An Information-Driven and Disturbance-Aware Planning Method for Long-Term Ocean Monitoring

Kai-Chieh Ma, Lantao Liu, Gaurav S. Sukhatme

***Abstract*— We propose an efficient path planning method for an autonomous underwater vehicle (AUV) used for the long-range and long-term ocean monitoring. We consider both the spatio-temporal variations of ocean phenomena and the disturbances caused by ocean currents, and design an approach integrating the information-theoretic and decision-theoretic planning frameworks. Specifically, the information-theoretic component employs a hierarchical structure and plans the most informative observation way-points for reducing the uncertainty of ocean phenomena modeling and prediction; whereas the decision-theoretic component plans local motions by taking into account the non-stationary ocean current disturbances. We validated the method through simulations with real ocean data.**

I. Introduction and Related Work

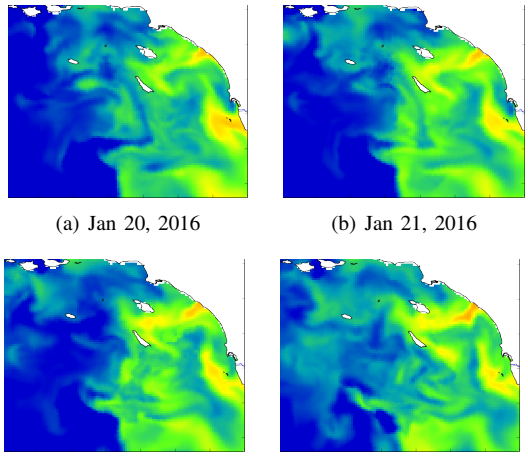
Ocean (environmental) monitoring and sensing allow scientists to gain a greater understanding of the planet and its environmental processes related to, e.g., physical, chemical or biological parameters [6]. We are interested in the problem of collecting data about a scalar field of important environmental attributes such as the temperature, salinity, and chlorophyll content of the ocean.

A key challenge of this research lies in the sensing, modeling, and predicting large-scale and spatially correlated environmental phenomena, especially when they are unknown and non-stationary [16]. Fig. 1 shows the variations of salinity data in the Southern California Bight region generated by the Regional Ocean Modeling Systems (ROMS) [20]. In practice, many sensing applications require continuous information gathering in order to provide a good estimate of the state of the environment at any time [13].

Traditionally, environmental monitoring is done by deploying static sensors which are used to collect data through the area of interest [15]. However, if the monitoring region is large, such a scheme is non-trivial to achieve due to a trade-off between the quantity of sensing resources and the quality of active sensing. Increasingly, a variety of autonomous robotic systems including marine vehicles [7], aerial vehicles [24], and ground vehicles [23], are designed and deployed for the environmental monitoring. Particularly, the autonomous underwater vehicles (AUVs) such as marine gliders are becoming popular due to their long-range (hundreds of kilometers) and long-term (weeks even months) monitoring capabilities [10, 14, 17].

We are interested in developing a path planning method that guides an AUV to collect ocean data in the most efficient way. By efficiency we mean the "informativeness" of collected data (i.e., reduction of phenomena modeling uncertainty) as well as the minimization of energy and time used to collect the data. Such a planning framework is also called *informative path planning* [13].

The authors are with the Department of Computer Science at the University of Southern California, Los Angeles, CA 90089, USA. {kaichiem, lantao.liu, gaurav}@usc.edu



(a) Jan 20, 2016 (b) Jan 21, 2016

(c) Jan 22, 2016 (d) Jan 23, 2016

Fig. 1. Ocean salinity data in the Southern California Bight region observed by the Regional Ocean Modeling Systems (ROMS) [20]. A brighter color indicates a larger value.

A variety of methodologies have been proposed to tackle the informative path planning problem, among which the most investigated approaches belong to the *nonmyopic* framework. Formally, the term *myopic* means that the path way-points are computed independently and greedily, without considering the cost and consequences of making observations in the future. Instead, a nonmyopic strategy performs optimization and computes a series of way-points by considering the effect of later time-steps [13]. Representative nonmyopic approaches include, for example, a *recursive-greedy* based algorithm [21] where the informativeness is generalized as a *submodular set function* built on which a sequential-allocation mechanism is designed in order to obtain subsequent way-points. This recursive-greedy framework has been extended by taking into account the avoidance of shipping lanes [2] and the diminshing returns [3]. Differing from above mechanisms where the path way-points are built by separate searching techniques with the informativeness as some utility function, Low [11] proposed a differential entropy based planning method in which a batch of way-points can be obtained through solving a dynamic program. Such a framework has been extended to approaches incorporating mutual information [5] and Markov-Information [12] optimization criteria. However,

these approaches are formulated with an assumption that the underlying map is in a regular shape (e.g., a skewed square) and the map is transected (sliced) column-wise, so that each algorithmic iteration computes way-points within a column and the navigation paths are obtained by connecting those way-points among the pairwise adjacent columns. In addition, there are also deterministic methods that optimize paths/tours under certain constraints (e.g., see [22, 25]). Deterministic optimization subject to path constraints is not a focus of this paper.

