry and Geographical Information Systems*. Taylor & Francis, 2002.

[12] J. Leonard, J.D. Tardós, S. Thrun, and H. Choset, editors. *Workshop Notes of the ICRA Workshop on Concurrent Mapping and Localization for Autonomous Mobile Robots (W4)*. ICRA Conference, Washington, DC, 2002.

[13] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4:333–349, 1997.

[14] R. Miller. *A 3-D color Terrain Modeling System For Small Autonomous Helicopters*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, 2002. Technical Report CMU-RI-TR-02-07.

[15] R. Miller and O. Amidi. 3-D site mapping with the CMU autonomous helicopter. In *Proceedings of the 5th International Conference on Intelligent Autonomous Systems*, Sapporo, Japan, 1998.

[16] D. Murray and J. Little. Interpreting stereo vision for a mobile robot. *Autonomous Robots*, 2001. To Appear.

[17] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, Quebec City, Canada, 2001. IEEEComputer Society.

[18] S. Saripalli, J.F. Montgomery, and G.S. Sukhatme. Visually-guided landing of an autonomous aerial vehicle. *IEEE Transactions on Robotics and Automation*, 2002.

[19] S. Thrun. Robotic mapping: A survey. In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann, 2002. to appear.

[20] S. Thrun, C. Martin, Y. Liu, D. Hähnel, R. Emery-Montemerlo, D. Chakrabarti, and W. Burgard. A real-time expectation maximization algorithm for acquiring multi-planar maps of indoor environments with mobile robots. *IEEE Transactions on Robotics and Automation*, 2003. To Appear.