

# A Regional Point Descriptor for Global Topological Localization in Flooded Subterranean Environments

David M. Bradley, David Silver, Scott M. Thayer  
Robotics Institute  
Carnegie Mellon University  
Pittsburgh, PA

**Abstract**—A regional point descriptor for global localization using natural landmarks in flooded subterranean environments is presented. Global localization in underwater environments is complicated by a lack of sensors that land robots depend on for position estimation such as GPS, LADAR, and wheel odometry. This descriptor, the slide image, is designed to take advantage of orientation references available in subterranean voids while tolerating expected pose estimation errors in the horizontal plane. It will serve as the basis of a robust topological navigation system for Minefish, an untethered, borehole deployable AUV designed to map flooded mines. Results are presented from tests on sonar data collected in the Wakulla springs tunnel system in Florida.

## I. INTRODUCTION

Robots have the potential to create more accurate and lower-cost maps of subterranean voids, such as abandoned mines, than current techniques allow. In the U.S. alone, tens of thousands of abandoned mines threaten nearby surface structures, ecosystems, and ongoing mining operations [1]. Proactive measures can prevent these threats from developing into disastrous situations; however, prevention relies upon knowledge of mine layout and condition. Mine maps are the primary source of such information. In the case of abandoned mines, these maps, when they exist, are often inaccurate if at all legible. Previous work has addressed robotic inspection and mapping of dry mines [2], human guided sonar mapping of underwater caverns [3], and autonomous underwater vehicle (AUV) operation in open water ([4] and [5]). However, no research so far has addressed the unique challenges encountered in mapping flooded mines.

Mine-mapping robots are forced to be autonomous because of the difficulty of communicating with the robot or retrieving it should it fail. A tether would quickly become tangled in a mine. Radio communication is generally precluded by water and the overburden of rock between the mine and the surface. Acoustic communication is plagued by multipath reflections. Additionally, since the cost of drilling a borehole is roughly proportional to the volume of material removed, many subterranean voids are only accessible through a narrow borehole.

Our prototype mine mapping system, Minefish, (Figure 1), is designed to fit down an 8" borehole, and the eventual goal is to miniaturize that to 4". Minefish will deploy on the end of a tether, detach and map the mine, and then redock with the tether for extraction. The narrow diameter of the borehole places severe constraints on the components that can be used,

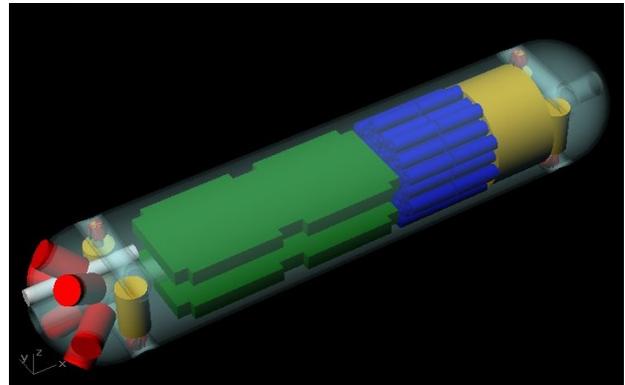


Fig. 1. Initial design for Minefish, our underwater autonomous mine mapping platform.

and virtually guarantees that there is no way of retrieving the robot should the autonomy fail after it has undocked from the tether. Not only does this require the autonomous navigation to be extremely robust, it also necessitates the use of an inexpensive, expendable robot.

Unfortunately most of the sensors used in previous underwater mapping efforts are neither small nor inexpensive. Inertial Measurement Units (IMUs) capable of accurate dead-reckoning navigation are far too costly to be practical. Active navigation beacons are problematic due to multipath interference and are difficult to deploy around the mine. Particle-based scan matching using data from a sonar rotating in the horizontal plane [6] has shown promising initial results but since it does not track persistent features it can accumulate pose errors over time. Once the robot has detached from its tether, its retrieval depends on correcting any pose errors well enough to find its tether again for redocking and extraction. This can be accomplished by identifying distinct natural landmarks in the environment, and storing compact descriptions (referred to as descriptors) of them in a map. If the landmarks are sufficiently distinct, they can be used for global localization.