Our work falls into the nonmyopic informative planning category. We employ a Gaussian Process [19] to model an underlying phenomenon, and utilize the mutual information between visited locations and the remainder of the space to characterize the amount of information collected. Related to the practical ocean monitoring scenarios, we also consider the AUV's action uncertainty due to disturbances caused by the non-stationary ocean currents, and extend the Markov Decision Process (MDP) [18] in continuous space to control AUV's motion.

The paper makes the following contributions: we introduce a long-term autonomy planning method that integrates information-theoretic and decision-theoretic planning frameworks, so that both the path informativeness and action uncertainty are taken into account simultaneously; The method adapts to the spatio-temporal variations of both information (entropy) and disturbances. Formally, the high-level information-driven planner uses a hierarchical structure for characterizing hotspot regions and plans navigation way-points based on the maximization of mutual information; whereas the low-level disturbance-aware planner makes action decisions within short-term and short-range to "reject" the external disturbances; We validated the method through extensive simulations with real ocean data and show that the method not only maximizes information-gain but also saves time and energy while exploring the non-stationary ocean.

II. Preliminaries

A. Gaussian Process based Uncertain Field

To model spatial phenomena, a common approach in spatial statistics is to use a rich class of Gaussian Processes [16, 19, 21].

Formally, let W denote a set of sampling points describing the environmental phenomenon of interest. Each point $w \in W$ is a d-dimensional feature vector associating with either a realized measurement z_w if observed (sampled) or a random measurement Z_w if unobserved. Let set $\{Z_w\}_{w\in W}$ denote a GP, then for every finite subset of $\{Z_w\}_{w\in W}$, it has a multivariate Gaussian distribution. The GP can be fully specified by its mean $\mu_w \triangleq \mathbb{E}(Z_w)$ and covariance $\sigma_{ww'|\theta} \triangleq cov(Z_w, Z_{w'}|\theta)$ for all $w, w' \in W$, where θ parameterizes the covariance function which models the spatial phenomenon (parameterization details are presented later).

Assume we are given an observed data set $D = \{(w_i, Z_{w_i}), i = 1 : |D|\}$, where $D \subset W$. GP can be used to predict the mean and covariance of measurements for any unobserved subset of $U \subset W \setminus D$. Based on the property that every subset of $\{Z_w\}_{w\in W}$ is a multivariate Gaussian distribution, the joint distribution of Z_U and Z_D can therefore be expressed as:

$$\begin{pmatrix} Z_D \\ Z_U \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_D \\ \mu_U \end{pmatrix}, \begin{pmatrix} K_{DD} & K_{DU} \\ K_{DU}^T & K_{UU} \end{pmatrix}\right). \tag{1}$$

where

$$\begin{aligned} K_{DD} &= \left(\sigma_{ww'|\theta}\right)_{|D|\times|D|}, \quad \forall w, w' \in D, \\ K_{DU} &= \left(\sigma_{ww'|\theta}\right)_{|D|\times|U|}, \quad \forall w \in D, w' \in U, \\ K_{UU} &= \left(\sigma_{ww'|\theta}\right)_{|U|\times|U|}, \quad \forall w, w' \in U. \end{aligned} \tag{2}$$

We then obtain the Gaussian posterior mean and covariance,

$$\begin{aligned} \mu_{U|D,\theta} &= \mu(W_U) + K_{DU}^T K_{DD}^{-1}(Z_D - \mu(W_D)), \\ \Sigma_{U|D,\theta} &= K_{UU} - K_{DU}^T K_{DD}^{-1} K_{DU}, \end{aligned} \tag{3}$$

where W_U, W_D denote the set of sampling points in U, D respectively, and Z_D denotes the realized measurements of D. Note that the posterior covariance matrix $\Sigma_{U|D,\theta}$ is independent of the measurements and it can be used to assess the uncertainty with respect to the predicted measurements.

A GP's behavior is controlled via specifying its prior covariance (also known as *kernel*) $\sigma_{ww'|\theta}$, which describes the relation between sampling points w and w'. A widely adopted choice is the *squared exponential* kernel function:

$$\sigma_{ww'|\theta} = \sigma_s^2 \exp(-\frac{1}{2}(w - w')^T \Lambda^{-1}(w - w')) + \sigma_n^2 \delta_{ww'} \tag{4}$$

where $\Lambda = diag(l_1^2, \ldots, l_d^2)$ and $\theta = \{\sigma_s^2, \sigma_n^2, l_1^2, \ldots, l_d^2\}$ is the set of hyper-parameters specifying the property of the w, w' pairwise relation. The parameters $l_1 \ldots l_d$ are the *length-scales* in each dimension of w and determine the level of correlation between points (each l_i models the degree of smoothness in the spatial variation of the measurements in the ith dimension of the feature vector w). σ_s^2 and σ_n^2 denote the variances of the signal and noise, respectively. $\delta_{ww'}$ is the Kronecker delta function which is 1 if $w = w'$ and zero otherwise.

B. Entropy and Mutual Information

To assess the level of measurement and prediction uncertainty, we adopt the concept of entropy and mutual information. In information theory, the entropy is defined to quantify the uncertainty of random variables while the mutual information is used to describe the mutual dependence between two variables.

