Selecting Landmarks for Localization in Natural Terrain

CLARK F. OLSON
University of Washington, Bothell, Computing and Software Systems, 18115 Campus Way NE, Box 358534, Bothell, WA 98011-8246, USA

Abstract. We describe techniques to optimally select landmarks for performing mobile robot localization by matching terrain maps. The method is based upon a maximum-likelihood robot localization algorithm that efficiently searches the space of possible robot positions. We use a sensor error model to estimate a probability distribution over the terrain expected to be seen from the current robot position. The estimated distribution is compared to a previously generated map of the terrain and the optimal landmark is selected by minimizing the predicted uncertainty in the localization. This approach has been applied to the generation of a sensor uncertainty field that can be used to plan a robot’s movements. Experiments indicate that landmark selection improves not only the localization uncertainty, but also the likelihood of success. Examples of landmark selection are given using real and synthetic data.

Keywords: mobile robot, localization, landmark selection, terrain map, maximum-likelihood estimation, sensor uncertainty field

1. Introduction

A robot must know where it is in the environment in order to perform useful tasks. Unfortunately, simple techniques for maintaining position knowledge, such as integrating odometry sensors and dead-reckoning are prone to significant errors. For this reason, position updates using other sensors such as vision or sonar are commonly used. Probabilistic methods such as Markov localization (Cassandra et al., 1996; Fox et al., 1998; Koenig and Simmons, 1996; Nourbakhsh et al., 1995; Simmons and Koenig, 1995; Thrun et al., 1998) are a useful tool for performing such updates. We have previously developed such a probabilistic method for performing localization by matching dense range maps of natural terrain (Olson, 2000a). This method operates under the assumption that the local range map is generated with a stereo range sensor that was pointed at recognizable terrain containing distinctive landmarks. In this paper, we extend this method to use a previously developed map of the terrain to automatically select the positioning of the sensor in order to improve the localization performance.

Several recent papers have discussed strategies for sensor placement or landmark selection for use in navigation and localization. A common approach is to consider which landmarks, from a pre-determined set of landmarks, will yield the best localization result. Sutherland and Thompson (1994) developed one of the earliest methods for landmark selection. They applied heuristic functions to select a landmark triple, from the set of such triples, that is likely to yield a good localization result. Greiner and Isukapalli (1996) learn a function to select landmarks that minimize the expected localization error. A related technique is given by Thrun (1998), who trains a neural network to learn landmarks that optimize the localization uncertainty. Fox et al. (1998) describe a strategy for both determining the robot motion and sensing direction in order to improve robot localization.

Yeh and Kriegman (1995) select the subset of possible features that minimizes a Bayesian cost of localization. Deng et al. (1996) select a set of landmarks in order to minimize the cost of sensing over a path segment. Murphy et al. (1997) first select candidate landmark triples using heuristics. The highest ranked candidate (and others, if necessary) are tested using experimentation with a robot. Sim and Dudek (1998) consider image locations with high edge density as possible landmarks, which are represented using an
appearance-based method. Landmarks are detected by matching in the image subspace and the resulting estimates are combined in a robust manner. Little et al. (1998) find stable landmarks by first detecting image corners. The corners that lie on depth discontinuities are eliminated using stereo vision.

Each of these papers considers a problem where the landmarks are selected from a pre-determined set of possible landmarks. Research that does not assume a pre-determined set of landmarks includes work by Simhon and Dudek (1998). They choose regions in which good metric maps can be established according to a distinctiveness measure. Grudic and Lawrence (1998) learn a mapping between an image and the robot location, but they do not address the problem of where to best place the camera to obtain the image.

The closely related problem of maintaining low uncertainty (in the position of the robot or the representation of the world) over some course of action has also been examined. Whaitie and Ferrie (1997) discuss autonomous exploration strategies that can plan the robot sensing steps in order to improve the robot's internal representation of the world. Roy et al. (1999) give techniques for determining robot paths that take into account the uncertainty along the path, thus reducing the average localization error.

Little of the research to date can be successfully applied to localization in unstructured outdoor terrain, which is the problem that we address. We describe a technique that selects the best landmark to view for localization when the robot has only an elevation map of the terrain and an estimate of the robot's position. We assume that the robot is equipped with a limited field of view (FOV) range sensor, such as stereo cameras or sonar. Our method selects the position to aim the range sensor in order to optimally perform localization in the unstructured three-dimensional terrain.

