# Towards Particle Filter SLAM with Three Dimensional Evidence Grids in a Flooded Subterranean Environment

Nathaniel Fairfield, George Kantor, David Wettergreen
Robotics Institute, Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
Email: {than, kantor, dsw}@cmu.edu

***Abstract*— This paper describes the application of a Rao-Blackwellized Particle Filter to the problem of simultaneous localization and mapping onboard a hovering autonomous underwater vehicle. This vehicle, called DEPTHX, will be equipped with a large array of pencil-beam sonars for mapping, and will autonomously explore a system of flooded tunnels associated with the Zacatón sinkhole in Tamaulipas, Mexico. Due to the three-dimensional nature of the tunnels, the Particle Filter must use a three dimensional map. We describe an extension of traditional two dimensional evidence grids to three dimensions. In May 2005, we collected a sonar data set in Zacatón. We present successful SLAM results using both the real-world data and simulated data.**

## I. Introduction

An Autonomous Underwater Vehicle (AUV) uses a combination of depth sensors, inertial sensors, and Doppler velocity logs (DVLs) to compute a dead-reckoned estimate of its position. With high accuracy attitude and depth sensors the only uncertainty in the AUV's 3D pose ($x$,$y$,$z$,roll,pitch,yaw) is in $x$ and $y$. Most underwater vehicle localization strategies are geared toward open-water situations where the ocean floor is feature-sparse, and where beacon-based approaches or periodic surfacing for GPS fixes is a reasonable solution. The vehicle relies on these external position aides to correct the inevitable integration error in its dead-reckoning solution. Simultaneous localization and mapping (SLAM) offers an attractive alternative method to bound the dead-reckoning error because it allows the vehicle to be completely self-contained and autonomous.

SLAM strategies depend on the extraction of features from sensor data, however in the underwater domain there are not many sensors which are capable of providing the resolution necessary to resolve and recognize features. There has been work with using tunnel cross-sections, or slide images, which can be derived from sparse sonar ranges [1] as long as the environment is tunnel-shaped. In the case where there are free floating artificial features, scanning sonars have been shown to have high enough resolution to support feature-based SLAM [2]. Alternatively, in clear water with good lighting, SLAM has been demonstrated via video mosaicing [3] and also a combination of vision-based feature detection and sonar [4]. However, many underwater environments are characterized by large monotonous regions, where there has been promising work with Synthetic Aperture Sonar (SAS) to support range-and-bearing SLAM [5]. Finally, there is the mine-mapping work by [6] done in similarly complex tunnels, although with a planar floor plan, using scan matching with a laser range finder to recover the 2D vehicle pose from which the 3D map is reconstructed in a post-processing step.

Fig. 1. The DropSonde being lowered into Zacatón.

Our application platform is an autonomous hovering vehicle, called DEPTHX, which will explore flooded *cenotes* (sinkholes) [7]. The mapping system for DEPTHX is an array of pencil-beam sonars that provides a constellation of ranges around the vehicle, but lacks the resolution, update rate, and point density of a laser scanner, making feature detection difficult. The deeper portions of the sinkholes are completely unexplored and may have unexpected geometries, which makes it even more difficult to design feature detectors. For these reasons we have selected a very data-driven representation, evidence grids, as our basic world representation [8]. In the 3D evidence grid representation, space is divided into a grid of cubic volume elements, or voxels, which contain the occupancy evidence inferred from sensors. While 2D evidence grid-based SLAM is well established in the indoor mobile robot domain, it has been found to have limited applicability in truly 3D environments – largely because the 2D map simplification is only suitable in "two and a half"

dimensional environments (such as indoors). The difficulty in generalizing the evidence grid approach to full 3D grows from the multiplication of the computational cost of accessing, modifying, and storing the map due to the third dimension.

Due to the uncertainties in sonar measurements and vehicle motion models, we chose to address the SLAM problem using a Bayesian framework. Particle filters (PFs), provide a proven implementation of Bayesian filtering for systems whose belief state, process noise, and sensor noise are modeled by arbitrary probability density functions (PDFs) [9]. Rao-Blackwellized particle filters (RBPFs), in which each particle contains both a position and a map, have proven an effective way to do SLAM with evidence grids and provide the basis for the approach taken here [10].

