Information Based Exploration with Panoramas and Angle Occupancy Grids

Daniel Mox, Anthony Cowley, M. Ani Hsieh, and C. J. Taylor

Abstract In this work we present a multi-robot information based exploration strategy with the goal of constructing high resolution 3D maps. We use a Cauchy-Schwarz Quadratic Mutual Information (CSQMI) based objective which operates on a novel angle enhanced occupancy grid to guide robots in the collection of RGBD panoramas, which have been shown to provide memory efficient high quality representations of space. To intelligently collect panoramas, we introduce the angle enhanced occupancy grid which emphasizes perspective in addition to coverage, a characteristic we believe results in the construction of higher quality maps than traditional occupancy grid methods. To show this, we conduct simulations and compare our approach with frontier exploration. Using our angle enhanced occupancy grid, only 11.4% of decimeter wall segments were covered by fewer than 20 pixels as compared with 33.5% for the frontier method.

1 Introduction

A central pillar of robotics research is the development of efficient autonomous mapping and exploration strategies with the attendant development of suitable representations of the environment. Suitable representations serve as a prerequisite for fundamental tasks such as exploration and localization, with a plethora of forms emerging to suit each case. Metric maps that render spatial attributes in terms of their location in a shared coordinate frame have enjoyed immense popularity in the literature. Examples of *de facto* standards for metric maps include occupancy grids for exploration [4] and landmark maps for simultaneous localization and mapping (SLAM) [10]. On the other hand, topological maps seek to relate semantic informa-

Daniel Mox and M. Ani Hsieh
Drexel University, Philadelpha, PA e-mail: {dcm64,mhsieh1}@drexel.edu

Anthony Cowley and C. J. Taylor
University of Pennsylvania, Philadelpha, PA e-mail: {acowley,cjtaylor}@seas.upenn.edu

tion about the environment gleaned from sensor observations, and have been used to aid in rapid exploration [3] and efficient task allocation [8].

In typical SLAM scenarios, localization and mapping are the motivating factors of the entire system. However, choices made in support of those goals may not be well-suited for other uses of map data. For example, in landmark based maps only a few patches of images representing regions of interest are extracted from source imagery, while in voxel based grids the fine details of a point cloud are lumped into a small number of cells. The effects of filtering out significant portions of sensor data have only been magnified with the emergence of low cost off the shelf RGBD cameras. These sensors combine depth and visual information that can be used to create dense, high fidelity three dimensional maps [11], but the torrent of available data must must be carefully navigated.

New approaches using RGBD cameras include the influential work of Newcombe in which dense mapping and tracking is achieved by fusing depth frames for surface reconstruction [7]. Henry employed surfels, *e.g.*, surface orientation, patch size, and color, gleaned from RGBD frames to build 3D maps of indoor environments [5]. Recent work by Taylor [9] focuses on panoramic views comprising multiple depth frames captured at a location. Panoramic depth images are advantageous for mapping because they provide both spatial information for motion planning and fine grained detail for object recognition. Furthermore, when compared to traditional grid based approaches, panoramic images have the potential to provide significantly more detail while consuming only a modest amount of memory resulting in better scaling characteristics.

In this work, we build upon [9] to develop a novel information maximizing exploration strategy for teams of robots to map an unknown environment. Recent advances in our understanding of information theory and how to apply it to the robot exploration problem have lead to a slew of new algorithms based on a powerful principle: the map provides hints about what observations to expect at different locations and vice versa. This notion is at the heart of mutual information, which seeks to quantify the amount of information one random variable contains about another [2]. Julian demonstrated that mutual information eventually drives the robot towards unknown space and used it as an objective function for autonomous exploration using a range sensor [6]. A similar approach was shown by Charrow who used Cauchy-Schwarz Quadratic Mutual Information (CSQMI), a metric closely related to mutual information, to guide a robot equipped with an RGBD camera to map an unknown environment [1].