The landmark selection technique that we use is based upon performing uncertainty estimation using a maximum-likelihood localization method (Olson, 2000a). This method uses a probabilistic technique to perform localization using dense range data in natural environments. Prior to performing localization, the robot analyzes the terrain in the global map to select a localization target, which is the position in the terrain that the robot senses in order to generate a local map to compare against the global map. We desire a location that has distinctive terrain and thus allows the localization to be performed with a low uncertainty. This assumes that the error in the robot position is not so large that the localization target will be outside of the view of the robot when it attempts to sense this location. Active vision techniques can be used if, after the robot attempts to sense the localization target, no distinctive terrain is seen.

The first step in determining the localization target is estimating the error present in the global map and the error expected from sensing the terrain at the robot's current position. These errors are encoded in a probability map of the terrain expected to be seen by the robot. Each cell in this map contains an estimate of the probability that the cell will be seen as occupied by the robot, if the robot performs sensing with the cell in the field-of-view. By treating this probability map as a terrain map, we can apply previously developed uncertainty estimation techniques (Olson, 1999) to predict the uncertainty that will occur for any target in the probability map. The location with the lowest predicted uncertainty is selected as the localization target.

In addition to improving localization, these techniques can be applied to determining a sensor uncertainty field for the robot. The sensor uncertainty field is a concept introduced by Takeda et al. (1994) that measures the expected distribution of errors in the robot position as the robot moves through some environment, performing sensing periodically in order to improve localization. Given the uncertainty estimation and target selection methods, we can determine the expected localization uncertainty for any robot position in the environment.

These techniques have been applied to localization for Rocky 7, which is a research prototype for Mars exploration and science operations. In addition to body-mounted stereo cameras on the front and back, Rocky 7 has a stereo pair of cameras on a retractable mast that allows it to survey the terrain. See Fig. 1. We, thus, concentrate on localization using stereo range data. However, this methodology is also applicable to other range sensors. Experiments on Rocky 7 and with synthetic data indicate that the landmark selection not only decreases the robot's localization uncertainty, but also increases the probability of achieving a qualitatively correct localization result.

In the following section, we review the terrain matching methodology that we have previously developed for performing localization using dense terrain maps in natural environments. Section 3 describes a new representation for the terrain maps using a probabilistic scheme to predict the probability of each terrain location being seen as occupied by the robot after
moving and performing sensing at a new location and Section 4 describes the use of these probability maps in our localization framework for selecting landmark positions that are likely to yield good localization results. Section 5 describes results that we have achieved using synthetic data and gives an example of the use of this technique with real images. Section 6 summarizes and concludes the paper.

2. Terrain Matching

The basic localization technique that we use is to compare a map generated at the current robot position (the local map) to a previously generated map of the environment (the global map) (Olson and Matthies, 1998). This section reviews this method. The following section generalizes these maps using the concept of a probability map that captures both the error present in the global map and the error expected in generating a new local map. The pose likelihood function and uncertainty estimation technique described here will then be applied to the probability map in order to predict the terrain patch that can be viewed in order to achieve the lowest localization uncertainty.

We generate both the local map and the global map (which may be the combined result of previous local maps) using stereo vision on-board the robot. The range image is converted into a digital elevation map under the assumption that we know the robot orientation through other sensors, although this restriction can be removed, if desired (Olson, 2000a). To further simplify the problem, we use a high-pass filter on the terrain heights, so that the search for the robot position only needs to be performed in the x and y directions. The generated representation is then converted into a three-dimensional occupancy grid.

2.1. Map Similarity Measure

We formulate the map matching problem in terms of maximum-likelihood estimation. A convenient set of measurements that can be used for this problem are the distances from the occupied cells in the local map to their closest occupied cells in the global map according to some relative position between the maps. These distances are a function of the estimated robot position, since this affects the relative position between the maps. We maximize the likelihood of the observed distances over the space of possible robot positions in order to determine the maximum-likelihood position estimate. Denote the distances $D_{X_1}, \ldots, D_{X_n}$ for the robot position $X$, where $n$ is the number of occupied cells in the local map. The likelihood function for the robot position can be formulated as the product of the probability distributions of these distances (Olson and Matthies, 1998). We find the maximum of this pose likelihood function in order to localize the robot. For convenience, we work with the logarithm of the likelihood function:

$$\ln L(X) = \sum_{i=1}^{n} \ln p(D_X^i).$$

(1)

For the uncertainty estimation to be accurate, it is important that we use a probability distribution function (PDF) that closely models the sensor uncertainty. This can be accomplished using a PDF that is the weighted sum of two terms (Olson, 2000b):

$$p(d) = \alpha p_1(d) + (1 - \alpha) p_2(d).$$

(2)