In Section II we describe the May 2005 field expedition to Zacatón, as well as some of its salient features. Sections III and IV provide a brief description of 3D evidence grids and our implementation of an RBPF for SLAM. The next section describes experimental results from the Zacatón sonar data and a simulated tunnel system (Section V). Finally, we close with future work and conclusions.

## II. Exploring Zacatón

We have discovered that the Zacatón *cenote* in Tamaulipas, Mexico, is roughly a truncated cone 110m wide, and at least 300m deep (see Figures 1,2). Prior to the May 2005 DEPTHX expedition it was unmapped except for wire drops to establish the depth. Zacatón is the deepest of a series of hydrothermal *cenotes* in the region. It produces microbial mats in the photic zone and has exotic geochemical features which make it an excellent match for the exploration and sampling mission of DEPTHX.

The DEPTHX team used a sonar test platform (see Figure 3), called the DropSonde, to build a preliminary map of the first 200m of Zacatón. In addition to a data logging computer, the onboard sensors include 32 spirally arranged sonars, a Honeywell HG2001AC RLG IMU, and three depth sensors. The sonars were fired at about 2Hz, and have a 10cm precision. The IMU has about a 1m/s accelerometer drift rate, which precludes its use as a dead-reckoning navigation system, but only a 0.1 degree/hour gyro drift rate, giving use excellent heading and attitude information. The depth sensors had individual accuracies of about 10cm.

During a week-long field expedition in May 2005, the DEPTHX team lowered the DropSonde progressively down to 200m, the maximum depth rating for several of the components. The DropSonde was lowered on a winch from a stabilized barge, and the locations of the seven “drops” were recorded with surveying equipment, which together with the data from the onboard sensors allowed the sonar data to be registered. By spinning the DropSonde during its slow ascent and descent, we acquired excellent sonar coverage of the walls. For the first 250m, the *cenote* is remarkably smooth and cylindrical, although there are several shelves (see Figure 2).

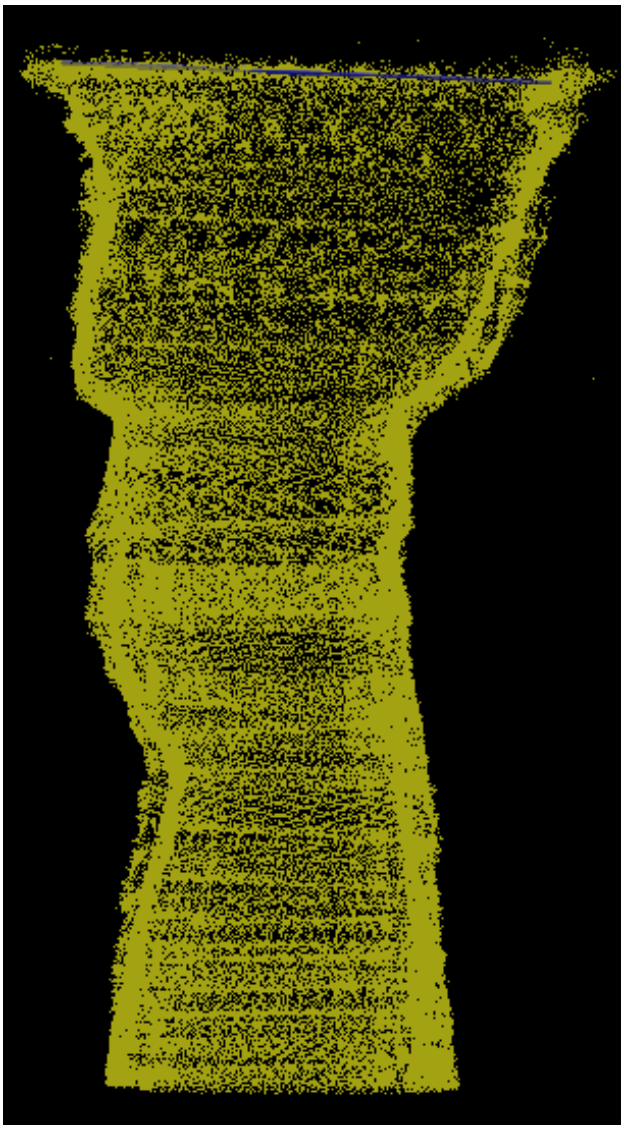

Fig. 2. An north-facing side view of the first 200m of Zacatón, plotted as a sonar point cloud in orthographic projection.