Different from existing work, we present a novel spatial grid representation that emphasizes *perspective* in addition to coverage to guide robots in exploring unknown spaces while collecting useful panoramas to create a detailed map of the environment. We leverage existing grid maps and their simplifying power for reasoning about exploration tasks to guide a group of robots to collect panoramas which can then be used to construct a detailed map of the environment. In our strategy, we employ an occupancy grid based representation of the environment to enable each robot to execute a CSQMI based collaborative exploration strategy that uses minimal communication bandwidth. A high fidelity representation of the environment is then reconstructed using the collective sequence of panoramic images.

This paper is organized as follows: A detailed description of our methodology is provided in Section 2. Simulation and experimental results are presented in Section 3. We conclude with a summary of our work and a discussion on directions for future work in Section 4.

2 Methodology

The objective is to efficiently explore an unknown environment using multiple robots to collect high resolution panoramic images that can be used to produce a detailed, memory conscious representation of space.

Using panoramas to represent the environment offer a number of advantages over traditional grid based maps [9]. First, they capture the surface structure of interest and nothing more as free space is implicit in the representation, offering a high level of detail for the memory used as compared to an occupancy grid. While memory usage may not be of significant concern when the workspace is small, it quickly becomes a significant issue when exploring and mapping realistically detailed and expansive spaces. Our work capitalizes on the unique perceptual data provided by RGBD panoramas presented in [9]. By capturing such views of the space surrounding a robot at specific locations in the workspace, a detailed map of the environment can be created. Our work focuses on the development of a suitable exploration strategy and we refer the interested reader to [9] for the details on the synthesis of the panoramas and an overview of comparable state-of-the-art mapping techniques.

From [9], the panoramas are pieced together from a series of color and depth images collected at regular intervals while the robot turns in place. As such, the images can be stitched and refined locally on the robot, alleviating the need for high bandwidth networking. As such, our exploration strategy is decomposed into two main components: 1) collecting and storing panoramas for reconstruction and updating local maps maintained for planning purposes and 2) determining the next best location to collect more panoramas. We briefly describe our approach.

2.1 Angle Enhanced Occupancy Grid

While many exploration techniques focus on coverage of free space, the ideal map should also afford clear and diverse views of surfaces within that space. Consider a scenario where a robot is advancing down a hallway. A panorama is captured, and a sign on the wall just manages to fall within the cameras range. The free space has been identified and, from a traditional free space coverage standpoint, no more value can be gleaned from observing the robot's immediate surroundings. However, from a surface reconstruction standpoint, the face of the sign was captured at a shallow angle and may be rendered illegible in the resulting map. Here the utility of capturing space from advantageous perspectives is made plain as can be seen in Fig. 1.

To encourage perspective in addition to coverage, we employ an *angle enhanced* 2D occupancy grid such that each point in space contains four values representing the cardinal directions. As new observations are integrated into the map, the appropriate bin is updated according to the angle at which the cell was observed by the sensor. This creates an incentive for the robot to collect different views of the same physical space. Since free space contains the same information from all perspectives this binning strategy is not applied to unoccupied cells.

The principle behind this choice in map representation is that an occupancy grid naturally breaks environmental surfaces into segments. These segments are, at a fine enough scale, well approximated by star convex functions that we can drastically subsample by only considering a small number of view angles between sensor and surface. The primary tuning parameter then is the occupancy grid resolution, but not, as is typical, to directly capture finer geometry. Instead, occupancy grid resolution must be sufficient for the above convexity assumption to hold. Thus complex surface geometry that may be captured in an image complements the efficient, if coarse, geometric approximation offered by the underlying occupancy grid used for planning.