Formally, given a vector of sampling points A of size k, the joint *differential entropy* of the corresponding vector Z_A of random measurements is

$$\begin{aligned} H(Z_A) &= -\int p(Z_A) \log p(Z_A) d(Z_A) \\ &= \frac{1}{2} \log\left((2\pi e)^k |\Sigma_{AA}|\right). \end{aligned} \tag{5}$$

For arbitrary two vectors of sampling points A, B, the mutual information between A and B can be expressed in terms of (conditional) entropy

$$I(Z_A; Z_B) = I(Z_B; Z_A) = H(Z_A) - H(Z_A|Z_B) \quad (6)$$

where the conditional entropy $H(Z_A|Z_B)$ is

$$H(Z_A|Z_B) = \frac{1}{2}\log\Big((2\pi e)^k|\Sigma_{A|B}|\Big). \quad (7)$$

Because the field is modeled with a GP, the conditional covariance matrix $\Sigma_{A|B}$ can essentially be calculated from the posterior covariance matrix described in Eq. (3).

III. Technical Approach

Oftentimes the informative path planning starts with certain prior knowledge. For example, such knowledge may come from other AUVs that have already traversed the region or from remote sensing data [2]. With this prior knowledge, the planner computes a path for the AUV which gives us the most additional information (reducing predictive uncertainty). As time elapses, the uncertainty of visited/measured regions may increase again. The informative paths are thus repetitively generated based on the spatially and temporally varying ocean phenomenon.

A. Environment Representations & Methodology Framework

To represent the ocean environment, we discretize the environment into a grid map with certain resolution. Each grid at a location stores a mean value that predicts the phenomenon interested as well as a variance that measures the uncertainty of such prediction. We assume noise-free observation and when a location is sampled/observed, the variance of this grid is reduced to a very small value. In this work, we call a region with prediction uncertainty larger than certain threshold a *hotspot*. In many environmental processes, hotspots distribute non-uniformly and adjacent hotspots may be clustered as clouds, which are of particular interest to be explored.

We characterize the most informative regions by using a hierarchical structure, which is illustrated in Fig. 2. In greater detail, we start with a grid map with a very low resolution (i.e., the environment is tessellated into a few large regions), and apply the information-driven planner to get the first batch of observation points that are most informative. These observation points are too large to be used as actual path way-points, but they well characterize the hotspot regions with the maximum informativeness. We then recursively tessellate the hotspot regions and apply the information-driven planner to get new batches of observation points at finer resolutions. The process is repeated until the specified bottom layer of the hierarchy is reached.

Note that, the basic informativeness maximization procedure (details are presented later) only outputs batches of observation points, but does not convey any information of paths which are a sequence of ordered way-points. Therefore, we post-process these observation points with an existing Travelling Salesman Problem (TSP) solver to generate a meaningful paths and route the AUV from its initial location to visit all generated way-points with the minimized/shortest path length.

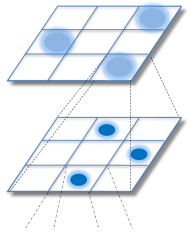


Fig. 2. Hierarchical planning framework consisting of multiple layers. A grid on a non-bottom layer can be expanded to a sub-map at a lower layer.

Lastly, the low-level motion planner takes into account the AUV's motion uncertainty caused by the ocean current disturbances. We map the discrete state space of MDP to continuous motion space and integrate the external disturbance into the stochastic transition model. The way-points generated from the information-driven planner are projected onto a fine grid map representing MDP's state space, and are used as local goal states for the purpose of local decision-making (navigation).

B. Hierarchical Information-Driven Planner

Given a desired total number of way-points n and a desired number of layers l, a series of sub-maps can be constructed hierarchically, as illustrated in Fig. 2. The target number of way-points in each sub-map is $n_l = \sqrt[l]{n}$. Let $W_k^{(l)}$ denote the whole sampling set of a sub-map at layer l with an index number k. The objective on each sub-map is to find a subset of sampling points, $P \subset W_k^{(l)}$ with a size $|P| \triangleq n_l$, which gives us the most information. It's equivalent to finding the maximum of the mutual information between the selected subset and the rest (unobserved part) of the sub-map. Thus the optimal P^* with maximal mutual information is

$$P^* = \arg\max_{P\in\mathcal{X}} I(Z_P; Z_{W_k^{(l)}\setminus P}) \quad (8)$$

where $\mathcal{X}$ represents all possible combinatorial sets each of which has a size n_l in the sub-map. In order to find P^*, one way is to enumerate all possible combinations in $\mathcal{X}$. The time complexity of this naive approach would be exponential and hence intractable in practice. Fortunately, P^* can be computed following a *dynamic programming* structure which dramatically reduces the computational time. Details are described below.