## III. 3D Evidence Grids

A 3D evidence grid is based on the same set of concepts and Bayesian update equations as the classic 2D evidence grids which have been so well described in the literature (see [8] [11]). Each 3D cell, or voxel, in the evidence grid represents the probability that the corresponding volume in the real world is occupied. As measurements are collected, the evidence they provide about the occupancy of each voxel is entered into the map. A sonar beam model defines how a single range measurement can be inserted into the evidence grid. There are several methods which can be used to construct a beam model, including deriving it from physical first principles [12] or learning it [8]. We chose to use the simplest reasonable approximation – a cone with a cap (see Figure 4). Likewise, the simplest method to query a sonar range from the 3D evidence grid is to trace a single ray from the origin of

Fig. 3. The DropSonde – the three main pressure vessels, from top to bottom, contain the IMU, the batteries, and the sonar system.

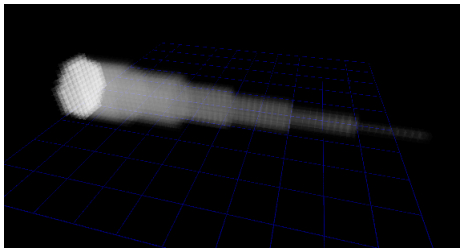

Fig. 4. An example of the conic sonar beam model, with a 6 degree beam-width.

the sonar query until some threshold or other terminating condition is met.

Evidence grids have strong independence assumptions: each voxel is independent of its neighbors, and every sonar measurement is independent of all other measurements. While these assumptions are clearly an over-simplification, they are precisely what allow the map to be updated using Bayesian update rules [8] (see the next section for details).

## IV. Particle Filter SLAM

A PF approximates an unknown distribution with a cloud of particles. In order to achieve this approximation, the particles must be spread over the dimensions of the distribution. In the SLAM case, this unknown distribution encompasses both the particle pose $s_t^m$ (with $m = 0 : N_{\text{particles}}$) at time $t$ and the map $\Theta^m$, conditioned on the range measurements ${}^n r_{0:t}$ (with $n = 0 : N_{\text{sonars}}$) and the vehicle dead-reckoned position innovation $u_{0:t}$ (for an overview of notation see Table I):

$$p(s_t, \Theta | r_{0:t}, u_{0:t})$$

The key insight the RBPF is that it is possible to factor this distribution into two parts and analytically marginalizing out one of the parts, conditioned on the other part, thereby reducing the number of dimensions that have to be covered by the particles [10]. Applying this insight to SLAM requires marginalizing the map distribution $p(\Theta | s_{0:t}, r_{0:t}, u_{0:t})$ (now also conditioned on $s_{0:t}$), and using particles to estimate a posterior over vehicle trajectories $p(s_{0:t} | r_{0:t}, u_{0:t})$[1], in order to compute

$$p(s_t, \Theta | R_{0:t}, u_{0:t}) = p(\Theta | s_{0:t}, R_{0:t}, u_{0:t}) p(s_{0:t} | R_{0:t}, u_{0:t}).$$

As described in the previous section, evidence grids offer a method for analytically estimating $p(\Theta | s_{0:t}, r_{0:t}, u_{0:t})$. Since it is unrealistic to recompute the evidence grid $\Theta$ at each timestep, in actual implementation each particle consists of a position $s_t$ and a map $\Theta$, which is kept up to date at each timestep.