Minefish will use a pressure sensor, an IMU, and a rotary sonar scanning in the horizontal plane for pose estimation. The effects of accumulated pose error will be minimized using a topological approach to underwater navigation described in the next section. In addition to the horizontal sonar, a rotary sonar scanning radially around the AUV will collect

a 3-D point cloud that will be used for localization and 3-D mapping. Section III provides an overview of previous work on descriptors created from point clouds. In section IV we propose a descriptor, the slide image, for recognizing natural landmarks in tunnel segments that is designed specifically to be robust to the sources of error found in the subterranean environment. Section V gives results from tests of slide images on sonar data collected in the flooded tunnel system of Wakulla Springs, Florida.

## II. UNDERWATER NAVIGATION

Underwater environments present serious navigational challenges. Laser range finders and wheel odometry are unavailable. Although sonar is much more accurate in water than in air, rotary sonars still have much coarser resolutions and slower update rates than the laser range finders that have been used successfully in previous mine-mapping efforts [2]. Fortunately, three of the six dimensions of the AUV's pose can be measured directly. The Z coordinate (depth) of the AUV's pose can be measured with a pressure sensor, and a standard tilt sensor can be used to measure roll and pitch angles. That leaves only motion in the horizontal plane (X, Y, and yaw). This can be estimated over a short time period using an IMU and a model of the AUV's thrusters and control surfaces, although without an extremely accurate and expensive IMU, error will accumulate rapidly in this estimate. An overview of common methods for AUV position estimation can be found in [7].

It is not necessary for Minefish to know explicitly where it is in order for it to explore a mine and return to its tether for extraction. As long as Minefish has reliable low-level navigational skills such as corridor following and obstacle avoidance, it only needs to build and navigate with a topologically correct map – one that describes how different locations in the mine are connected. One example of topological navigation is driving directions, which are commonly given as a list of landmarks ("3rd street", "stop sign", etc...) and actions to take at those locations ("turn right", "go straight", etc...). Mines, like roads, are ideal environments for topological navigation because for the most part they consist of well defined corridors and intersections. Creating reliable and distinct landmarks becomes the most significant challenge, as a mine-mapping robot must be able to recognize when it has closed a loop in order to explore efficiently.

The traditional approach to topological navigation [8] creates a topological graph of the world with nodes at distinctly recognizable places in the environment and edges that represent the paths connecting them. Usually nodes used in topological maps correspond to actual topological features of the environment, such as nodes of the generalized Voronoi graph (GVG) [9], but it is only necessary that the nodes are reliably recognizable places. Adding intermediate nodes of degree two to topological edges, where these nodes correspond to reliable landmarks, creates a topologically correct map with additional information about metric position. The challenge examined in the next section is how to create a reliable

landmark by capturing the unique aspects of a short segment of subterranean tunnel in a compact representation which is robust to the noise inherent in the observation process.

After the robot has returned to its tether, accurate but computationally intensive batch methods can be used to create the actual metric map of the mine from logged sensor data. For an overview of such methods see [10].

## III. REGIONAL POINT DESCRIPTORS

Localization in a subterranean tunnel system can be thought of as an object recognition problem where short sections of the tunnels serve as the objects to be recognized. As the robot explores, descriptions of the tunnel segments it passes through are placed into a library. Before a tunnel segment is added to the library it is compared against the tunnel segments already in the library for matches (from closing a loop or going back along a previously explored tunnel for example). This comparison step contains the main challenge. First, the sensor measurements of a tunnel segment (in this case a 3-D point cloud from a sonar scanning radially around the vehicle) must be converted into a compact yet discriminative feature vector (descriptor). Secondly, a particular segment of tunnel must reliably generate the same descriptor no matter where it is viewed from or how much error has accumulated in the robot's pose estimate. Finally, an accurate and computationally efficient distance metric must be employed that minimizes the differences between descriptors from the same location and emphasizes the differences between descriptors from different locations.