The first term describes the error distribution when the cell is an inlier (in the sense that the terrain position...
under consideration in the local map also exists in the global map). In this case, the distance \( d \) is a function of errors in both the local map and the global map at this position. To evaluate the likelihood function (1) for a position \( X \), we use the values \( D_1^x, \ldots, D_n^x \) as \( d \) in this PDF. In the absence of additional information with respect to the sensor error, we approximate \( p_\delta(d) \) as a normal distribution:

\[
p_\delta(d) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{d^2}{2\sigma^2}}. \tag{3}
\]

The second term in Eq. (2) describes the error distribution when the cell is an outlier. In this case the position represented by the cell in the local map does not appear in the global map (e.g. due to range shadows or stereo outliers). In practice, we have found that modeling this term as a constant is both convenient and effective (Olson, 2000b).

\[
p_\delta(d) = K. \tag{4}
\]

Although, \( p_\delta(d) \) is not a probability distribution (it does not integrate to one), using the expected probability density for a measurement generated by a random outlier point yields excellent results:

\[
K = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(d) dx \, dy. \tag{5}
\]

This value can be estimated quickly through examination of the Euclidean distance transform of the map (Olson, 2000b).

In Eq. (2), \( \sigma \) is the probability that any particular cell in the local map is an inlier. For our occupancy grids, we assume that this value is relatively large (\( \sigma = 0.95 \)). In practice, the localization is insensitive to small changes in this value. Finally, \( \sigma \) is the standard deviation of the measurements that are inliers. This value can be determined from the characteristics of the sensor, or it can be estimated empirically by examining real data, which is the method that we have used for localization with Rocky 7.

### 2.2. Search Strategy

We locate the most likely robot position by adapting a multi-resolution search strategy that has been applied to image matching using the Hausdorff distance (Huttenlocher and Rucklidge, 1993; Olson and Huttenlocher, 1997; Rucklidge, 1997). We first test the nominal position of the robot given by dead-reckoning (or any other position, if no initial estimate is available) so that we have an initial position and likelihood to compare against. Next, the pose space is divided into rectilinear cells. Each cell is tested using a conservative test to determine whether it could contain a position that is better than the best position found so far (or any threshold, in general). Cells that cannot be pruned are divided into smaller cells, which are examined recursively. Further details can be found in (Olson, 2000a).

### 2.3. Uncertainty Estimation

We determine the uncertainty in the localization estimate by fitting a parameterized surface to the likelihood function in the neighborhood of the highest peak (Olson, 1999). Since the likelihood function measures the probability that each position in the pose space is the actual robot position, the uncertainty in the localization is measured by the rate at which the likelihood function falls off from the peak.

We assume that the likelihood function can be approximated as a normal distribution in the neighborhood around the peak location. Fitting a normal distribution to the computed likelihoods yields both an estimated variance in the localization estimate and a subpixel estimate of the peak location. We, thus, fit the peak in the likelihood function with:

\[
\frac{1}{2\pi \sigma x \sigma y} e^{-\frac{1}{2\sigma^2} \left( \frac{(\mu_x - \mu)^2}{\sigma_x^2} + \frac{(\mu_y - \mu_y)^2}{\sigma_y^2} \right)},
\]

where \( \mu_x \) and \( \mu_y \) represent the subpixel position estimate, \( \sigma_x \) and \( \sigma_y \) are the standard deviations along the axes, and \( \rho \) describes the orientation of the axes with respect to the global coordinate frame. The function is fit using the peak value and the eight neighboring values using a least-squares criterion in the log-likelihood domain.

In addition to estimating the uncertainty in the localization estimate, we can use the likelihood scores to estimate the probability of a failure to detect the correct position of the robot (Olson, 1999). This is particularly useful when the terrain yields few landmarks for localization and, thus, many positions appear similar to the robot.
3. Probability Mapping

In order to predict the uncertainty achievable by sensing some location (or combination of locations) in the terrain, we make a probabilistic prediction of the appearance of the terrain to the sensor. Each cell in this map stores a probability estimate that the cell will be seen as occupied in the sensed map. We call this the probability map of the terrain. This mapping should encompass the errors present both in the generation of the global map and the expected errors in the new local map.

For the case of stereo vision, Matthies and Shafer (1987) has found that the position errors are well approximated by a two-dimensional normal distribution with the major axis aligned away from the cameras. We, thus, convolve the global map with two normal distributions, one representing the error in the local map and one representing the error in the global map. Therefore, must continue to vary the distribution as a function of the location in the environment. The expected error grows with the square of the distance to the camera. We must, therefore, allow the width of the normal distributions to vary with the position in the environment.