2.2 Exploration Planning

To select useful exploration goals, it is important to estimate how much future observations will tell us about the space they can cover. This notion is captured by computing the mutual information between a predicted sensor reading and the map given by

$$I_{MI}[m;z] = H[m] - H[m \mid z]. \tag{1}$$

Since a future measurement is implicitly conditioned on the robot's position, an objective function naturally arises wherein mutual information is calculated for a collection of candidate poses and used to determine the quality of a candidate panorama capture location.

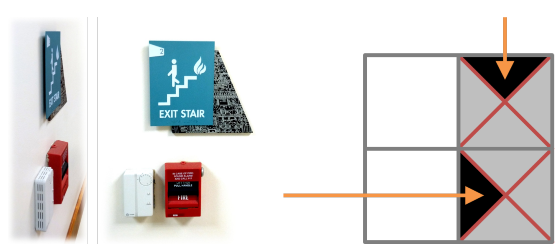


Fig. 1 Objects of interest on display in a hallway captured at different perspectives (left). Illustration of cells in an angle occupancy grid begin intercepted by beams, depicted as orange arrows (right).

In this work, we employ the Cauchy-Schwarz Quadratic Mutual Information (CSQMI) between a predicted measurement and the map to determine the next best position for the robot to obtain a panorama during its exploration of the space. In general, mutual information can be computationally expensive since it requires integrating the likelihood over entire space, often represented by a grid. Since a closed form expression for the CSQMI exists [1], it enables the quantification of the value of future measurements in a computationally efficient manner. Furthermore, CSQMI has been shown to produce similar results to mutual information [1]. Thus, when coupled with our angle enhanced 2D occupancy grid representation of the workspace for navigation purposes, the result is a scalable exploration strategy with well bounded computational and memory complexities for any given workspace. We briefly summarize the computation of CSMI using the proposed angle enhanced occupancy grid and refer the interested reader to [1] for the details.

Let z denote the set of random variables representing future sensor measurements that model the distance a beam at image pixel k travels before encountering an obstacle. Since our sensor measurement is a panoramic RGBD image, we evenly distribute k beams on the interval $[0,2\pi)$ originating at pose x. To model the noise inherent in a depth measurement, we use the piecewise normal distribution given by

$$p(z_k = z \mid d) = \begin{cases} \mathcal{N}(z-0,\sigma^2) & d \le z_{min} \\ \mathcal{N}(z-z_{max},\sigma^2) & d \ge z_{max} \\ \mathcal{N}(z-d,\sigma^2) & \text{otherwise} \end{cases} \tag{2}$$

where z_{min} and z_{max} are the minimum and maximum range of the sensor respectively and $\mathcal{N}(z-\mu,\sigma^2)$ is a Gaussian with mean μ and variance σ^2. While more complex beam based models exist [10], this simple model is sufficient for our purposes.

Using the beam model given by (2), a distribution over possible measurements denoted by $p(z_k)$ can be found by computing the marginal distribution over the list of c cells belonging to the map m intercepted by the beam

$$p(z_k) = \sum_{i=0}^{C} p(c = e_i) p(z_k \mid c = e_i) \tag{3}$$

where for $i > 1$, e_i means that the i^{th} cell is the first occupied cell in c, e_0 means that no cells in c are occupied, and $C = |c|$.

The CSQMI between the map, m, and a predicted observation, z, collected at a location in space, x, is expressed as

$$I_{CS}[m; z \mid x] = \frac{(\sum \int p(m,z \mid x) p(m) p(z \mid x) dz)^2}{\sum \int p^2(m,z \mid x) dz \sum \int p^2(m) p^2(z \mid x) dz} \tag{4}$$

where each sum is over all possible maps and the integrals are over all possible measurements.