The first two challenges are highly inter-dependent. If a consistent origin, orientation, and scale can be found for an object, then a very detailed description of the object's point cloud can be used. Possible robot pose estimation errors, however, make finding this difficult. Fortunately, much of the previous point descriptor work has focused on tasks, such as recognizing vrml models, where the pose and scale of the object to be recognized are unknown. Various techniques have overcome this ambiguity in orientation by estimating an orientation vector from the local surface normal, or by forgoing orientation information altogether and matching point statistics.

A number of techniques use the local surface normal as an orientation reference when describing the shape of a point cloud  $P$  in a support region around a reference point  $p$ . Spin images [11], and 3-D shape contexts [12] divide the support region around the reference point into bins and use a histogram of how many points fall into each bin as a descriptor by arranging the bin counts into a feature vector. Two linearly-independent vectors are required to define an orientation in 3-D space, however, and the estimated surface normal at  $p$  provides only one of these. This leaves one remaining degree of ambiguity in orientation. Spin images avoid this orientation ambiguity by collapsing the extra rotational degree of freedom. The surface normal is used as the axis of a cylindrical support region around  $p$ , which is then sliced into cylindrical shells based on distance along the axis and

distance radially from the axis. 3-D shape contexts use a spherical support region which is oriented with the local surface normal as the north pole and divided into bins by azimuth, elevation, and radial distance. Because the azimuth orientation has not been constrained, comparing 3-D shape contexts requires the brute force approach of trying  $L$  different azimuth orientations around the surface normal. Not only is this computationally expensive, it also increases the chance of false matches. Finally, harmonic shape descriptors [13], [12] use the same spherical histogram as 3-D shape contexts, but compute a rotation-invariant descriptor from it by applying a harmonic transform and using the magnitude of the frequency-domain coefficients of the 3-D histogram as a descriptor. Because rotations of the 3D histogram in the spatial domain are phase shifts in the frequency domain, they do not affect the magnitude of the harmonic coefficients.

Accurate surface normals can be hard to compute when, as is often the case with sonar data, the point cloud is noisy and sparse. To avoid having to estimate surface normals, another class of techniques sacrifices information about the object’s orientation entirely and instead creates feature vectors from the statistical occurrence of features of the object’s surface. This has been tried with histograms of geometric relationships like the distance between randomly selected pairs of points on the surface [14], and more complex local features that also take into account the relative surface orientations of pairs of points [15].

Section IV describes a feature, the slide image, that takes advantage of the subterranean environment to solve the orientation ambiguity problem. Specifically the feature assumes that subterranean voids are generally tunnels, estimates one orientation vector from the axis of the tunnel, and uses the gravity vector as the other orientation vector.

#### IV. SLIDE IMAGES

The ideal descriptor would be (i) highly discriminative, (ii) robust to sensor noise, (iii) independent of the global coordinate frame, and (iv) compact. In the underwater environment, depth, pitch, and roll can be sensed absolutely. However, uncertainty can accumulate in yaw orientation and position in the horizontal (X-Y) plane. Good descriptors will be invariant to possible errors in these dimensions of the robot’s pose, while still retaining as much identifying information as possible about the tunnel segment. Because the robot’s estimated yaw pose may be wrong, either a consistent yaw reference point must be estimated or the feature must be made invariant to yaw orientation, which necessarily makes it less discriminative. Fortunately the subterranean environments that we are interested in are largely made up of tunnels. Tunnel environments have a natural axis along the length of the tunnel, and this natural orientation can be used to line up tunnel segments in a consistent way regardless of long term drift in the estimated yaw pose. However, there may still be inaccuracies in the ability to judge how far and in which direction the robot has moved down the tunnel, so we eliminate the need to measure those parameters by histogramming the

