Proceedings of 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation July 29 - August 1, 2001, Banff, Alberta, Canada

Collaborative Exploration for Map Construction

Ioannis Rekleitis[1], Robert Sim[1], Gregory Dudek[1], and Evangelos Milios[2]

[1]Centre for Intelligent Machines, McGill University
3480 University St, Montréal, Canada H3A 2A7
[2]Faculty of Computer Science, Dalhousie University
Halifax, Nova Scotia, Canada

Abstract

We consider the problem of map learning while maintaining ground-truth pose estimates. Map learning is important in tasks that require a model of the environment or some of its features. As a robot collects data, uncertainty about its position accumulates and corrupts its knowledge of the positions from which observations are taken. We address this problem by employing cooperative localization*; that is, deploying a second robot to observe the other as it explores, thereby establishing a* virtual tether*, and enabling an accurate estimate of the robot's position while it constructs the map. This paper presents our approach to this problem in the context of learning a set of visual landmarks useful for pose estimation. In addition to developing a formalism and concept, we validate our results experimentally and present quantitative results demonstrating the performance of the method.*

1 Introduction

Many robotic tasks require that the robot learn a representation, or map, of some aspect of the environment. Examples of such maps include measures of radiation hot-spots, magnetic declination, sonar representations and visual maps [1, 2, 3]. A significant issue faced by many map-building schemes is the accumulation of positional error as the robot collects observations from the environment. That is, as the robots collects successive measurements the certainty of their pose estimates with each new measurement decreases. In some cases where the observation lies on a high-dimensional manifold, correlation between dimensions allows for an expectation-maximization approach to correcting the observation poses [4, 5]. However, it is often the case that either there is insufficient geometric constraint in the observations to produce confident pose estimates even *post hoc*, or that the computational cost of making the appropriate inferences is infeasible. Other methods such as Kalman filtering can reduce the severity of the problem, but certainly do not eliminate it.

Our approach to the mapping problem entails the use of multiple robots working together. Several authors have considered the use of marsupial robots or robot teams either in theory or practice [6, 7]. Our work seeks to exploit robot collaboration for an explicitly quantitative mapping task. As such, it addresses algorithmic issues, implementation questions and design tradeoffs, and a experimental validation.

Figure 1: The two robots during the exploration of our laboratory.

This paper addresses the problem of establishing accurate pose estimates in the context of robotic mapping. The pose estimates can be used to collect accurate calibrated measurements to be used in their own right, or as a precursor to a system that builds a map to be used subsequently. The robot collecting measurements for the map operates in concert with a second robot that acts as an active observer. In our *cooperative localization* scheme, this second robot tracks the motions of the first as it collects data and provides it with the information required to prevent odometric error from accumulating. In this sense, a *virtual tether* is established between the two robots that enables the task of mapping to be accomplished without significant error. In principle, more than one of these active observers could be used simultaneously, although this is not elaborated in this paper. Beyond presenting the details of the approach and its implementation, this paper provides a quantitative evaluation validating the effectiveness of this methodology.

In prior work we have proposed the use of a multi-robot search strategy for a large environment which is either larger than the range a sensor can span [8], or small enough to permit sensing across the diameter of the environment [9]. In other work we have considered the use of attention-based learning for the construction of maps for pose estimation [10]. In this paper, we consider a synthesis of these approaches in a new problem domain, and examine the experimental performance of the hybrid approach.

The remainder of this paper is structured as follows: Section 2 discusses related work that addresses the problem of minimizing localization error during exploration and Section 3 describes our approach to the problem. We then discuss a particular application of our approach to the task of visual landmark learning in Section 4 and experimental results are presented in Section 5 Finally, we discuss open questions and future directions in Section 6.

2 Related Work

The problem that we have described is closely related to the problem of simultaneous localization and map-building, wherein the robot is tasked to explore its environment and construct a map. The advantages of collaborative behavior have been examined extensively in the context of biological systems [11, 12].

