Adaptive Exploration and Sampling by Heterogeneous Robotic Team

Sandeep Manjanna[1], Alberto Quattrini Li[2], Ryan N. Smith[3], Ioannis Rekleitis[2] and Gregory Dudek[1]

I. INTRODUCTION

Physical sampling of water for off-site analysis is necessary for many applications like monitoring the quality of drinking water in reservoirs, understanding marine ecosystems, and measuring contamination levels in fresh-water systems. Robotic sampling enables to strategically collect water samples based on real-time measurements of physical and chemical properties gathered with onboard sensors. In this paper, we present a multi-robot, data-driven, water-sampling strategy, where autonomous surface vehicles plan and execute water sampling using the chlorophyll density as a cue for plankton-rich water samples.

Our focus is to address the problem of monitoring a region and collecting water samples with emphasis on selecting good sampling locations, but without *a priori* knowledge of where these locations might be. We use a heterogeneous robotic team composed of two robotic boats, an *explorer* that can measure variables that suggest sample utility and a *sampler* that can collect physical samples (Figure 1). Das et al. [1] proposed a probabilistic method for a single AUV that can monitor and sample. In our case, we divide the task between two robots. This provides an efficient trade off between system complexity, payload capacity, and run time, besides improving the quality of the collected samples – where quality is expressed as the sum of measured values over samples collected.

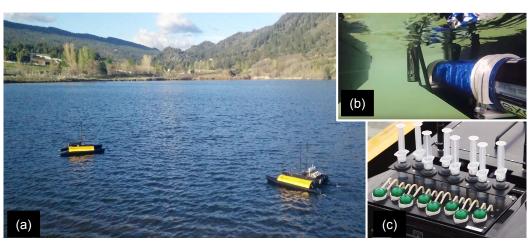

Fig. 1. Two Clearpath Heron ASVs (a), one equipped with a water quality sensor (b), another with a water sampling apparatus (c).

In particular, such a task leads to two related subproblems: *exploration* and *sampling*. We propose an exploration strategy for the *explorer* – the robot with the water quality sensor – that makes real-time observations to create a preliminary map. The sampler is then informed about the potential locations for sampling. Our method is based on the concept of *frontier*-based exploration, similar to that introduced by Yamauchi [2] for indoor map building and exploration. The robot decides according to the latest information and this approach scales well with the size of the region, unlike some common coverage approaches that employ a lawn-mower coverage pattern. Notice that the absence of prior information on the spatial distribution of the data of interest prevents us from using alternative powerful selective coverage methods [3], [4].

[1]S. Manjanna and G. Dudek are with School of Computer Science, McGill University, Montreal, Canada msandeep,dudek@cim.mcgill.ca

[2]A. Quattrini Li and I. Rekleitis are with the Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA albertoq,yiannisr@cse.sc.edu

[3]R.N. Smith is with Fort Lewis College, Durango, CO 81301, USA. rnsmith@fortlewis.edu

II. PROBLEM STATEMENT

Two robotic boats are deployed in a continuous two-dimensional area of interest $\mathcal{E} \subset \mathbb{R}^2$ with a user pre-defined boundaries. We assume that such an area is obstacle-free, as in many marine science expeditions. Both of them move via differential drive, are using GPS to localize, and can communicate continuously via a WiFi channel. As the mission evolves, the explorer selects a series of destination poses where to get more measurements and builds a more reliable model of the area, that is a map that has low uncertainty; at the same time, the sampler receives measurements from the explorer and uses this information to decide where to take a sample. The mission progresses up to the mission duration T_m, which generally depends on the specific logistics of the mission. All k units of the water sampling apparatus should be used in such a timeframe. This process leads to two related problems:

1) *Exploration*: explorer selects a sequence of poses $Q = \langle q_0, q_1, \ldots, q_n \rangle$, with $q_i \in \mathcal{E}$, so that the model of the area converges to the true phenomenon. Note that this process can be run online, and the explorer can take decisions as new measurements y_i associated with GPS locations $\mathbf{x}^i$ are collected. The efficiency is determined by traveled distance and quality of the map.
2) *Sampling*: based on all the measurements $\mathbf{Y}$, the sampler selects a number of locations $\mathcal{L}$, where to take physical samples, where $|\mathcal{L}| = k$ and $l \in \mathcal{L} \iff \exists y^i \in \mathbf{Y} | x^i = l$. The final objective is to maximize the sum of the values at sampled locations ($\sum_{l \in \mathcal{L}^*} f(l)$) within the maximum duration of the mission T_m.

Intuitively, the better the performance of the explorer, the better the performance of the sampler.

III. INFORMED STRATEGIC SAMPLING

The proposed system is based on using a variant of frontier-based exploration by the explorer, while a variant of the secretary hiring problem for the sampler.

A. Gaussian Process Frontier-based Exploration

Starting with zero knowledge about the spatial phenomenon in the given region, the explorer's objective is to select locations $L^* = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^m]$ over time such that the phenomenon is mapped efficiently. Note that while the robot is traveling to those locations, measurements $\mathbf{Y} = [y^1, y^2, \ldots, y^t]$ with associated GPS locations $\mathbf{X} = [\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^t]$ are collected at the frequency rate of the sensor. The goal is to optimize the time and the traveled distance to create a good model $\hat{f}(\mathbf{x})$ of the phenomenon $f(\mathbf{x})$.