Our particle filter implementation has the following steps, closely following the Sequential Importance Sampling with Resampling filter [13]:

- **Predict:** Particle positions $s_t$ are predicted using the vehicle motion model $h(s_{t-1}, u_t)$. The particles are now an approximation of the prior PDF of the vehicle position, which is why this step is often called Sampling.
- **Weight:** The sonar measurements $r_t$ and the particle map $\Theta$ are used to compute the weight $w$ for each particle. The weights of the particles now approximate the importance PDF of the vehicle position.
- **Resample:** These two distributions (prior PDF and importance PDF) are merged to yield the posterior PDF of the vehicle position, by resampling the particles according to their weights.
- **Update:** The particle maps $\Theta$ are updated using the sonar ranges $r$, and the particle positions $p_t$.
- **Estimate:** A position estimate is computed from the particle cloud.

### A. Predict

The dead-reckoned position innovation $u_t$ is computed using the navigation sensors (IMU, DVL and depth sensor). A new position $s_t$ is predicted for each particle using the vehicle motion model:

$$s_t = h(s_{t-1}, u_t, N(0, \sigma_s)$$

which includes a zero mean noise model with standard deviation $\sigma_s$. The distribution of particles is now an approximation of the prior PDF of the vehicle position.

### B. Weight

The weight for each particle is computed by

$$w = \eta \prod_{n=1}^{N_{\text{sonar}}} p({}^n r_t | s_t, \Theta),$$

where $\eta$ is a normalizing factor. Since the simulated ranges $\hat{r}$ are generated by the measurement model

$$\hat{r} = g(s_t, u_t, N(0, \sigma_r)),$$

[1]In the DEPTHX application, the particles are primarily distributed over $x$ and $y$, due to the very low error in the IMU heading and attitude, as well as in the depth sensors.

which has a normal noise model,

$$p(r|s,\Theta) = \frac{1}{\sqrt{2\pi\sigma_r^2}} e^{\frac{-(\hat{r}-r)^2}{2\sigma_r^2}}.$$

Taking the logarithm of both sides shows that maximizing this weight metric is very close to minimizing the intuitive sum squared error metric:

$$\log w = N_{sonar} \log\left(\sqrt{2\pi\sigma^2}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{N_{sonar}} (\hat{r}^i - r)^2.$$

### C. Resample

The $O(n)$ algorithm described in [9] is used to resample the particle set $S$ according to the weights $w$:

$$S_t = \text{resample}(S_{t-1}, w)$$

In actual implementation, much of the computational cost of the algorithm is in simply copying the particle maps during the resampling step.

### D. Update

The particle maps $\Theta^m$ are updated by using the standard log-likelihood evidence grid update equation [8] to update the evidence of all the voxels $\theta = {}^{ijk}\Theta$ which lie in the conic sonar beam model of each ${}^{n}r$:

$$\begin{aligned}\log(\theta) &= \log\left(\frac{p(\theta|r)}{1-p(\theta|r)}\right) \\ &\quad + \log\left(\frac{1-p(\theta)}{p(\theta)}\right) + \log(\theta)\end{aligned}$$

The map is initialized using

$$\log(\theta) = \log\left(\frac{p(\theta)}{1-p(\theta)}\right).$$

If the prior $p({}^{ijk}\theta) = 0.5$, the second term is zero, and the initialization simply sets all voxels to zero. The remaining term involving $p({}^{ijk}\theta|r_t)$ is provided by the sonar beam model, generally called the inverse sensor model [14]. The update equation for each voxel can be reduced to simply summing the value of the sonar model with current voxel evidence. Using matrix transformations for each voxel is too computationally expensive for operations such as filling in evidence cones or simulating ranges by tracing lines. These tasks can be decomposed into raster operations, which can be performed by a 3D variant of the classic 2D Bresenham line drawing algorithm.

### E. Estimate

Although many methods exist for generating a position estimate from the particle cloud, it has been found that simply taking the mean of the particle positions consistently produces the lowest estimation error [15].