Due to the camera's high resolution and the fact that panoramas comprise many images stitched together, the number of beams k in the measurement z_k could be quite large. However, exactly representing each pixel in a panorama with a beam is not necessary. Depending on the resolution of the grid, beams very close together often intercept many of the same cells which leads to the double counting of information gained from each beam. As such, we only consider the subset of beams that can be reasonably considered independent of one another given the range of the camera and the resolution of the grid. Assuming independence between elements of this subset of beams, (4) is computed by summing the individual contribution of each beam as follows

$$\begin{aligned} I_{CS}[c;z_k \mid x] &= \log \sum_{l=0}^{C} w_l \mathcal{N}(0,2\sigma^2) \\ &+ \log \prod_{i=1}^{C} (o_i^2 + (1-o_i)^2) \sum_{j=0}^{C} \sum_{l=0}^{C} p(e_j) p(e_l) \mathcal{N}(\mu_l - \mu_j, 2\sigma^2) \\ &- 2\log \sum_{j=0}^{C} \sum_{l=0}^{C} p(e_j) w_l \mathcal{N}(\mu_l - \mu_j, 2\sigma^2). \end{aligned} \tag{5}$$

Here $o_i = p(c_i = 1)$ is the probability that the i^{th} cell in c is occupied and $p(e_j)$ is the probability that the first occupied cell in c is c_j. Additionally, each w_l is calculated as follows

$$w_l = p^2(e_l) \prod_{j=l+1}^{C} (o_j^2 + (1-o_j)^2). \tag{6}$$

with $0 < l < C$, $w_0 = p(e_0)$, and $w_C = p^2(e_C)$. The final result then becomes

$$I_{CS}[m;z \mid x] = \sum_{i=0}^{k} I_{CS}[c;z_k \mid x]. \tag{7}$$

CSQMI is computed for a list of poses, χ, sampled from known free space. Fig. 2 shows the CSQMI reward surface computed for the corresponding workspace using a traditional 2D occupancy grid and using an angle enhanced occupancy grid. By incorporating the notion of perspective into the occupancy grid, the CSQMI reward surface computed using the angle enhanced grid results in better coverage of the all accessible sides of objects and obstacles.

To choose between the highest CSQMI poses, travel costs in the form of distance along a path calculated using A* from the robot to the candidate pose are used. Thus, the next best position is given by

$$x^\star = \operatorname*{arg\,max}_{x \in \chi} \frac{I_{CS}[m;z \mid x]}{Cost(x)} \tag{8}$$

where $Cost(x)$ is the path distance from the robot's current pose to the candidate goal, x. Integrating the travel cost results in emphasizing completion of locally accessible space before advancing towards uncharted territory.

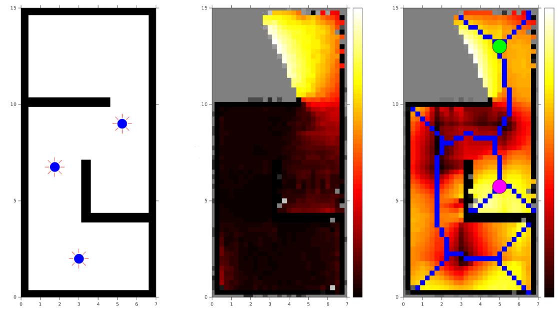


Fig. 2 A sample environment after three panoramas have been captured (left) and the resulting CSQMI reward surface for a traditional 2D occupancy grid (center) and the angle enhanced occupancy grid (right).

2.3 Multi-robot Strategy

From (4) we see that our approach provides a computationally efficient strategy to determine next best locations for measurements in a scalable way. Accordingly, our strategy is particularly well suited for small, resource constrained platforms since each robot only has to compute the CSQMI objective function once to determine the next best position to capture a panorama. This enables us to develop a simple coordination strategy where the single robot exploration strategy is executed in parallel by a team of robots.