Let $w_i \in W_k^{(l)}$ denote an arbitrary sampling point at stage i and $w_{a:b}$ represent a sequence of sampling points from stage a to stage b. With Eq. (6), the mutual information between P and the unobserved part at the final stage n_l can then be written as $I(Z_{w_{1:n_l}}; Z_{W_k^{(l)}\setminus\{w_{1:n_l}\}})$. This mutual information can be expanded using the chain rule:

$$\begin{aligned} I(Z_{w_{1:n_l}}; Z_{W_k^{(l)}\setminus\{w_{1:n_l}\}}) &= I(Z_{w_1}; Z_{W_k^{(l)}\setminus\{w_{1:n_l}\}}) \\ &+ \sum_{i=2}^{n_l} I(Z_{w_i}; Z_{W_k^{(l)}\setminus\{w_{1:n_l}\}}|Z_{w_{1:i-1}}). \end{aligned} \quad (9)$$

One can utilize this form of mutual information to calculate w_i step by step, however, at every stage i before the final

stage, the entire unobserved set $W_k^{(l)} \setminus \{w_{1:n_l}\}$ is not known in advance, therefore we make an approximation

$$I(Z_{w_{1:n_l}}; Z_{W_k^{(l)} \setminus \{w_{1:n_l}\}}) \approx I(Z_{w_1}; Z_{W_k^{(l)} \setminus \{w_1\}}) + \sum_{i=2}^{n_l} I(Z_{w_i}; Z_{W_k^{(l)} \setminus \{w_1,\ldots,w_i\}} | Z_{w_{1:i-1}}), \tag{10}$$

which can be formulated with a recursive form, i.e., for stages $i = 2, \ldots, n_l$, the value $V_i(w_i)$ of w_i is:

$$V_i(w_i) = \max_{w_i \in W_k^{(l)} \setminus \{w_1,\ldots,w_{i-1}\}} I(Z_{w_i}; Z_{W_k^{(l)} \setminus \{w_1,\ldots,w_i\}} | Z_{w_{1:i-1}}) + V_{i-1}(w_{i-1}), \tag{11}$$

and the base case for this recursion is:

$$V_1(w_1) = I(Z_{w_1}; Z_{W_k^{(l)} \setminus \{w_1\}}). \tag{12}$$

Note the optimal way-point in the last stage n_l is

$$w_{n_l}^* = \arg\max_{w_{n_l} \in W_k^{(l)}} V_{n_l}(w_{n_l}). \tag{13}$$

With the optimal solution in the last stage, $w_{n_l}^*$, we can backtrace all optimal sampling points till the first stage w_1^*, and get the whole set of observation points $w_{(k)}^* = \{w_1^*, w_2^*, \ldots, w_{n_l}^*\}$ in sub-map k. If these points are not generated from the sub-map at the bottom layer, we recursively calculate finer observation points at each of their resided regions in $w_{(k)}^*$. Otherwise the union of all obtained observation points, $\bigcup\{w_{(k)}^*\}$, $\forall k$, are connected as a path/tour using any existing TSP solver based on their spatial descriptions. In this way, the traveling cost is minimized, and we show in the experimental section that the information gain is also better than the myopic greedy scheme.

The whole computational process for the hierarchical information-driven planner is pseudo-coded in Alg. 1.

Algorithm 1: Hierarchical Information-Driven Planner

```
Given the number of way-points n, and the number of layers l, calculate the number of way-points in each sub-map n_l = ∛n
/* from the top layer */
foreach w ∈ W_k^(l) do
    V_1(w) = I(Z_w; Z_{W_k^(l) \ {w}})
foreach i = 2 to n_l do
    foreach w ∈ W_k^(l) do
        initialize V_i(w) = −∞
    foreach w_{i−1} ∈ W_k^(l) do
        foreach w_i ∈ W_k^(l) \ {w_1, ..., w_{i−1}} do
            V_i(w) =
              max(I(Z_{w_i}; Z_{W_k^(l) \ {w_1,...,w_i}} | Z_{w_{1:i−1}}) +
              V_{i−1}(w_{i−1}), V_i(w))
w*_{n_l} = arg max_{w_{n_l} ∈ W_k^(l)} V_{n_l}(w_{n_l})
Backtrace to get w*_(k) ≜ {w*_1, w*_2, ..., w*_{n_l}}
if bottom layer has not been reached then
    expand the next layer, go to line 3
With the set of final way-points ⋃{w*_(k)}, calculate a path/tour with an existing TSP solution
```

C. Disturbance-Aware Motion Planner

One unique feature for the AUVs lies in their (extremely) uncertain motion outcomes. This is because the AUVs such as marine gliders normally have an operating speed of less than 0.5 m/s, which is at the same magnitude of the ocean current speed. In such a scenario, tracking a pre-planned trajectory is difficult, inefficient, or even impossible. Consequently, we design the low-level motion planner based on the decision-theoretic framework, and develop an adaptive disturbance-aware planning method built upon the MDP [1, 18]. Specifically, the aforementioned high-level information-driven planner produces a series of informative path way-points. By setting the succeeding way-points as the short-horizon goal states, the low-level disturbance-aware motion planner generates policies for the local guidance.

Formally, let an MDP be $\mathcal{M} =< S, A, T, R >$, where $S = \{s\}$ and $A = \{a\}$ represent the state space and action space, respectively. The stochastic transition model for an agent can be modelled with $T_a(s, s') = \Pr(s'|s, a)$, which is a probability mass function that leads the agent to future state s' when it executes the action a from s. The forth element $R_a(s, s')$ is a positive reward scalar for performing action a on s and reaching s'.