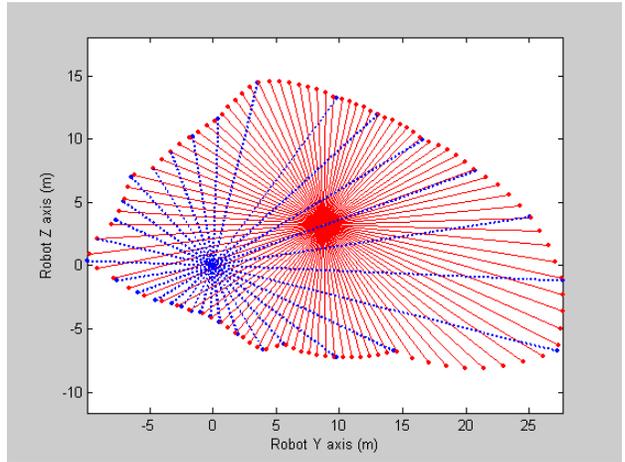


Fig. 2. Points collected with a radially scanning sonar (blue) have a higher density on closer walls of a tunnel. Resampling from the geometric center of the tunnel cross section (red) reduces this bias. Shown is a typical sonar scan from the Wakulla Springs tunnel system [3]. Points are plotted with respect to the robot Y and Z axes.

points making up the tunnel walls with respect to distance from the central axis of the tunnel and roll angle around the central axis (figure 3).

Our descriptor, the slide image, is basically the dual of the spin image. In the subterranean environment we have two orientation vectors at our disposal, the axis of the tunnel and the gravity vector, and we expect inaccurate positioning within the horizontal plane to be our major source of error. Hence we collapse the along-axis dimension rather than the around-axis dimension to make a descriptor robust to slides – translations and rotations on the horizontal plane – rather than spins.

#### A. Ensuring View Independence

Before computing the histogram, however, we first resample the sonar readings in order to correct for the position of the vehicle within the tunnel. The sonar ring samples points at a uniform radial density, meaning that if the sonar ring is not in the center of the tunnel, the returns will have a higher point density on the near walls and a lower point density on the far walls (Figure 2). To correct this imbalance, the geometric center of each scan  $x_c$  is computed as the weighted mean of the points in the scan (1) where the weight  $w_i$  of each point  $x_i$  with distance  $r_i$  from the vehicle is given by (2).

$$x_c = \frac{\sum_{i=1}^n w_i x_i}{n} \quad (1)$$

$$w_i = \frac{2\pi r_i}{\sum_{k=1}^n w_k} \quad (2)$$

The axis of the tunnel is then computed with the covariance matrix decomposition approach used in [16]. First, a covariance matrix is calculated from the center points,  $\{X_{c_i}\} = \{(x_{c_i}, y_{c_i}, z_{c_i})\}$ , of the  $N$  scans taken in the last 10m of the tunnel as defined in (3). Even though the descriptor is only generated from the last 5m of the tunnel, 10m is used to

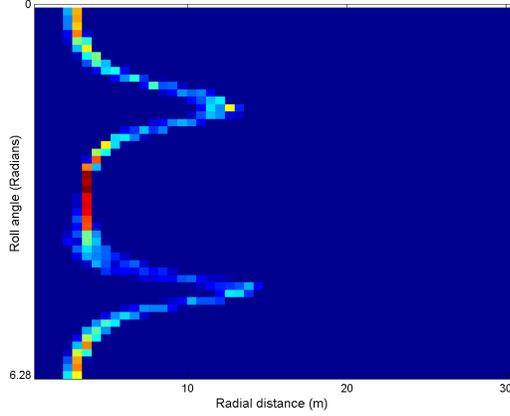


Fig. 3. Points collected with a radially scanning sonar are binned based on distance from and roll angle around the estimated tunnel axis to create the histogram shown.

provide a more robust estimation of the tunnel axis.

$$\frac{1}{N} \sum_{i=1}^N (X_{c_i} - \bar{X}_c)(X_{c_i} - \bar{X}_c)^T \quad (3)$$

Next, the principal components of the covariance matrix  $\epsilon_0$ ,  $\epsilon_1$ , and  $\epsilon_2$  are computed along with their associated eigenvalues  $\lambda_0$ ,  $\lambda_1$ , and  $\lambda_2$ , where  $\lambda_0 \geq \lambda_1 \geq \lambda_2$ . The principal component with the largest eigenvector,  $\epsilon_0$ , is taken to be the axis of the tunnel. The relative magnitudes of the eigenvalues  $\lambda_0$  and  $\lambda_1$  provide insight into how stable is  $\epsilon_0$  as an orientation vector. If  $\lambda_0 \gg \lambda_1$ , the centerpoints all lie close to one line, and it is likely that the same line would be estimated in future passes though the same section of tunnel. If  $\lambda_0 \approx \lambda_1$  the direction of  $\epsilon_0$  is sensitive to corruption by sensor noise. For the test data set used in the next section it was found that filtering out descriptors where  $\lambda_0 < 40\lambda_1$  (about 50% of the total) significantly increased system performance.