With finite time and finite battery life of the robot, it is not feasible to take measurements at every location in the region of interest $\mathcal{E}$. Hence, we use Gaussian Processes (GP) [5] to model the spatial field. Mean and covariance functions should be formulated to completely define a GP. As done in the mainstream approach, mean is assumed to be zero, and a radial basis kernel is used as covariance function. Our exploration technique uses a one-step look ahead, where the robot decides on a set of locations to visit at epoch m only after reaching the chosen location of epoch $m-1$. We propose two methods to generate a list of locations (Figure 2). One of the approaches is to consider locations on the outer-most contour between a region with high variance and a region with low variance (Figure 2(a)). An easier method is to consider all the locations on a fixed planning window centered on the current position of the robot (Figure 2(b)). The location with highest predicted variance and least distance is chosen as the current target.

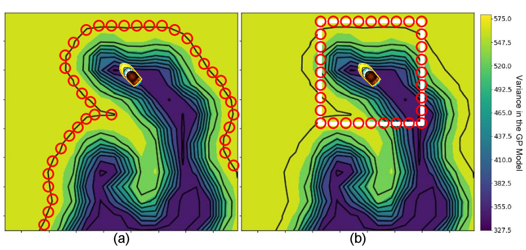

Fig. 2. Candidate locations generated by two techniques at a mission time step. Red circles represent the potential candidate locations l. Black lines show the contours.

B. Look-back Selective Sampling

As formalized in Section II, given M measurements i.e., candidate sampling positions, we need to choose k sample locations that optimize the quality of the final result. Since we are looking at simultaneous decision making along with the explorer, there is a need for optimal stopping criteria – in other words, when does the sampler decides to use one of the remaining water sampling units? This problem has similarities with the classic *Secretary Problem* that uses optimal stopping theory. Secretary Problem algorithm suggests we reject first $\frac{n}{e}$ candidates and then stopping at the first candidate with a higher ranking than all the ones evaluated until current time. Our problem is a variant of this problem as we need to choose k sample points instead of just one, hence we use a stopping threshold of $\frac{n}{ke}$. Unlike Secretary Problem, we have an advantage of looking back and choosing an old candidate if there is no better candidate location in the future. We want to maximize the sum of the values at sampled locations ($\sum_{l \in \mathcal{L}^*} f(l)$) with a minimum distance constraint (T_d) thus preventing acquisition of spatially neighboring samples. The value for T_d is application specific and also depends on the possible error in robot localization. We still need a stopping rule to make our decision.

IV. EXPERIMENTS AND DISCUSSION

We evaluated the system both in simulation and in the field on real robots. We have used three different setups to extensively evaluate the proposed system: 1) Simulated robots exploring and sampling from a synthetically created world, 2) Real world data (chlorophyll concentration in the flood plains of Amazon) used to create a world for simulated robots, and 3) deployment of two robotic boats in a reservoir to map the chlorophyll density distribution in the reservoir and collect water samples rich in chlorophyll content. Due to space constraint, we only present Figure 3 that illustrates the performance of the whole system, explorer and the sampler working together to achieve good sample quality.

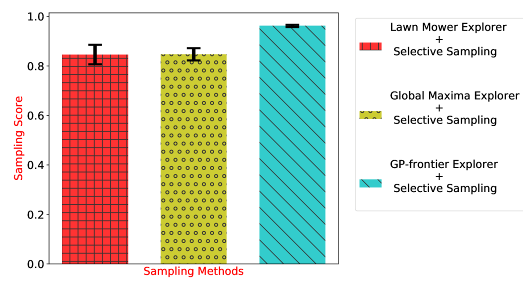

Fig. 3. Sampling scores achieved by the complete system, using different combinations of explorers and the look-back selective sampler.

We compared our proposed system – *GP-frontier Explorer + Look-back Selective Sampling* – to two other methods used in practice. The results show that the multirobot system with our proposed components performs well by achieving samples with high sampling scores.

With respect to future and ongoing work, we are scaling up the approach for application over larger regions in more challenging outdoor environments. The consideration of time-varying models will also be an interesting step towards more large-scale deployment in marine environments.

REFERENCES

[1] J. Das, F. Py, J. B. J. Harvey, J. P. Ryan, A. Gellene, R. Graham, D. A. Caron, K. Rajan, and G. S. Sukhatme, "Data-driven robotic sampling for marine ecosystem monitoring," *Int. J. Robot. Res.*, vol. 34, no. 12, pp. 1435–1452, 2015.

[2] B. Yamauchi, "Frontier-based exploration using multiple robots," in *Proc. International Conference on Autonomous Agents*, 1998, pp. 47–53.

[3] S. Manjanna, N. Kakodkar, M. Meghjani, and G. Dudek, "Efficient terrain driven coral coverage using gaussian processes for mosaic synthesis," in *Proc. CRV*. IEEE, 2016, pp. 448–455.

[4] K. H. Low, J. M. Dolan, and P. Khosla, "Adaptive multi-robot wide-area exploration and mapping," in *Proc. AAMAS*. International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 23–30.

[5] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.