However, the standard model described above cannot be implemented directly in our ocean monitoring scenario. This is because: (1) MDP is built on discrete state space S, whereas the robot motion is continuous; and (2) MDP's transition model $T_a(s, s')$ is a function of multiple resources and can be time varying. We address these issues as follows.

First, in the continuous space, we re-define these variables:

- State $\mathbf{x}$ is the counterpart of s but in continuous state space. When $\mathbf{x}$ coincides at s, we denote such state as $\mathbf{x}(s)$ in continuous space. We can also map $\mathbf{x}$ back to discrete space: $\mathbf{x} \mapsto s$ if $||\mathbf{x} - s||$ is less than space partition resolution;
- Local control reference $\mathbf{a}(s)$ at s is a vector and it is mapped from MDP's action $a \in A$;
- Vector $\mathbf{d}(s)$ expresses an environmental/external disturbance at s.

Second, there are two components that contribute to the distribution of transition model $T_a(s, s')$. One is the control of the robot, and the other is the external disturbances caused by the ocean currents. Oftentimes the robot's action/control $\mathbf{a}(s)$ and external disturbance $\mathbf{d}(s)$ are addable (e.g., forces, velocities) and produce a resultant/net vector $\mathbf{r}(s) = \mathbf{a}(s) + \mathbf{d}(s)$ applied on the robot. We assume both the action and the disturbance contain noises subject to independent Gaussian distributions: $\mathbf{a}(s) \sim \mathcal{N}(\mu_a, \Sigma_a)$, $\mathbf{d}(s) \sim \mathcal{N}(\mu_d, \Sigma_d)$. Thus the robot's resultant state $\mathbf{x}$ after applying $\mathbf{a}(s)$ and being disturbed by $\mathbf{d}(s)$ also follows a Gaussian distribution:

$$p_a(\mathbf{x}) = \mathcal{N}(\mu_x, \Sigma_x), \tag{14}$$

where $\mu_x = \mu_a + \mu_d$ and $\Sigma_x = \Sigma_a + \Sigma_d$ are the mean and covariance of $\mathbf{x}$, respectively. It is worth mentioning that the MDP may produce multiple optimal actions (with equal optimal value) at some state. In such a case, a *mixture distribution* can be used,

$$p_{\{a_1,\cdots,a_k\}}(\mathbf{x}) = \sum_{i=1}^{k} \lambda_i p_{a_i}(\mathbf{x}) = \frac{1}{k}\sum_{i=1}^{k} p_{a_i}(\mathbf{x}), \qquad (15)$$

where the weighting parameter λ_i for component PDFs are identical as actions have the same optimal value. Let $\{\mathbf{a}^*(s)\}$ be the set of optimal actions at state s, the transition model thus can be expressed as

$$T_a(s, s') = \Pr\big(s' \mid s, \{\mathbf{a}^*(s)\}, \mathbf{d}(s)\big). \qquad (16)$$

In practice, such discrete probability mass function is approximated by integrating Eq. (14) over discretized grids.

The value $V(s)$ of a state s can be formulated as a recursive equation

$$V^*(s) = \max_{a\in A} \sum_{\forall s'\in S} T_a(s, s')\Big(R_a(s, s') + \gamma V^*(s')\Big), \quad (17)$$

where $\gamma \in (0, 1)$ is an infinite-horizon discount factor for discounting future costs at a geometric rate. From Eq. (17), the optimal action policy $\pi^*(s)$ can be obtained

$$\pi^*(s) = \arg\max_{a\in A} \sum_{\forall s'\in S} T_a(s, s')\Big(R_a(s, s') + \gamma V^*(s')\Big). \qquad (18)$$

Employing Bellman's principle of optimality avoids enumerating solutions naively, and we utilize the *value iteration* [1] to calculate the optimal policy.

Note that, the disturbances $\mathbf{d}(s)$ caused by the ocean currents can be time-varying. The AUV updates the estimate of $\mathbf{d}(s)$ periodically to catch up with the ocean dynamics. (For example, a marine glider takes advantage of the forecasts by the ROMS, and updates the ocean current information every time when it surfaces.)

D. Time Analysis

Assume the number of grids in each sub-map is K, the most expensive part for computing the differential mutual information lies in the computation of matrix inversion and determinant, which is approximately $O(K^{2.4})$ for a matrix of dimension K if state-of-the-art matrix manipulation algorithms are used. The dynamic programming process requires K^2 and we have a total of n_l way-points (equivalently n_l stages), thus the complexity for each sub-map is $O(n_l K^2) \times O(K^{2.4}) = O(\sqrt[l]{n}K^{4.4})$. The total number of recursive sub-maps is $\sqrt[l]{n}^0 + \sqrt[l]{n}^1 + \cdots + \sqrt[l]{n}^{l-1} = \frac{n-1}{\sqrt[l]{n}-1}$. Therefore, the total time complexity of our hierarchical information-driven planner is $O(\frac{(n-1)\sqrt[l]{n}K^{4.4}}{\sqrt[l]{n}-1})$. (Note that, the TSP is NP-complete but there are numerous approximation solutions with polynomial time complexities. Generally an efficient TSP approximation solver costs $O(n^3)$ [9].)