### B. Descriptor Comparison

Once the natural axis of the tunnel is found, the point cloud is binned by radius  $r$  from the axis and roll angle  $\theta$  around the axis. It is then scaled to a range of 0-255 and quantized to 8-bit values. The histograms can then be compared with a variety of distance metrics. The next section includes results for the sum of absolute differences (4), the sum of squared distances (5), the correlation coefficient (6), and the symmetric form of the Kullback-Leibler Divergence (7).

$$SAD(D_q, D_r) = \sum_i |D_q(i) - D_r(i)| \quad (4)$$

$$SSD(D_q, D_r) = \sum_i (D_q(i) - D_r(i))^2 \quad (5)$$

$$r(D_q, D_r) = \frac{\sum_i (D_q(i) - \bar{D}_q)(D_r(i) - \bar{D}_r)}{\sqrt{\left(\sum_i (D_q(i) - \bar{D}_q)^2\right)\left(\sum_i (D_r(i) - \bar{D}_r)^2\right)}} \quad (6)$$

$$\kappa(D_q, D_r) = \sum_i (D_q(i) - D_r(i)) \ln \left( \frac{D_q(i)}{D_r(i)} \right) \quad (7)$$

## V. EVALUATION OF FEATURE PERFORMANCE

### A. Performance Criteria

For our testing we largely adopted the methodology presented in [17] and [18], which concentrates on the Receiver Operating Characteristics (ROC) of probability of target detection versus probability of generating false alarms. ROC curves are created by comparing test “query” descriptors against a library of “reference” descriptors computed from a separate set of sensor data. To perform this comparison we first define a distance function  $d(D_a, D_b)$  which gives a scalar output for every pair of descriptors  $D_a$  and  $D_b$ . The query descriptor is said to match any reference descriptors in the library for which the feature space distance  $d(D_q, D_r)$  is less than an arbitrary threshold  $\alpha$  (8). Thus, every query descriptor  $D_q$  returns a (possibly empty) set of reference descriptors  $M_F$  that are within  $\alpha$  of it in feature space.

$$M_F = \{D_r | d(D_q, D_r) < \alpha\} \quad (8)$$

We then compare the set of feature space matches against the set of reference descriptors that are actually within  $\beta$  meters of the query descriptor (as measured along the estimated tunnel axis). For the results presented  $\beta$  was chosen to be 1m. The probability of detection is computed as the number of correct matches retrieved divided by the total number of correct matches (9). The false positive rate is the number of reference descriptors falsely classified as matches divided by the total number of reference descriptors examined (10). Since the value of  $\beta$  used was significantly less than the length of the tunnel segment used to generate a slide image, two descriptors can share up to 80% of the same tunnel segment without being considered a match. Because of this problem, reference descriptors between 1m and 5m of the query descriptor were not counted as false positives. Repeating the recall test for multiple values of  $\alpha$  produces the ROC curve.

$$p_d = \frac{\# \text{ of correct matches retrieved}}{\# \text{ of matches}} \quad (9)$$

$$p_{fa} = \frac{\# \text{ of false matches retrieved}}{(\# \text{ reference descriptors})(\# \text{ queries})} \quad (10)$$

### B. Localization in an underground spring

In 1998 and early 1999 the U.S. Deep Caving Team used a custom sonar buoy with a sophisticated IMU to map a large section of the underwater cave system that feeds Wakulla Springs in Florida [3]. Certain sections of the tunnel system were mapped multiple times in separate runs, allowing for a test of localization performance on real-world subterranean sonar data.