TABLE I

PARTICLE FILTER NOTATION.

| | |
|---|---|
| $s_t^m$ | vehicle pose of the $m^{\text{th}}$ particle at time $t$ |
| | $= (roll, pitch, yaw, x, y, z)$ |
| ${}^{n}r_t$ | range measurement of the $n^{\text{th}}$ sonar at time $t$ |
| $u_t$ | vehicle dead-reckoned innovation at time $t$ |
| ${}^{ijk}\Theta^m$ | evidence grid voxel $ijk$ of the $m^{\text{th}}$ particle |
| $S_t$ | particle set at time $t$ |
| | $= \{(s_t^1, \Theta^1), ..., (s_t^m, \Theta^m)\}$ |
| $w_t^m$ | weight of the m$^{\text{th}}$ particle at time $t$ |
| $h(s_{t-1}, u_t, N(0, \sigma_s))$ | vehicle motion model with normal noise |
| $g(s_t, \Theta, N(0, \sigma_r))$ | range measurement model with normal noise |

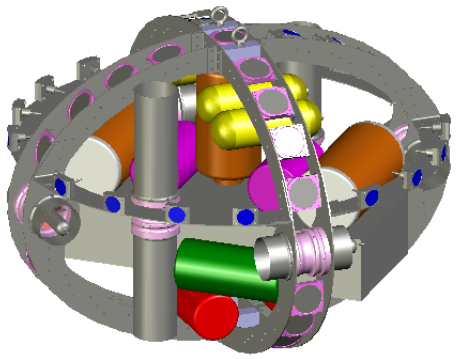

Fig. 5. A CAD model of the DEPTHX vehicle with the fairing removed to expose the sonar three circular sonar arrays.

## V. EXPERIMENTAL RESULTS

The SLAM position error on a single round-trip, or drop, from the Zacatón dataset, are shown in Figure 8, together with dead-reckoned position error. The DropSonde was lowered to 200m and then raised back to the surface on a cable, so although there was no significant change in $x$ and $y$, we simulated velocity noise with a zero mean and 0.05m/s standard deviation normal distribution, which roughly corresponds to real-world sensor noise. The SLAM solution shows considerable drift – we believe this is because the ring sonar array geometry does not allow the SLAM system to register itself to previously mapped areas: the sonar only looks at a slice of the tunnel at a time, and any position error between slices will accumulate.

The final DEPTHX vehicle will have a more sophisticated sonar array geometry, closely approximating three perpendicular great circles (abbreviated 3GC, see Figure 5), which has been show to be a good sonar geometry for the types of tunnels we expect to encounter [15]. Since the Zacatón data was collected with the ring sonar geometry, we used simulation test the performance of the 3GC geometry. For this purpose, we constructed a "dog's leg" tunnel with three segments (see Figure 6). A very simple simulator was used to generate the vehicle's trajectory from near the surface to the bottom of the well and back. In order to test the convergence performance

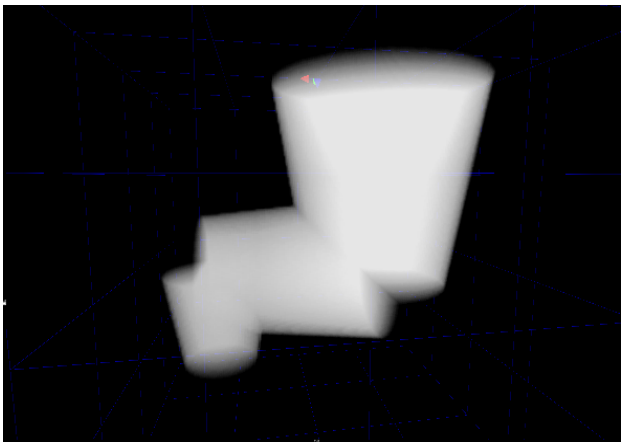

Fig. 6. The dog's leg map used for simulated experiments. The top of the tunnel is the at the "surface", and the bottom is 150m deep.

Fig. 7. A particle map $\Theta$ after running SLAM on the simulated dog's leg.