The time complexity for the disturbance-aware motion planner is bounded by $O(\tau|A||S||N_s|)$, where τ is the number of iterations, $|N_s|$ is the maximal number of states that a state can transit to, and $|A|$, $|S|$ are the numbers of MDP actions and states, respectively. Since each goal state is the succeeding way-point which is usually very close to the AUV's situated state, the value propagation horizon is thus short and τ can be approximated at the same magnitude of the propagation horizon (i.e., the number of propagation "hops" between the two states).

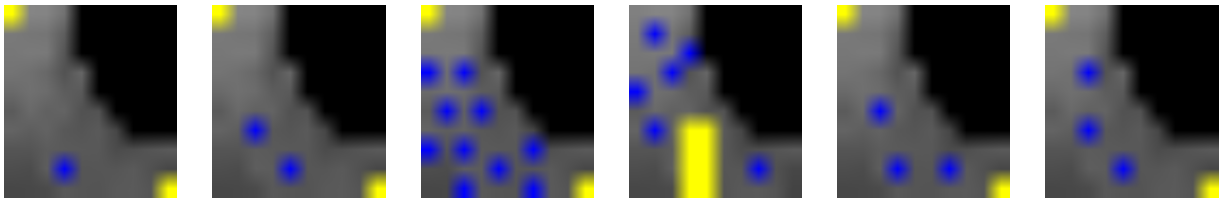
IV. Experimental Results

We validated our method in the scenario of ocean monitoring. A simulator written in C++ was built in order to test the proposed planning framework. The robot used in simulation is an underwater glider with a simplified kinematic model. The simulation environment was constructed as a two dimensional ocean surface and we tessellated the environment into grid maps. Our method applies for generic ocean phenomena. In our experiments, we use salinity and ocean current data observed in the Southern California Bight region. The raw data is obtained from ROMS [20].

For the hierarchical information-driven planner, we use the open-source library *libgp* [4] to model the latent phenomenon. Then we compute the observation points and the final sampling path following the procedure described in Alg. 1. For the low-level disturbance-aware planner, we represent the center of each grid as a state, where each non-boundary state has a total of nine actions, i.e., a non-boundary state can move in the directions of N, NE, E, SE, S, SW, W, NW, plus an idle action (returning to itself). Time varying ocean currents are external disturbances for the robot and are represented as a vector field. Specifically, vector $\hat{\mathbf{u}}(\mathbf{x})$ denotes the *easting* velocity component (along latitude axis) and vector $\hat{\mathbf{v}}(\mathbf{x})$ denotes the *northing* component (along longitude axis). The disturbance at $\mathbf{x}$ can thus be written as $\mathbf{d}(\mathbf{x}) = \hat{\mathbf{u}}(\mathbf{x}) + \hat{\mathbf{v}}(\mathbf{x})$. As mentioned earlier, a continuous state $\mathbf{x}$ can be mapped to a discrete tessellated state $s \in S$ within tessellation resolution.

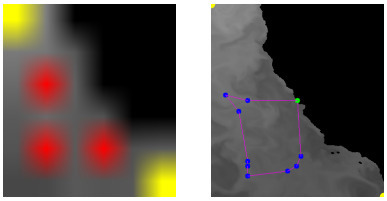
Fig. 3 demonstrates a set of observation points that maximize mutual information. For better illustration, the map contains only one layer and the layer is discretized into 10×10 grids. In the figures, the black regions represent lands while the gray areas denote oceans. The yellow dots and blue blobs represent prior sampling points and resultant observation points, respectively. Fig. 3(a)–3(d) show varying numbers of observation points with some prior/known samples that are set manually.

We compare our method with an existing popular methodology. Specifically, Guestrin *et al* [8] showed that, if the discretization of space is fine enough and the GP satisfies mild regularity conditions, the solution is near optimal for the myopic scheme where the observation points are computed independently and greedily based on immediate evaluations without considering the consequence of future path way-points. Following the evaluation design of prior works [5, 12], we thus compare the proposed method with the myopic greedy framework. Fig. 3(e) and 3(f) are observation points obtained from myopic greedy method (left) and the



(a) (b) (c) (d) (e) (f)

Fig. 3. Demonstration of observation points that maximize mutual information. (a)–(d) Varying numbers of observation points (blue blobs) with some prior/known samples (yellow blobs); (e)(f) Observation points from myopic greedy method (left) and the proposed method (right). A comparison reveals that our method generates observation points with a better field/exploration coverage.



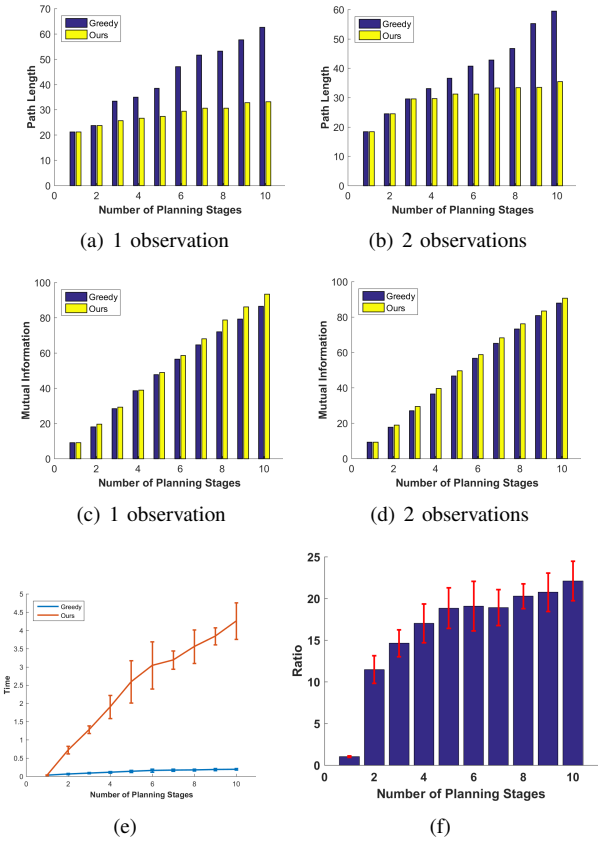
(a) (b)