The deployment of multiple robots speeds up the exploration process, especially when robots are tasked to focus on distinct regions of the workspace. Since travel costs are already accounted for during planning, our coordination strategy simply requires individual robots to share their local maps and goals to ensure panorama capture locations do not overlap. Assuming the relative pose of each agent is available, a small grid circumscribing the latest panorama is broadcasted to the other robots and integrated into their local maps. The active goals of other agents are also considered during planning to prevent multiple robots from traveling to the same region. We note that this communication strategy only requires the communication of local maps and goals which are only updated each time a panorama is captured and as such are only transmitted occasionally, reducing the required bandwidth. The result is a distributed strategy for cooperative mapping of an unknown environment. Critically, this approach relies on the exchange of small occupancy grids rather than accumulated imagery or detailed surface reconstructions. This is the difference between megabytes-per-second and kilobytes-per-hour in terms of communication cost.

3 Results

To validate our approach, simulations and experiments were conducted in a variety of indoor environments with teams of one to five robots. An Asus Xtion Pro Live RGBD camera was used to provide observations of the environment. Each panorama comprised of 36 images, one for every 10 degrees of rotation. Throughout our simulations and experiments, we assume the pose of each robot in a global coordinate frame is provided. While robot localization remains a non-trivial problem, our objective is validating the proposed exploration strategy. As such, we assume robot localization can be achieved via existing on-board localization methods.

3.1 Simulations

To evaluate the proposed angle aware exploration strategy, we compare our approach with the well established frontier method [12]. We use two simulated environments shown in Fig. 3a and 3b to compare the resulting exploration locations. Fig. 3c and 3d show the capture locations of panoramas in a simple indoor environment spanning a space of 8×14 m^2. High level results of each iteration of the simulation are summarized int Table 1.

The benefit of the angle enhanced occupancy grid approach may be seen by a detailed analysis of image coverage of the 2D surfaces in the Hallway environment.

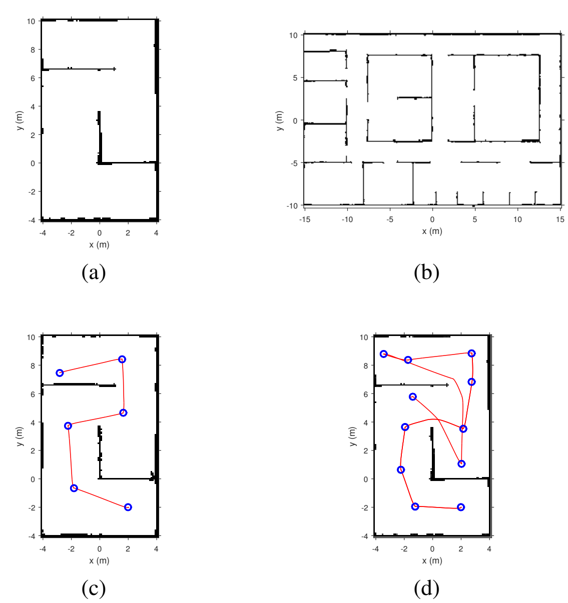


Fig. 3: (a) Curving hallway environment used with 1 robot; (b) larger office environment used with 5 robots; and simulation results for the curved hallway for frontier (c) and angle aware (d) planning methods. The red line shows the path of the robot and the blue circles represents panorama capture locations.

Table 1: Summary of the area covered and panoramas captured by each method in the Hallway (Fig. 3a) and Office (Fig. 3b) environments.

Environment	Method	Area Covered (m^2)	Panoramas Captured
hallway	angle aware	110.9	11
hallway	frontier	105.9	6
office	angle aware	572.0	140
office	frontier	571.3	38

As a proxy for data quality, we consider the maximum image size of every 10 cm segment of wall across all simulated panoramas using a linear imaging resolution of 8229 pixels for each 360° panoramic image as used by the experiments in [9]. Visibility is calculated for each decimeter as a whole by ray-casting from the capture location to the center of a wall segment. This is a simplification of the benefits of multiple observations, but is immediately applicable to common tasks such as optical character recognition of writing on walls, or face recognition applied to pictures on walls. Image size relates to our approach in so far as shallow observation angles correspond with lower resolution imaging of a surface. Put plainly, though we are not optimizing specifically for this metric, we expect some correlation. Note that while the numbers used throughout our analysis are arbitrary, they provide a context to compare each method that proves valuable regardless of the exact image resolution chosen.