In the context of terrain coverage, in particular, Balch and Arkin were among the first to quantitatively evaluate the utility of inter-robot communication [13]. Mataric was another pioneer in considering utility of inter-robot communication in space coverage [14]. Dudek, Jenkin, Milios and Wilkes proposed a multi-robot mapping strategy akin to that proposed here, but they only considered certain theoretical aspects of the approach as it applied to very large groups of robots. Several authors have also surveyed the range of possible approaches for collaborative robot interactions [15, 16, 17, 18].

A number of authors have considered pragmatic multi-robot map-making in particular. Most existing approaches operate in the sonar domain, where it is relatively straightforward to transform observations from a given position to expected observations from nearby positions, thereby exploiting structural relationships in the data [19, 20, 21, 7]. These approaches successfully apply the probabilistic *expectation maximization* (EM) paradigm[22] to the task by iteratively refining the map and the estimates of the observation points.

In other work, Rekleitis, Dudek and Milios have demonstrated the utility of introducing a second robot to aid in the tracking of the exploratory robot's position[8]. In that work, the robots exchange roles from time to time during the exploration of a polygon-shaped world, thus serving to minimize the accumulation of odometry error. The authors refer to this procedure as *cooperative localization*. This paper builds on these results by Rekleitis *et al* by considering the task of exploring the visual domain. In the following section, we describe the method employed for tracking the position of the robot as it explores.

3 Robot Tracker

We have constructed a tracking device that can estimate the position and orientation of a mobile robot relative to a base robot equipped with a laser range-finder. The motion planning strategy is such that at any time one of the robots is stationary while the other robot is moving. The stationary robot acts as an artificial landmark in order for the moving robot to recover its pose with respect to the stationary one. Therefore, a detectable landmark is provided without any modification of the environment. We call this approach *Cooperative Localization*. Different types of sensors could be used depending on the required precision of the specific task. In earlier work a visual tracker with a helical pattern on the target robot was used, resulting into a 3-5cm accuracy in the position and a $3-7^{\circ}$ accuracy in the orientation [8, 9]. Currently we employ an *AccuRange* laser range-finder mounted on one robot and a three plane target mounted on the observed robot (see Figure 1). The target is a set of three vertical planes extending from the center of the target at three distinct angles (approximately $100^{\circ}, 120^{\circ}, 140^{\circ}$). At any given orientation of the target robot at least two vertical planes are “visible”. The intersection of the two planes define a unique point in a fixed position with reference to the observed robot. Further on the angle between the two planes combined with their orientations provides an estimate for the orientation of the robot.

The precision of the system AccuRange-target is much higher than the precision of the visual tracker. The position estimation is accurate to half a centimeter and the orientation estimation error is below one degree.

The Robot Tracker returns a triplet of $T = < \rho\ \phi\ \theta >$ that represent: ρ the distance between the two robots, ϕ the angle at which the observing robot sees the observable robot (eg. in Figure 2 the angle the stationary robot sees the moving robot), and θ the orientation of the observed robot as observed by the observing robot (eg. the orientation of the moving robot in Figure 2). As can be seen in Figure 2 both configurations are feasible (any of the two robot could observe). If the stationary robot is equipped with the laser then the Pose ($\mathbf{X_m}$) of the moving robot is given by equation 1, where $< x_s, y_s, \theta_s >$ is the pose of the stationary robot. If the moving robot is equipped with the laser then its Pose ($\mathbf{X'_m}$) is given by equation 2.

$$\mathbf{X_m} = \begin{pmatrix} x_m \\ y_m \\ \theta_m \end{pmatrix} = \begin{pmatrix} x_s - \rho * \cos(\theta_s + \phi) \\ y_s - \rho * \sin(\theta_s + \phi) \\ \pi + \phi + \theta_s + \theta \end{pmatrix} \quad (1)$$

$$\mathbf{X'_m} = \begin{pmatrix} x'm \\ y'm \\ \theta'm \end{pmatrix} = \begin{pmatrix} x_s - \rho * \cos(\phi - theta_s) \\ y_s - \rho * \sin(\phi - \theta_s) \\ \pi + \phi + \theta_s - \theta \end{pmatrix} \quad (2)$$