Fig. 4. Demonstration of hierarchical planning results with a total of 9 way-points. (a) The top layer of way-points (red) with priors on the top-left and bottom-right corners. (b) A complete tour connecting the way-points (blue) generated from each sub-map at the bottom layer. The green dot on the shore represents AUV's starting location.

proposed method (right), respectively. A close comparison reveals that our method generates observation points with a better field/exploration coverage.

Information-driven planning results from a two-layer framework is shown in Fig. 4. In this example, we chose nine way-points and discretized the first layer into a 6×6 grid map. The red blobs are the observation points produced from the first layer, based on which the second layer generates a series of nine sampling points. By connecting with the AUV's starting location, a TSP tour is computed, as shown in Fig. 4(b).

Fig. 5 shows statistical comparisons between the myopic greedy approach and this proposed nonmyopic method. In the figures, the x-axis corresponds to the number of planning stages, which is equivalent to the number of way-points to be generated. Fig. 5(a) and 5(b) show that our approach generates paths with much shorter path lengths. We then compare the informativeness. Fig. 5(c) and 5(d) reveal the total mutual information gained after completing the whole paths. We can observe that our method is slighter superior to the greedy strategy and the advantage becomes more significant as the number of planning stages increases. Additionally, the computational time is also investigated. Fig. 5(e) and 5(f) show the running times as well as the ratio between the greedy and our approach. Both the running times grow linearly as the planning stages increase, and our method is more costly due to the optimization process using the dynamic programming. (All statistics are obtained after 20 trials, where each time we randomly select some prior observation locations.) Therefore, we can conclude that, with some cost of the computational resource, our method performs better than the myopic greedy method in terms of the information gain, and much superior from the perspective of saving traveling time and energy.



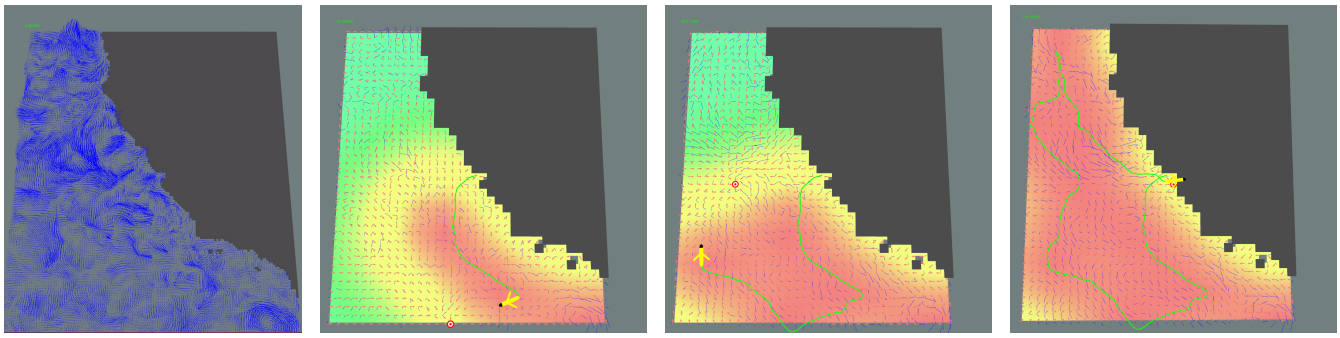
(a) 1 observation (b) 2 observations

(c) 1 observation (d) 2 observations

(e) (f)

Fig. 5. (a)-(d) Comparisons in terms of path length and mutual information given one (left) and two (right) prior observations; (e) Running times of the greedy and the proposed approaches. (f) The ratio of the running time of our approach against the greedy method.

Finally, we validated the proposed approach with real ocean current data. Fig. 6(a) depicts the raw ocean current data obtained from ROMS. The ocean current data is integrated into the disturbance-aware planning component to construct the MDP's stochastic transition model. Note that the ocean currents are non-stationary, thus the planner periodically updates the disturbance information at some fixed period. With the navigation way-points output from the hierarchical information-driven planner, the AUV follows local decisions represented by the MDP's optimal policy until it finishes the current batch of way-points. As the latent phenomenon is varying spatially and temporally, the planning procedure repeats and continues as a long-term process. A resultant trajectory of the AUV is illustrated in Fig. 6(b)–6(d). The colormap in these figures denote the information-



(a) (b) (c) (d)

Fig. 6. Disturbance-aware planning of an AUV under non-stationary ocean currents. (a) Raw ocean currents predicted by ROMS; (b)–(d) the AUV follows a series of way-points which cover the most uncertainty regions. The colormap depicts the variance of phenomenon prediction, where a warmer color represents a smaller variance. Blue lines are ocean currents and read arrows denote the MDP action policy.

gain (informativeness), from which we can see that the proposed method produces informative path that explores and covers most of uncertain regions.