Note, however, that the error in the map is a function of the convolution window, and

\[ P(i, j) = \sum_{x=-W}^{W} \sum_{y=-W}^{W} P_{g}(x+i, y+j)N(x, y; i, j). \] (9)

Of course, the instances of \( N(x, y; i, j) \) in (8) and (9) will be somewhat different, since the expected standard deviations and orientations will be different for the points in the global map versus the local map.

4. Landmark Selection

Given the probability map of the terrain, we can now estimate the uncertainty that will result from pointing the range sensor at some location in the environment and performing localization using the visible terrain. This is performed by treating the corresponding terrain patch in the probability map as a potential local map and comparing it the previously generated global map using the uncertainty estimation equations described above. The landmark selection step considers every possible terrain patch of some specified size in the probability map and performs uncertainty estimation at the correct position of the terrain patch in order to determine the terrain patch that yields the lowest uncertainty. This operation can be time-consuming if it is not performed carefully. We use several methods to improve the efficiency of this method.

In our implementation, we approximate the general normal distribution used to model the sensor error as a rotationally-symmetric 2-D normal distribution. While the error due to stereo vision is much greater along the direction parallel to the camera axes, error in the robot’s knowledge of its orientation will increase the error in the perpendicular direction. The shape of the error surface will, thus, become less elliptical due to the uncertainty knowledge of the robot orientation. Furthermore, our experiments indicate that the precise shape of the distribution does not have a large effect on the landmark selected. The use of rotationally-symmetric normal distributions makes the convolution kernel separable and this allows the convolutions to be performed efficiently by examining the x- and y-directions sequentially. On the other hand, it is crucial to use a wider and flatter distribution at locations further from the sensor, in order to model the increase in error with distance. We, therefore, must continue to vary the distribution as a function of the location in the space.
We can make the computation even more efficient by discretizing the space of allowable standard deviations for the normal distributions in Eq. 7. This allows us to pre-compute each of the distributions prior to computing the probability map. In our implementation, we select ten standard deviations (related by powers of $\sqrt{2}$). For each position in the terrain, we select the distribution with the closest standard deviation to the desired value in order to estimate the probability map. This approximation allows the probability map to be computed quickly upon demand for a region of the terrain map.

Finally, we use dynamic programming to compute the likelihood function from Section 2.1 for each of the terrain regions considered as a possible landmark for performing localization. This is performed at the optimal localization position for each possible landmark and the neighboring locations in the pose space in order to apply the uncertainty estimation techniques from Section 2.3. The terrain landmark yielding the lowest localization uncertainty is selected as the localization target.

The overall computation required by this method, including generation of the probability map, is linear in the number of cells examined in the occupancy grid representation of the global map. Note that when the occupancy grid represents a wide field of terrain, only those positions relatively close to the robot need to be examined, since the localization accuracy falls off with the distance of the landmark selected, in general. In experiments on a 10.28 m $\times$ 6.44 m map with square cells 2 cm wide, landmark selection requires approximately 2 seconds on a 333 MHz Sun UltraSPARC.

5. Results

This section describes the results of experiments where we have applied these techniques to perform localization using real and synthetic data.

5.1. Synthetic Data

An example using a synthetically generated elevation map is shown in Fig. 2. This case models a scenario where the robot is moving in a terrain consisting of rocks of various sizes, as a rover would encounter on the surface of Mars. The positions near large rocks are considered to be good targets, as shown by the uncertainty scores in Fig. 2(b). The target that is chosen is a position that contains not only a large rock, but also smaller rocks that are useful in performing the localization. The scale of this map is also 10.28 m $\times$ 6.44 m, with square cells 2 cm wide.

In order to test the localization performance when using the target selection techniques, we simulated localization problems by sampling local maps from the

![Figure 2](image)

**Figure 2.** Landmark selection example. The small boxes are the (interchangeable) robot beginning and ending positions. The selected target region is marked by a larger box. (a) Digital elevation map. (b) Estimated uncertainty with each possible landmark location. Light values are low uncertainties. (c) Three-dimensional terrain map.
distribution specified by the probability map of the terrain and then performing localization against the global map. This experiment selected robot positions at random from the terrain with the constraint that the beginning and ending positions were 5 m apart. We next performed target selection and, finally, localization using the selected target. In addition, we tested localization by sensing a target at the position directly between the robot starting and ending positions, and eight other targets on an evenly spaced grid around this position.

The results of this experiment are dramatic. When the target selection techniques were used in 1000 trials with the terrain shown in Fig. 2, the localization process found the qualitatively correct position in 97.8% of the trials, where the correct position was said to be found if the localization error was below 10 cm. However, when target selection was not used, the localization succeeded in only 29.5% of the trials, since much of the terrain provides little useful information for localization. In addition, the successful cases were, on average, 15.3% closer to the correct localization result when target selection was used, with an average error of 1.57 cm. This experiment demonstrates that target selection was not only useful in reducing the localization uncertainty, but also important in obtaining the correct qualitative position.