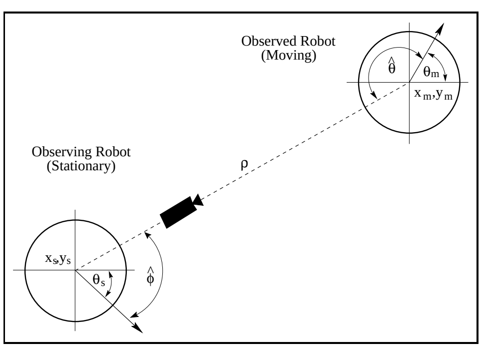

Figure 2: Observation of the Moving Robot by the Stationary Robot. Note that the “camera” indicates the robot with the *Robot Tracker*

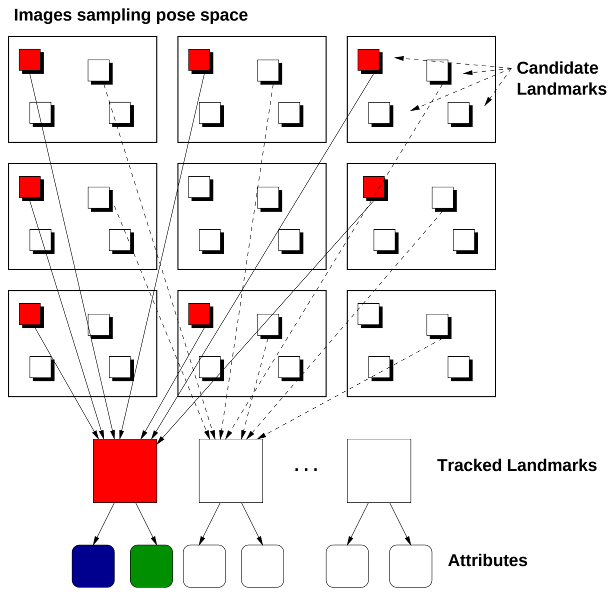

Figure 3: The off-line training method. Images (large rectangles) are collected sampling the pose space. Landmarks are extracted from the images and matched across the samples. The *tracked landmarks* are parameterized as a function of pose and saved for future pose estimation.

4 Application: Landmark Learning

In this section we demonstrate the effectiveness of our approach as it applies to the problem of learning visual landmarks which are useful for the task of pose estimation. The tracker is employed to provide “ground truth” positions while the robot explores the visual environment. We employ the landmark learning framework described in [23] and [10], which tracks the set of points output by an arbitrary model of visual attention and attempts to construct a representation of the landmark as a function of the pose of the robot. Such a representation can then be later exploited for the task of estimating the pose of the robot in the absence of the second robot or the tracker.

The learning method is depicted in Figure 3 and operates as follows (refer to the cited work for further details):

1. **Exploration:** One robot tracks the other as it collects images sampling a range of poses in the environment. The pose at which each image is taken is recorded as the estimate given by the tracker.

2. **Detection:** *Landmark candidates* are extracted

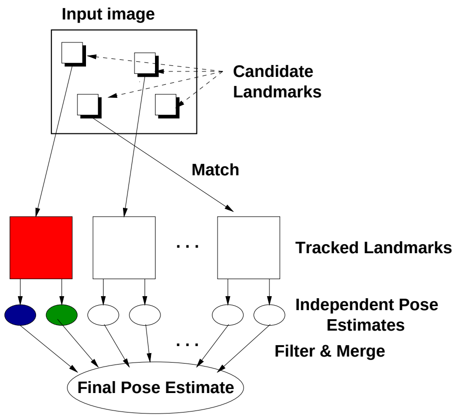

Figure 4: Pose estimation based on learned visual landmarks. Landmarks (small squares) are extracted from the current camera observation and matched to the previously learned tracked landmarks. Each match generates a pose estimate, which are filtered and combined to generate a final pose estimate.

from each image using a model of visual attention.