V. Conclusions

In this paper, we presented an informative path planning method for the long-term AUV ocean monitoring. The method takes into account both the spatio-temporal variations of ocean phenomena and the disturbances caused by the ocean currents, and integrates components from both information-theoretic and decision-theoretic frameworks. Specifically, the information-theoretic component employs a hierarchical structure and plans the most informative observation way-points; whereas the decision-theoretic component plans local motions by taking into account the non-stationary ocean current disturbances. Our simulation results show that the proposed approach is superior to the myopic schemes in terms of both the information gains and the vehicle's energy cost.

References

[1] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition, 2000.

[2] J. Binney, A. Krause, and G. S. Sukhatme. Informative path planning for an autonomous underwater vehicle. In *International Conference on Robotics and Automation*, pages 4791–4796, 2010.

[3] J. Binney, A. Krause, and G. S. Sukhatme. Optimizing waypoints for monitoring spatiotemporal phenomena. *International Journal on Robotics Research (IJRR)*, 32(8):873–888, 2013.

[4] M. Blum. libgp. `https://github.com/mblum/libgp`.

[5] N. Cao, K. H. Low, and J. M. Dolan. Multi-robot informative path planning for active sensing of environmental phenomena: A tale of two algorithms. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, pages 7–14, 2013.

[6] M. Dunbabin and L. Marques. Robotics for environmental monitoring: Significant advancements and applications. *IEEE Robot. Autom. Mag.*, 19(1):24 – 39, 2012.

[7] E. Fiorelli, P. Bhatta, and N. E. Leonard. Adaptive sampling using feedback control of an autonomous underwater glider fleet. In *Proc. 13th Int. Symposium on Unmanned Untethered Submersible Tech*, pages 1–16, 2003.

[8] C. Guestrin, A. Krause, and A. Singh. Near-optimal sensor placements in Gaussian processes. In *International Conference on Machine Learning (ICML)*, August 2005.

[9] G. Laporte. The traveling salesman problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59(2):231 – 247, 1992.

[10] N. E. Leonard, D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien, and F. Zhang. Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in monterey bay. *Journal of Field Robotics*, 27(6):718–740, 2010.

[11] K. H. Low. *Multi-robot Adaptive Exploration and Mapping for Environmental Sensing Applications*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2009.

[12] K. H. Low, J. M. Dolan, and P. Khosla. Active markov information-theoretic path planning for robotic environmental sensing. In *Proceedings of the 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS-11)*, pages 753–760, May 2011.

[13] A. Meliou, A. Krause, C. Guestrin, and J. M. Hellerstein. Nonmyopic informative path planning in spatio-temporal models. In *Proceedings of National Conference on Artificial Intelligence (AAAI)*, pages 602–607, 2007.

[14] T. Miles, S. H. Lee, A. Whlin, H. K. Ha, T. W. Kim, K. M. Assmann, and O. Schofield. Glider observations of the dotson ice shelf outflow. *Deep Sea Research Part II: Topical Studies in Oceanography*, 2015.

[15] L. M. Oliveira and J. J. Rodrigues. Wireless sensor networks: a survey on environmental monitoring. *Journal of Communications*, 6:143–151, 2011.

[16] R. Ouyang, K. H. Low, J. Chen, and P. Jaillet. Multi-robot active sensing of non-stationary gaussian process-based environmental phenomena. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, pages 573–580, 2014.

[17] D. A. Paley, F. Zhang, D. M. Fratantoni, and N. E. Leonard. Glider control for ocean sampling: The glider coordinated control system. *IEEE Transaction on Control System Technology*, 16(4):735–744, 2008.

[18] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1st edition, 1994.

[19] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

[20] A. F. Shchepetkin and J. C. McWilliams. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, 9(4):347–404, 2005.

[21] A. Singh, A. Krause, C. Guestrin, W. Kaiser, and M. Batalin. Efficient planning of informative paths for multiple robots. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, IJCAI'07, pages 2204–2211, 2007.

[22] D. E. Soltero, M. Schwager, and D. Rus. Generating informative paths for persistent sensing in unknown environments. In *IROS*, pages 2172–2179, 2012.

[23] M. Trincavelli, M. Reggente, S. Coradeschi, A. Loutfi, H. Ishida, and A. J. Lilienthal. Towards environmental monitoring with mobile robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2210–2215, 2008.

[24] A. C. Watts, V. G. Ambrosia, and E. A. Hinkley. Unmanned Aircraft Systems in Remote Sensing and Scientific Research: Classification and Considerations of Use. *Remote Sensing*, 4(6):1671–1692, June 2012.

[25] J. Yu, M. Schwager, and D. Rus. Correlated orienteering problem and its application to informative path planning for persistent monitoring tasks. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014.