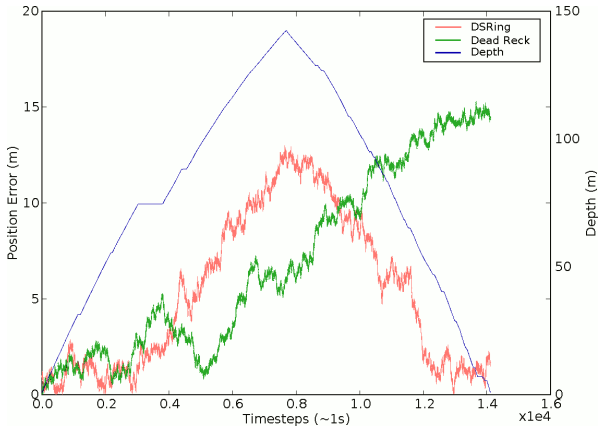


Fig. 8. Position error over time on a single round-trip 200m drop in Zacatón, which used the ring sonar geometry. The SLAM error demonstrates the weakness of the ring sonar geometry, though it correctly converges as the DropSonde is pulled back up to the surface.

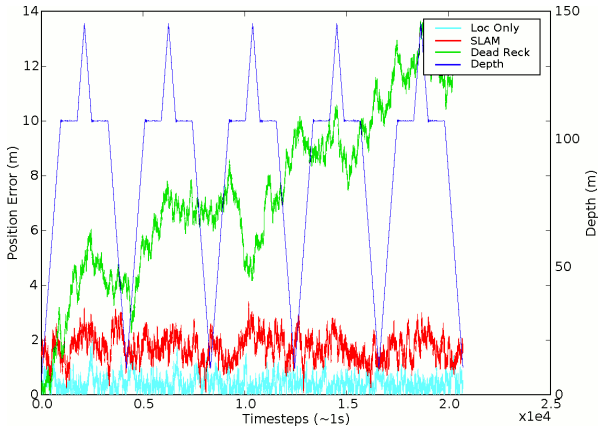


Fig. 9. Position error over time of the three different localization algorithms using the simulated dog's leg data and the 3GC sonar geometry. There are five round-trips.

of the SLAM system, the vehicle followed this trajectory five times. Using the same ray-trace ranging method described above and the simulated map, simulated ranges with normal noise distribution were generated for each vehicle position in the trajectory. The trajectory was also used to generate attitude, depth, and velocity sensor readings, also with normal noise distributions.

The SLAM algorithm used 20 particles with $200^3$ voxel maps at 1m resolution (see Figure 7 for the final map of a single particle). Each run took about 43 minutes on a modern desktop computer: about four times faster than real-time. We did five runs using SLAM, then localization only (by initializing the particle maps with the true map and performing no update step), then dead-reckoning only. The mean position error of these three methods is given in Figure 10. As expected, the SLAM error is higher than localization only, but much lower than dead-reckoning. More importantly, the SLAM error is bounded, as can be seen in Figure 9.

## VI. Future Work

The performance of the SLAM algorithm is clearly linked to the number of particles which can be supported, as well as the resolution of the particle maps. Currently, the algorithm is limited by the amount of memory required to store the particle maps, as well as the memory copy operation during the resample step – although insert and query operations are also significant. We are investigating more sophisticated data structures for storing and manipulating the 3D evidence grids.

A simple improvement to the current system would be to exploit an iterative exploration strategy by constructing a "known world" map off-line after each dive. This "known world" map could then be used on successive dives to localize the vehicle until it has reached the frontier of exploration, at which point the SLAM system could use smaller particle maps spliced onto the master map to allow the use of more particles.

A more general extension of this idea was implemented for 2D evidence grids in DP-SLAM [16], which explicitly keeps track of the shared map. Extending this concept to 3D evidence grids seems promising, as does some recent work, also using 2D evidence grids, on improving the proposal distribution and selective resampling [17].

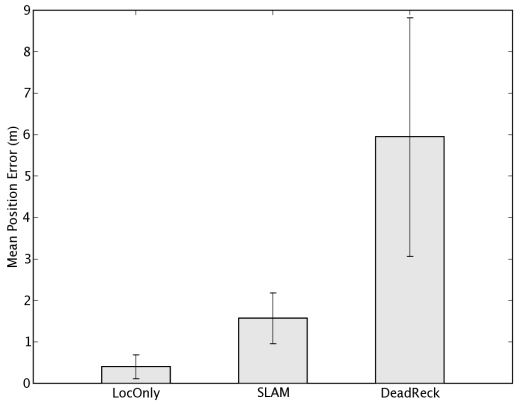


Fig. 10. Mean position error (distance between true position and estimated position) over five runs, and compares three cases: Localization Only, where the PF is initialized with the real map; SLAM, where the PF must build its own map; and Dead Reckoning, where the navigation sensors are used to generate a position.