5.2. Real Data

We have implemented these techniques to run on-board the Rocky 7 rover prototype (Fig. 1). The techniques run in real-time and require approximately 30 seconds on a Motorola 68060 CPU to perform landmark selection and rover localization (this neglects the time necessary to point the sensor) using a 10 m × 10 m global map represented by an occupancy grid with square cells 2 cm wide. In this scenario, it is useful to capture stereo images to one side of the desired direction of travel, since landmarks directly between the current rover position and desired goal often have range shadows overlapping the terrain that is visible from the other direction. An example of the use of these techniques to perform target selection for localization is shown in Fig. 3. In this test, the rover captured six images to the left of the direction of desired travel. The goal position is at the right side of the top right image in Fig. 3(a) approximately 5 m from the initial rover position.

Figure 3(b) shows the terrain map that was generated on-board the rover using the stereo data in this test. As can be observed, there is a significant amount of noise in the terrain map due to errors in stereo range estimation. The landmark selection techniques were still able to select a large rock (near the center of the top middle image) as the best landmark for performing localization. The rover was then commanded to move to the goal position using autonomous obstacle avoidance during the traverse. When the rover state estimate became close to the commanded position, the rover halted. In this case, the rover traverse was short of the desired location owing to errors in the odometry. Localization was then performed using the selected landmark and the correct position was determined with an error below 5 cm.

Similar experiments were conducted at several locations in the JPL Mars Yard, including cases with more complex terrain. In each case, a useful landmark was selected for localization and an improved localization estimate was achieved. While an insufficient number of experiments (less than 10) were performed to draw reliably statistical information from, these experiments suggest that our target selection methodology is robust to the noise and other errors that occur in real data.

5.3. Sensor Uncertainty Field

Our approach to predicting terrain appearance and performing uncertainty estimation can be used to generate a sensor uncertainty field (Takeda et al., 1994) for a known terrain map. This field is the expected distribution of error in the sensed robot position as a function of the robot location. While, in general, the uncertainty will depend on the path taken to each position, we consider the uncertainty as a function of only the robot position.

We can, of course, compute the sensor uncertainty field using a brute-force method, where the best landmark is selected for each location of the robot and the resulting expected uncertainty is stored for each. Unfortunately, this process requires much computation. Note, however, that the uncertainties change slowly as the robot position that is examined is moved in the pose space. Our strategy is to first sample the pose space at a coarse resolution and then examine locations of interest, such as those that yield low uncertainties, subsequently at a finer resolution.

Figure 4 shows an example where a sensor uncertainty field has been generated for the terrain shown in Fig. 2. As expected, lower uncertainties occur near large rocks. However, the uncertainty grows large at
locations directly on top of rocks, since we use a model that does not allow the robot to view the terrain directly beneath it.

6. Summary

We have described a method to select the location to sense for performing mobile robot localization by matching terrain maps. The localization method that we use constructs a likelihood function in the space of possible robot positions. The uncertainty is estimated for localization using a local map by fitting a normal distribution to the likelihood surface generated in the pose space. We select the best landmark for localization by minimizing the expected uncertainty in the robot localization. In order to predict the uncertainty obtained by localization using various landmarks, our method constructs a probabilistic representation of the terrain expected to be sensed at any position in the global map. Treating the patches of this “probability map” of the terrain as a local map allows the uncertainty expected by sensing the terrain patch to be estimated using the
surface fitting techniques. We have applied this technique to robot localization in rocky terrain with good results.

Acknowledgments

The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This paper is an expanded version of a paper that previously appeared in the International Conference on Robotics and Automation (Olson, 2000c).

Notes

1. If the global map was a probabilistic representation that modelled the errors, we would use only the second of these convolutions. However, we have used a binary representation for the map in previous work (Olson, 2000a)

2. Note that the probability map has continuous values, while the original algorithm operated on a binary-valued local map. The only change to the algorithm is that we must multiply the score for each position in the probability map by the value of the probability map at that position.

References


Clark Olson received the B.S. degree in computer engineering in 1989 and the M.S. degree in electrical engineering in 1990, both from the University of Washington, Seattle. He received the Ph.D. degree in computer science in 1994 from the University of California, Berkeley. After spending two years doing research at Cornell University, he moved to the Jet Propulsion Laboratory, where he spent five years working on computer vision techniques for Mars rovers and other applications. Dr. Olson joined the faculty at the University of Washington, Bothell in 2001. His research interests include computer vision and mobile robotics.