For the set of simulated experiments shown in Fig. 3c and 3d, the frontier approach fully explored the Hallway environment after collecting 6 panoramas, while the angle aware approach selected 11 locations to faithfully capture all surfaces. With these panorama collection locations, we can calculate that 33.5% of wall segments observed during the frontier exploration strategy were captured by fewer than 20 pixels in any given panorama. For the angle enhanced occupancy grid approach, only 11.4% of wall segments were observed below the same threshold resolution. If we evenly subdivide the frontier exploration trajectory to result in an equivalent number of panoramas ("augmented frontier"), 17.9% of wall segments are observed below the resolution threshold. This demonstrates that angle sensitivity in the exploration strategy results in *substantially* fewer "blind spots" along surfaces in the environment as it has more information with which to plan. A naïve increase in panorama locations alone, however, *does not* result in the same coverage gains.

The prevalence of low resolution blind spots in the augmented frontier strategy is visible in the overlayed histograms in the lower-right of Fig. 4. In this figure, the green histogram bars of the augmented frontier observations stand out on the left of the chart. Precisely where these blind spots arise may be seen in the top row of Fig. 4 where the light blue wall segments on the right side of the map approximately one third of the way up the image, for example, indicate poor coverage.

Lastly, Fig. 5 shows the panorama capture locations for a team of five robots deployed in the office environment shown in Fig. 3b. Initially each robot balances capturing local space with pushing into uncharted territory. As each robot explores more space, the simple assignment strategy ensures that each robot does not interfere with its neighbor.

3.2 Experiments

We have also conducted live experiments in the space shown in Fig. 6b which covers an area of approximately $5 \times 5\ m^2$. The two differential drive ground robots shown in Fig. 6a traversed the environment to collect panoramas, sharing maps and goal locations as described in Section 2.3. Each robot was equipped with an Asus Xtion Pro Live RGBD camera providing synchronized color and depth frames for the panoramas and spoofing a laser rangefinder for construction of the angle occupancy grid used during planning. All software was written in C++ and executed on