The test set used consists of sonar data from two separate mapping runs that share a common segment. After the first 210 meters the runs diverge and follow different tunnels (see figure 4 for a view of the point clouds of both runs). First, a library of reference descriptors was created from the longer of the two

runs. Then query descriptors were generated from the other run and compared against the library of reference descriptors. In this case 1,108 query descriptors (of which 396 had matches) were compared against a database of 2,963 descriptors.

Figure 5 shows the ROC curves generated for several different binning resolutions, and a closeup of the important leftmost section of the ROC curves is shown in figure 6. The best performance at low false positive rates was obtained with a histogram composed of 400 bins: 20 in radial distance, 20 in roll angle, and only one in the along-axis direction. Binning along the tunnel axis makes the feature more sensitive to noise and errors in positioning, which may explain the recall degradation at low false positive rates. Similarly, dividing the radial distance and roll-angle dimensions into more than 20 bins seemed to extract more noise than actual information about the environment.

Using the best binning resolution, slightly less than half of the 4862 possible correct matches were recalled without a single false positive. Because there are on average 12.2 reference descriptors within 1m of each query from the shared tunnel segment, not all of the possible matches have to be recalled in order to match each location. For reliable localization, the set of feature space matches  $M_F$  of each query should contain at least one correct match and no false positives. Figure 7 shows the probability that a query with matching reference descriptors in the library will be matched to at least one of them. This probability is displayed against the probability that  $M_F$  will contain at least one false positive. Results are displayed for the different distance functions discussed in section IV-B. The symmetric form of the Kullback-Liebler divergence proves to be the most effective, but it also is the most computationally intensive. Correlation shows better performance than SAD or SSD, but SSD may eventually provide the fastest searches.

Additionally, the recall performance in Figure 7 has been increased by taking into account the depth of the center of segment used to create each descriptor. The observed center of a segment can vary with measurement noise, but none of the 4,862 pairs of matching segments in the Wakulla dataset had centers more than 3.5m apart in depth. Only considering reference descriptors within 3.5m of depth of the query descriptor reduced the number of comparisons by over 80%, significantly increasing both the speed and the precision of the algorithm. The effect of depth filtering on localization performance can be seen in Figure 8.

## VI. CONCLUSION AND FURTHER WORK

We have presented a regional point descriptor, the slide image, designed for global topological localization in flooded subterranean environments. We take advantage of the fact that subterranean voids, particularly mines, are structured as a series of tunnels, and use the axis of the tunnel along with the gravity vector to create a consistent reference orientation. Having this stable orientation allows us to design a descriptor that is robust to pose estimation errors in the horizontal plane.

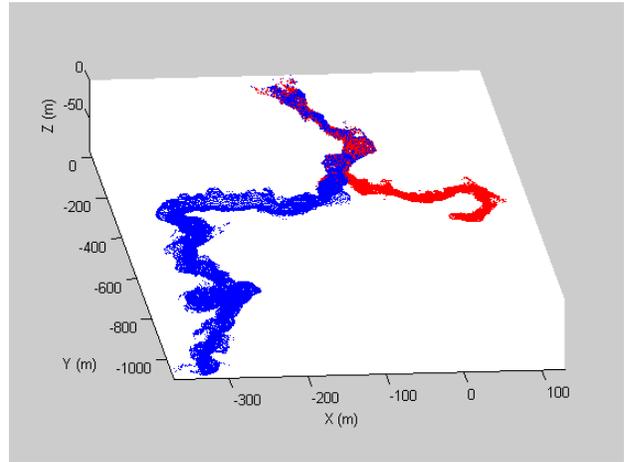


Fig. 4. Point clouds from two different mapping runs through the Wakulla Springs tunnel system [3]. Both runs start in the same tunnel but diverge after approximately 210 meters. For feature performance evaluation the run on the left was used to generate a database of reference descriptors, and descriptors from the other run were matched to this library.