## VII. Conclusion

We have demonstrated SLAM using a Rao-Blackwellized particle filter with 3D evidence grids using both real data collected from the Zacatón *cenote*, as well as a simulated tunnel system. Preliminary timing results indicate that the current system can be run in real-time with about 50 particles.

## References


[1] D. Bradley, D. Silver, and S. Thayer, "A regional point descriptor for global localization in subterranean environments," in *IEEE conference on Robotics Automation and Mechatronics (RAM 2005)*, December 2004.

[2] S. B. Williams, P. Newman, M. W. M. G. Dissanayake, and H. Durrant-Whyte, "Autonomous underwater simultaneous localisation and map building," in *Proceedings of IEEE International Conference on Robotics and Automation*, San Francisco CA, USA, April 2000, pp. 22–28.

[3] R. Eustice, H. Singh, J. Leonard, M. Walter, and R. Ballard, "Visually navigating the RMS Titanic with SLAM information filters," in *Proceedings of Robotics Science and Systems*, Cambridge, MA, June 2005, pp. Accepted, To Appear.

[4] S. B. Williams and I. Mahon, "Simultaneous localisation and mapping on the great barrier reef," in *IEEE International Conference on Robotics and Automation*, New Orleans, USA, April 26-May 1 2004.

[5] P. M. Newman, J. J. Leonard, and R. J. Rikoski, "Towards constant-time slam on an autonomous underwater vehicle using synthetic aperture sonar," in *Eleventh International Symposium of Robotics Research (ISRR)*, ser. Springer Tracts in Advanced Robotics (STAR), D. Paolo and R. Chatila, Eds. Siena, Italy: Springer Verlag, 2003, vol. 15, pp. 409–420.

[6] S. Thrun, D. Haehnel, D. Ferguson, M. Montemerlo, R. Triebel, W. Burgard, C. Baker, Z. Omohundro, S. Thayer, and W. R. L. Whittaker, "A system for volumetric robotic mapping of abandoned mines," in *Proceedings of the 2003 IEEE International Conference on Robotics and Automation (ICRA '03)*, May 2003.

[7] W. Stone, D. Wettergreen, G. Kantor, M. Stevens, E. Hui, E. Franke, and B. Hogan, "DEPTHX (deep phreatic thermal explorer)," in *Proceedings of the 14th International Symposium on Unmanned Untethered Submersible Technology (UUST)*, 2005.

[8] M. C. Martin and H. Moravec, "Robot evidence grids," Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-96-06, March 1996.

[9] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[10] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, "Rao-blackwellised particle filtering for dynamic bayesian networks," in *Uncertainty in AI*, 2000, pp. 176–183.

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

[12] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. New York, NY: McGraw-Hill, 1983.

[13] A. Doucet, "On sequential simulation-based methods for bayesian filtering," Cambridge University Department of Engineering, Tech. Rep. CUED/F-INFENG/TR. 310, 1998.

[14] S. Thrun, "Learning occupancy grid maps with forward sensor models." *Auton. Robots*, vol. 15, no. 2, pp. 111–127, 2003.

[15] N. Fairfield, G. A. Kantor, and D. Wettergreen, "Three dimensional evidence grids for SLAM in complex underwater environments," in *Proceedings of the 14th International Symposium of Unmanned Untethered Submersible Technology (UUST)*. Lee, New Hampshire: AUSI, August 2005.

[16] A. I. Eliazar and R. Parr, "DP-SLAM 2.0," in *Proceedings of ICRA*, 2004.

[17] G. Grisetti, C. Stachniss, and W. Burgard, "Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling," in *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, Barcelona, Spain, 2005, pp. 2443–2448.