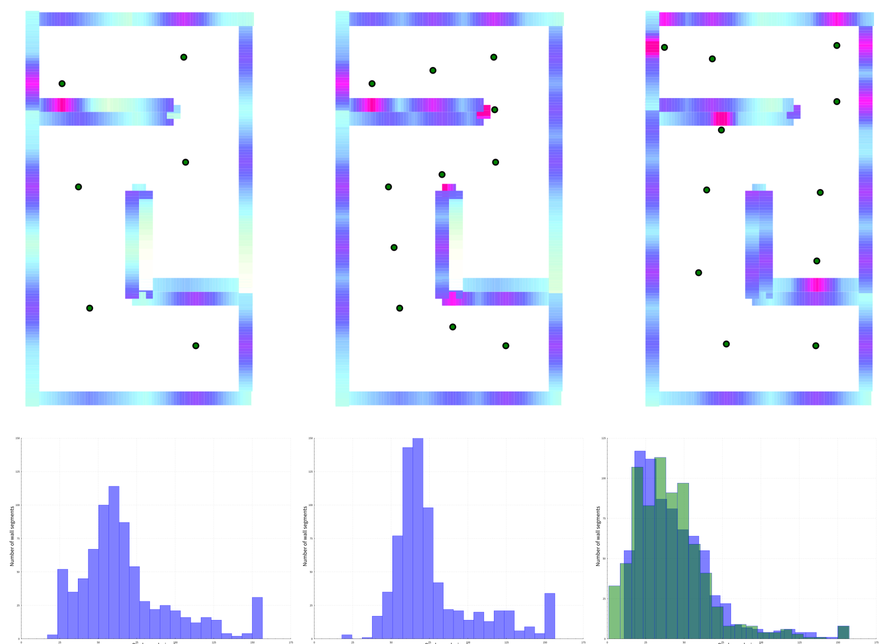


Fig. 4: Top row: map renderings with walls colorized by the resolution at which they were covered for frontier (left), augmented frontier (center), and angle aware (right) approaches. Hue indicates imaging resolution with red indicating high resolution, and blue indicating low resolution. In grayscale, the less saturated wall segments indicate poor coverage. Panorama capture locations are denoted by green circles. Bottom row: histograms of pixels per decimeter for augmented frontier (left), angle aware (center), and augmented frontier overlaid on angle aware (right).

an Odroid-XU4 single board computer on-board the robot running Linux and the Robotic Operating System (ROS). Localization was provided by an external motion capture system.

In order to predict future measurements and compute CSQMI, several parameters were specified. Panoramas were modeled as a collection of 360 beams distributed over the interval $[0, 2\pi)$. Given a map resolution of 0.05 m, increasing the number of beams only resulted in additional dependent beams which intercepted the same cells and would be factored out of the CSQMI calculation. Values of $z_{min} = 0.5\ m$ and $z_{max} = 4.5\ m$ were used for the minimum and maximum range of the Xtion, and the noise was modeled with $\sigma = 0.03$.

The results of the experiment can be see in Fig. 6c with the two robots beginning at $\{-2, 0\}$ and $\{-2, 1.5\}$ respectively. Starting at the same time, the team executed our exploration strategy simultaneously and collected a total of 13 panoramas. Operating in proximity, the team successfully avoided capturing panoramas from the same location demonstrating the effectiveness of our proposed multi-robot coordination strategy.

4 Conclusion and Future Work

Panoramas comprised of images collected from an RGBD sensor provide a new way to generate high fidelity representations of environments without demanding significant memory resources. This opens up the need for new exploration strategies that leverages the unique perceptual characteristics provided by the sensor. Our angle aware exploration strategy enables a team of robots to effectively map an unknown environment by searching for locations that yield high information gain while accounting for diverse perspectives when creating a comprehensive 3D map. Since the approach is computationally efficient, it can be used by a wide variety

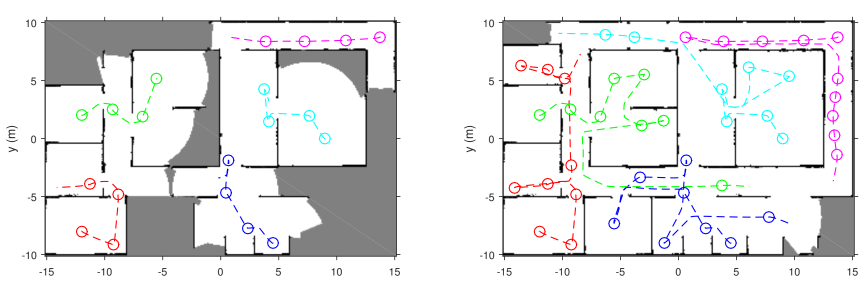
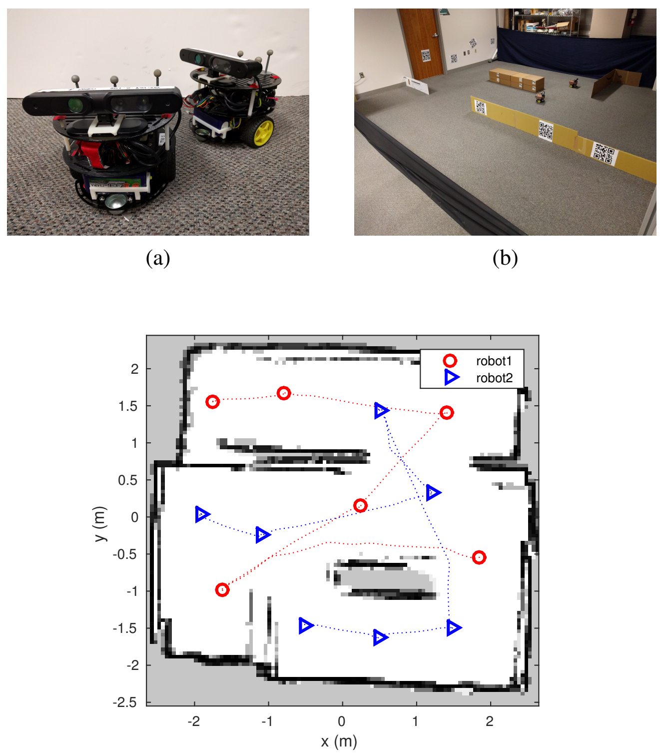