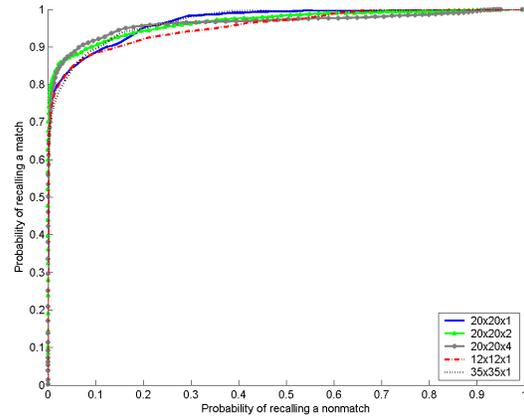


Fig. 5. ROC curves for different histogram binning schemes. Curves are labeled with a triplet of numbers describing the number of uniformly distributed bins used in the radial, around axis, and along axis directions respectively. A 20 by 20 bin slide image (one bin along the tunnel axis) seems to give the best performance at for the Wakulla Springs dataset. Reference descriptors within 1m of the query location are considered to be matches.

The performance of the descriptor has been demonstrated on real sonar data collected in the Wakulla springs tunnel system.

Now that slide images have shown their potential for subterranean navigation, the emphasis has shifted to creating an efficient implementation for testing and use on Minefish. The results presented in this paper were generated using an unoptimized matlab implementation that performed an exhaustive search of the entire library for each query. Various indexing and feature compression techniques exist to speed up this search without significantly compromising accuracy. In [11] the comparison of spin images has been accelerated by first compressing the descriptors with principal components analysis (PCA) and then comparing the smaller compressed descriptors. It even may be possible to skip the step of

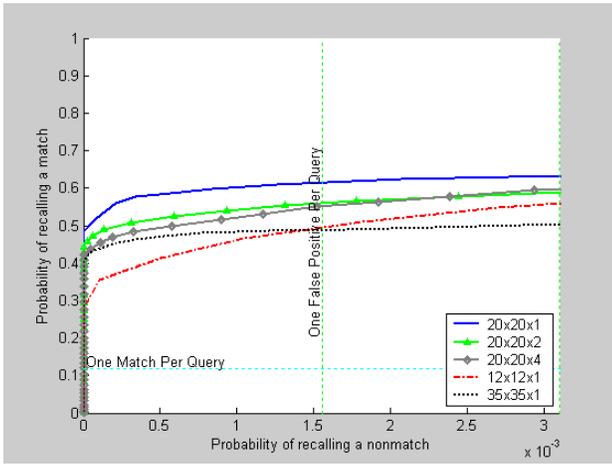


Fig. 6. Zoomed in version of figure 5. Because there are usually multiple reference descriptors within 1m of the query descriptor, only a relatively low recall rate is needed to produce good localization performance.

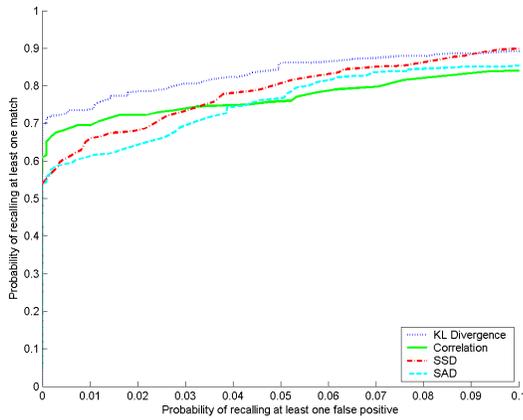


Fig. 7. Comparison of the performance of the distance functions described in section IV-B on a global localization test. Shown is the probability that at least one correct match will be returned for a query, if there is a match, vs. the probability that at least one false positive will be returned for the query.

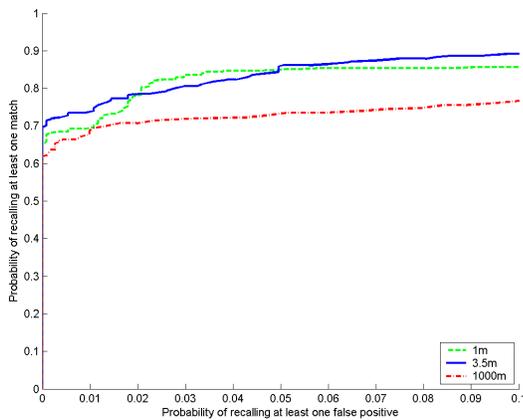


Fig. 8. Restricting the search to only those reference descriptors close in depth to the query descriptor increases recall performance and speed. Results are shown for vertical distance thresholds of 1m, 3.5m, and 1000m (no depth filtering).

computing the principal components. The authors of [19] report good results using random projections to compress spin images.