3. **Matching:** *Tracked landmarks* are extracted by tracking visually similar candidate landmarks over the configuration space.

4. **Parameterization:** The tracked landmarks are parameterized on the basis of a set of computed landmark attributes (for example, position in the image, intensity distribution, edge distribution, etc), and then measured in terms of their *a priori utility* for pose estimation.

5. The set of sufficiently useful tracked landmarks is stored for future retrieval.

When the robot requires a pose estimate without the aid of the tracker, it can obtain a camera image and extract candidate landmarks using the same model of visual attention that was employed during training. The candidates are matched visually to the learned landmarks and the previously constructed pose-dependent representation of each matched landmark contributes to the final pose estimate, as illustrated by Figure 4.

5 Experimental Results

In this section we present the results of deploying the tracking method for the task of landmark learning. Our environment consisted of a laboratory partitioned into two "rooms", with an open doorway connecting them (Figure 5). At the outset, one robot remained stationary while the other used a seed-spreader exploration procedure [24] across the floor, taking image samples at 40cm intervals. When the robot had completed the first room, it moved beyond the door and the stationary robot followed it to the threshold, where it remained stationary while tracking the exploratory robot as it continued its exploration of the second room.

Figure 5: Views of the two "rooms" as seen by the robot.

The trajectory of the exploratory robot was defined at the outset by a user. However, as the robot explored, accumulated error in odometry resulted in the robot straying from the desired path. The tracking estimate of the stationary robot was provided to the moving robot in order to correct this accumulated error. Figure 6 plots the uncorrected odometric trajectory (plotted as 'x') and the actual trajectory of the robot, as measured by the tracker (plotted as 'o'). For the sake of clarity, the odometric error was reset to zero between the first and second rooms. The accumulated odometric error is plotted for the second room versus total distance traveled in Figure 7.

Once image samples were obtained using the tracker estimates as ground truth position estimates,

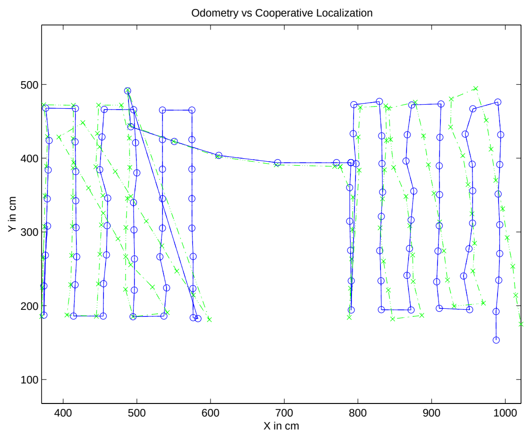

Figure 6: Odometric (x) vs Tracker-corrected (o) trajectories of the robot.

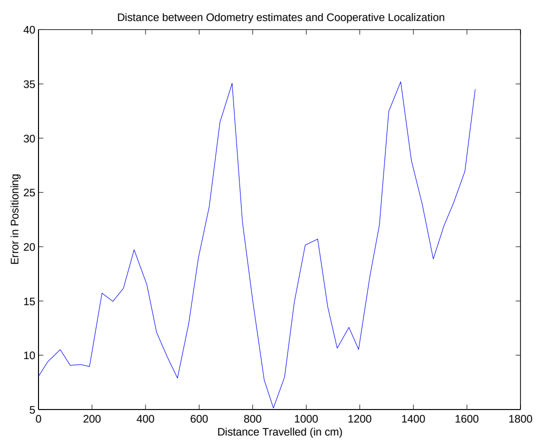

Figure 7: Odometric error versus distance traveled.

it was possible to apply our landmark learning framework to the image samples in order to learn a mapping between appearance-based landmarks and the pose of the robot. Figure 8 shows the discrepancies between the pose estimates from the tracker (marked as circles) and the vision-based estimator (marked as x's). At each position, the two 2-D projections of the alternative pose estimates are joined by a line. While the tracker is clearly more accurate, the utility of the image-based landmarks is sufficient for situations where only one robot is present.