Fig. 5: Five robots collaboratively exploring an office environment after 4 panoramas (left) and after 10 panoramas (right).

of robotic platforms, even ones with limited computational resources. Furthermore, the approach places minimal demands on inter-agent wireless communication and computation at both the planning and coordination level.

As depth panoramas contain rich information about the environment, an alternative approach could be constructing and using them online directly in the planning phase. This has the added benefit that the final map is included in the planning loop which ensures some level of quality. Additionally, extending this approach to platforms not constrained to the plane such as unmanned aerial vehicles would require modifications to the angle enhanced occupancy grid employed. However, we believe the proposed exploration strategy will enable heterogeneous teams of robots to better leverage the distinct perceptual and mobility capabilities of the various sensing resources within the team. These are all directions we are pursuing for future work.



(a) (b)

(c)

Fig. 6: The ground robots (a), lab space used for live experiments (b), and experimental results showing the panorama capture locations for the two robot team (c).

Acknowledgement

The authors gratefully acknowledge the support of ARL grant W911NF-08-2-0004.

References

1. Benjamin Charrow, Sikang Liu, Vijay Kumar, and Nathan Michael. Information-theoretic mapping using Cauchy-Schwarz Quadratic Mutual Information. 2015 IEEE International Conference on Robotics and Automation (ICRA), pages 4791–4798, 2015.
2. Thomas M. Cover and Joy a. Thomas. Elements of Information Theory. 2005.
3. Anthony Cowley, Camillo J. Taylor, and Ben Southall. Rapid multi-robot exploration with topometric maps. Proceedings - IEEE International Conference on Robotics and Automation, pages 1044–1049, 2011.
4. Alberto Elfes. Using occupancy grids for mobile robot perception and navigation. Computer, 22(6):46–57, 1989.
5. Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments. Springer Tracts in Advanced Robotics, 79:477–491, 2014.
6. B. J. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. The International Journal of Robotics Research, 33(10):1375–1392, 2014.
7. Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. 2011 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011, (July 2016):127–136, 2011.
8. Cyrill Stachniss, Óscar Martínez Mozos, and Wolfram Burgard. Efficient exploration of unknown indoor environments using a team of mobile robots. Annals of Mathematics and Artificial Intelligence, 52(2):205–227, 2008.
9. Camillo J. Taylor, Anthony Cowley, Rafe Kettler, Kai Ninomiya, Mayank Gupta, and Boyang Niu. Mapping with Depth Panoramas. Intelligent Robots and Systems (IROS 2015), 2015 IEEE/RSJ International Conference on, pages 6265–6272, 2015.
10. Sebastian Thrun, Wolfram Burgard, and Dieter Fox. Probabilistic robotics. MIT press, 2005.
11. Thomas Whelan, Michael Kaess, Maurice Fallon, Hordur Johannsson, John Leonard, and John McDonald. Kintinuous: Spatially extended kinectfusion. 2012.
12. Brian Yamauchi. Frontier-based exploration using multiple robots. In Proceedings of the second international conference on Autonomous agents, pages 47–53. ACM, 1998.