#### ACKNOWLEDGMENTS

Chuck Wittaker, Zachary Omohundro, Mike Sokolsky, and Chris Atwood, for their efforts in developing Minefish.

George Kantor and David Wettergreen, for continued feedback and commentary.

The US deep caving team for their work mapping the Wakulla Springs tunnel system and their willingness to share their data.

Senator Arlen Specter for his continued support of this work.

#### REFERENCES

- [1] J. Belwood and R. Waugh, "Bats and mines: Abandoned does not always mean empty," *Bats*, vol. 9, no. 3, 1991.
- [2] C. Baker, A. Morris, D. Ferguson, S. Thayer, C. Whittaker, Z. Omohundro, C. Reverte, W. Whittaker, D. Hähnel, and S. Thrun, "A Campaign in Autonomous Mine Mapping," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, New Orleans, LA, 2004.
- [3] W. Stone, B. Ende, F. Wefer, and N. Jones, "Automated 3d mapping of submarine tunnels," in *Robotics 2000: ASCE Conference on Robotics for Challenging Environments*, Albuquerque, NM, February 2000.
- [4] S. Williams and I. Mahon, "Simultaneous localization and mapping on the great barrier reef," in *International Conference on Robotics and Automation*, New Orleans, LA, April 2004.
- [5] S. Williams, G. Dissanayake, and H. Durant-Whyte, "Towards terrain-aided navigation for underwater robotics," *Advanced Robotics*, vol. 15, no. 5, 2001.
- [6] D. Silver, D. Bradley, and S. Thayer, "Scan matching for flooded subterranean voids," in *Submitted to RAM 04*, 2004.
- [7] L. Whitcomb, "Underwater robotics: Out of the research laboratory and into the field," in *IEEE International Conference on Robotics and Automation*, 2000.
- [8] B. Kuipers and Y. Byan, "A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations," *J. Robot. Auton. Syst.*, vol. 8, pp. 47–63, 1991.
- [9] H. Choset and K. Nagatani, "Topological simultaneous localization and mapping (slam): towards exact localization without explicit localization," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 2, pp. 125–137, Apr. 2001.
- [10] S. Thrun, "Robotic mapping: A survey," 2002.
- [11] A. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 433–449, 1999.
- [12] A. Frome, D. Huber, R. Kolluri, T. Bulow, and J. Malik, "Recognizing objects in range data using regional point descriptors," in *Proceedings of the European Conference on Computer Vision (ECCV)*, May 2004.
- [13] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3d shape descriptors," in *Symposium on Geometry Processing*, Aachen, Germany, June 2003.
- [14] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Transactions on Graphics*, vol. 21, no. 4, pp. 807–832, October 2002.
- [15] E. Wahl, U. Hillenbrand, and G. Hirzinger, "Surflet-pair-relation histograms: A statistical 3d-shape representation for rapid classification," in *International Conference on 3-D Digital Imaging and Modeling*. IEEE, 2003.
- [16] N. Vandapel, D. Huber, A. Kapuria, and M. Hebert, "Natural terrain classification using 3-d ladar data," in *IEEE International Conference on Robotics and Automation*, April 2004.
- [17] C. Schmid and K. Mikolajczyk, "A performance evaluation of local descriptors," in *CVPR*, vol. 2, 2003, pp. 257–263.
- [18] G. Caneiro and A. Jepson, "Phase-based local features," *ECCV*, vol. 1, pp. 282–296, 2002.
- [19] M. M. S. Ruiz-Correa, L. Shapiro, "A new signature-based method for efficient 3-d object recognition," in *CVPR*, 2000.