Our final experiment involved navigating the robot to a series of random positions and acquiring image- and tracker-based pose estimates, which are plotted together in Figure 9. This final experiment illustrates the use of a multi-sensor estimator in removing outliers. Assuming that the tracker-based position is correct, the mean error in the image-based estimate was 33cm, a large part of which can be attributed to the two significant outliers from nearly the same position.

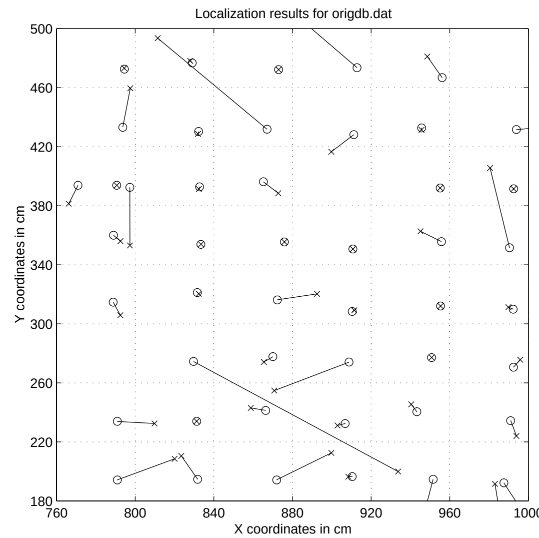

Figure 8: Tracker estimates vs Vision-based estimates for training images.

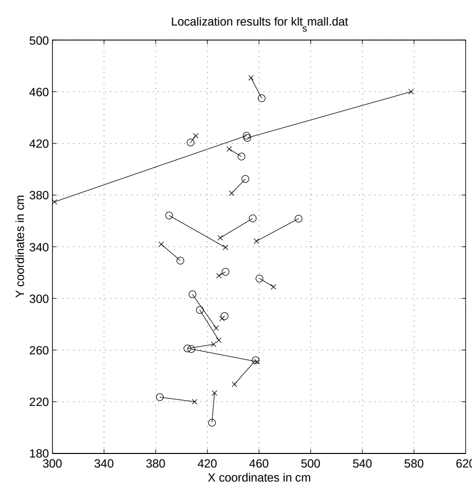

Figure 9: Tracker estimates vs Image-based estimates for a set of 21 random positions.

6 Conclusions

We have presented a method for the automatic mapping of an arbitrary environment which utilizes *cooperative localization* in order to maintain a *virtual* *tether* between two robots as one explores the environment and the other tracks its pose. The method relies on a mounted target whose pose can be estimated using a laser range-finder. The need for such an approach to maintaining a "ground truth" estimate of the exploring robot is validated by the magnitude of the accrued odometric error in our experimental results. Furthermore, we validate the utility of a set of learned landmarks for localization when the second robot cannot be deployed.

References

[1] B. Kuipers and Y.-T. Byun, "A robot exploration and mapping strategy based on a semantic hierachy of spatial representations", *Robotics and Autonomous Systems*, vol. 8, pp. 46–63, 1991.

[2] Howie Choset and Joel Burdick, "Sensor based planning, part 11: Incremental construction of the generalized voroni graph", in *Proc. of IEEE Conference on Robotics and Automation*, Nagoya, Japan, May 1995, pp. 1643 – 1648, IEEE Press.

[3] F. Dellaert, W. Burgard, D. Fox, and S. Thrun, "Using the condensation algorithm for robust, vision-based mobile robot localization", in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. June 1999, IEEE Press.

[4] Feng Lu and Evangelos E. Milios, "Optimal global pose estimation for consistent sensor data registration", in *Proc. of IEEE Conference on Robotics and Automation*, Nagoya, Japan, May 1995, pp. 93 – 100, IEEE Press.

[5] S. Thrun, D. Fox, and W. Burghard, "A probabilistic approach to concurrent mapping and localization for mobile robots.", *Autonomous Robots*, vol. 5, pp. 253–271, 1998.

[6] Robin R. Murphy, Michelle Ausmus, Magda Bugajska, Tanya Ellis, Tonia Johnson, Nia Kelley Jodi Kiefer, and Lisa Pollok, "Marsupial-like mobile robot societies", in *Agents*, 1999, pp. 364–365.

[7] R. Kurazume and S. Hirose, "Study on cooperative positioning system - optimum moving strategies for cps-iii", in *Proc. IEEE Int. Conf. on Robotics and Automation*, IEEE, Ed., 1998, vol. 4, pp. 2896–2903.

[8] Ioannis M. Rekleitis, Gregory Dudek, and Evangelos Milios, "Multi-robot collaboration for robust exploration", in *Proceedings of International Conference in Robotics and Automation*, San Francisco, CA, April 2000.

[9] Ioannis M. Rekleitis, Gregory Dudek, and Evangelos Milios, "Multi-robot collaboration for robust exploration", *Annals of Mathematics and Artificial Intelligence*, vol. 0, pp. 1–34, 2000, To appear.

[10] R. Sim and G. Dudek, "Learning and evaluating visual features for pose estimation", in *Proceedings of the Seventh IEEE International Conference on Computer Vision(ICCV)*, Kerkyra, Greece, sept 1999, IEEE Press.

[11] Niko Tinbergen, *The animal in its world; explorations of an ethologist, 1932-1972*, Harvard University Press, 1972.

[12] B.L. Partridge, "The structure and function of fish schools", *Scientific American*, pp. 114–123, June 1982.

[13] Tucker Balch and Ronald C. Arkin, "Communication in reactive multiagent robotic systems", *Autonomous Robots*, vol. 1, no. 1, pp. 27–52, 1994.

[14] M.J Mataric, "Environmental learning using a distributed representation", *IEEE International Conference on Robotics and Automation*, vol. 1, pp. 402–406, 1990.

[15] Gregory Dudek, Michael Jenkin, Evangelos Milios, and David Wilkes, "A taxonomy for multi-agent robotics", *Autonomous Robots*, vol. 3, pp. 375–397, 1996.

[16] Y. U. Cao, A. S. Fukunaga, A. B. Kahng, and F. Meng, "Cooperative mobile robotics: Antecedents and directions", in *Proc. IEEE/RSJ IROS*, Pittsburgh, PA, 1995, vol. 1, pp. 226–234.

[17] Gregory Dudek, "Environment mapping using multiple abstraction levels", in *Proceedings of the IEEE*, Nov. 1996, vol. 84.

[18] Gregory Dudek and Michael Jenkin, *Computational Principles of Mobile Robotics*, Number ISBN: 0521568765. Cambridge University Press, May 2000.

[19] John J. Leonard and Hugh F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons", *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 376–382, June 1991.

[20] D. Fox, W. Burgard, and S. Thrun, "Active markov localization for mobile robots", *Robotics and Autonomous Systems*, 1998, To appear.

[21] Wolfram Burgard, Dieter Fox, Mark Moors, Reid Simmons, and Sebastian Thrun, "Collaborative multi-robot exploration", in *Proceedings of the IEEE International Conference on Robotics and Automation*, San Francisco, CA, May 2000, IEEE Press.

[22] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm (with discussion)", *Journal of the Royal Statistical Society series B*, vol. 39, pp. 1–38, 1977.

[23] R. Sim and G. Dudek, "Learning visual landmarks for pose estimation", in *Proceedings of the IEEE International Conference on Robotics and Automation(ICRA)*, Detroit, MI, May 1999, IEEE Press.

[24] V. Lumelsky, S. Mukhopadhyay, and Kang Sun, "Sensor-based terrain acquisition: The 'sightseer' strategy.", in *Proceedings of the IEEE Conference on Decision and Control Including The Symposium on Adaptive Processes*, IEEE Service Center, Piscataway, NJ, USA, 1989, IEEE, vol. 2, pp. 1157–1161, IEEE.